

Szkoła Główna Gospodarstwa Wiejskiego w Warszawie

Katedra Ekonometrii i Statystyki

**METODY ILOŚCIOWE
W BADANIACH
EKONOMICZNYCH**

Nr X (2009)

Redakcja naukowa:
Dorota Witkowska

Wydawnictwo SGGW
Warszawa 2009

Each human activity requires decision making and it is connected with risk. Therefore the methodology supporting decision processes and risk control has been developed in many fields including economics. This methodology bases on quantitative methods that let us to employ mathematical theories and models to describe the objects and phenomena or solve the problems regarding socio – economic matters.

The main problem that appears in socio - economic research is the measurement of the properties that characterize the analyzed objects, phenomena and relations since many of them cannot be directly measured or even observed. The process of measurement is central to quantitative research because it provides the fundamental connection between empirical observation and mathematical expression of quantitative relationships. Another important question is to apply the method that is adequate to:

- the aim of investigation,
- the problem that is to be solved, and
- the type of measurements that is used.

The minimization of risk and decision optimization need proper information. According to the progress in computer science and globalization process, huge data bases are available for decision-makers. That requires the development of adequate tools to process the data and to extract from them necessary information during the period of time that is short enough to choose the right decision. Therefore it is extremely important to explore issues and discuss problems of application statistic, econometric and other mathematical methods to solve properly defined problems in economics and finance, considering micro, mezzo and macro scales.

The papers that are presented in this issue of *Metody Ilościowe w Badaniach Ekonomicznych* discuss both theoretical questions and application of already known methodology to solve problems from practice.

Dorota Witkowska

Spis treści

1. Władysław Welfe, <i>Long-term macroeconomic model for the Polish economy</i>	5
2. Agnieszka Bezat, <i>Comparison of the deterministic and stochastic approaches for estimating technical efficiency on the example of non-parametric DEA and parametric SFA methods</i>	20
3. Agata Binderman, <i>Zależność oceny zróżnicowania rolnictwa w Polsce od wybranych mierników syntetycznych</i>	30
4. Mariusz M. Borawski, <i>The analysis of unemployment variation measures in the EU Member States</i>	42
5. Jadwiga Bożek, Karol Kukula, Danuta Bogocz, <i>Taxonomic analysis of areal structure of farms in Poland within 2000 – 2007</i>	50
6. Barbara Dańska-Borsiak, <i>Szacowanie i modelowanie TFP w przemyśle polskim na podstawie danych panelowych</i>	58
7. Małgorzata Doman, Ryszard Doman, <i>Forecasting the end-of-the-day realized variance</i>	67
8. Marek Gruszczyński, <i>Quantitative methods in accounting research</i>	76
9. Mariusz Hamulczuk, Katarzyna Hertel, <i>AGMEMOD model - structure and application for analysis and simulation of Polish agricultural sector</i>	88
10. Władysław Hoffmann, <i>Metoda AHP w ocenie stosowalności metod oceny zdolności kredytowej w systemie ekspertowym</i>	99
11. Stanisław K. Jabłonowski, Andrzej Kluza, <i>The multinomial logit models' usefulness to examine the relationship between agricultural enterprises' ownership form and their economic results</i>	107
12. Alina Jędrzejczak, Jan Kubacki, <i>Application of EBLUP estimation to the analysis of small areas on the basis of polish household budget survey</i>	121
13. Tomasz M. Kolanowski, <i>Predictive marketing mix modelling in FMCG packaged food category including price and advertising impact</i>	131
14. Witold Luciński <i>Analysis of banking performance in Polish commercial banks. Symptoms of threats</i>	142

15. Małgorzata Łatuszyńska, *Simulation investigation of consequences following from transport decisions* 153
16. Agnieszka A. Majewska, Sebastian M. Majewski, *Results of mistaken time period in analysis in the case of framing effect for some capital market's models* 163
17. Natalia Nehrebecka, Sylwia D. Grudkowska, *Application of epsilon method to modeling expectations in construction* 176
18. Kesra Nermend, *Normowanie zmiennych zawierających obiekty nietypowe* 185
19. Jacek Nowak, *Is multiple linear regression the proper tool of modelling a behaviour of real systems?* 194
20. Jitka Poměnková, *Analysis of economic activity movements in the Czech Republic – frequency approach* 207
21. Wacława Starzyńska, Justyna Wiktorowicz, *Application of selected statistical methods to ICT used in public procurement system in Poland* 221
22. Paweł Strawiński, *Łączenie danych z dynamicznym obciążeniem. Wyniki wstępne* 232
23. Tadeusz Waściński, Grzegorz Przekota, Ludwik Sobczak, *Koniunktura na polskim rynku kapitałowym wobec sytuacji na rynkach wiodących* 243
24. Antoni W. Wiliński, *Badanie efektywności predykcyjnej modelu klasy GMDH opartego na behawioryzmie uczestników rynku walutowego* 252
25. Dorota Witkowska, Krzysztof Kompa, Marta Grabska, *Badanie informacyjnej efektywności rynku w formie silnej na przykładzie wybranych funduszy inwestycyjnych* 265
26. Wojciech Zieliński, *Uogólniona miara dopasowania w modelu liniowym* 286

LONG-TERM MACROECONOMETRIC MODEL FOR THE POLISH ECONOMY

Władysław Welfe

Katedra Modeli i Prognoz Ekonometrycznych, Uniwersytet Łódzki
e-mail: emfiws@uni.lodz.pl

Abstract: Development of market economies towards economies based on knowledge implies the construction of new long-term macroeconomic models. The structure and use of the long-term W8D model for Poland was shown. This is a complete model that explains production and the factors its growth including R&D and human capital and also final demand allowing for studies of potential disequilibria. It is closed by equations explaining prices, wages and financial flows. The simulation version includes 235 equations, of which 111 are stochastic. It is used in long-term growth scenarios up to the year 2030.

Keywords: Knowledge Capital, R&D, Human Capital, Macromodelling, Multipliers, Scenarios
JEL C51, C52, E17, E22, O11, O40, O47

INTRODUCTION

Over the last years the economics community agreed that the contemporary market economies tend towards knowledge-based economies. This concept has been formulated in contrast to an industrial economy system that prevailed in the last centuries (Smith [2002]).

There is vast literature based on theories of endogenous growth aimed to explain – at the world-wide level – the differences in the rates of growth of particular countries. It is based on an analysis of international cross-section data. Its non-technical excellent summary can be found in Helpman [2004]. The applications of the results of this research to studies of economic growth of single economies are rather exceptional, except for the US economy. (see Jorgenson et al. [2003] and Richards [2000]) and for Polish economy (see Welfe, ed. [2001], [2009a]). It has a practical aspect: the authorities and the scientific community of

a country need to have an instrument that will help to construct scenarios of long-term economic growth for 20–30 years ahead.

Nearly all these studies were concentrated on the supply side analyzing factors of growth and the resulting potential output. It may, however, considerably diverge from effective output which represents realizations of final demand, underlying business cycle fluctuations. Hence, to be used in empirical analysis and simulation exercises we need to construct a complete model that contains both the final demand and the total supplies. That makes possible to estimate the likely disequilibria: output gap, unemployment, foreign trade deficits etc. The model should be closed by introduction price, wage and financial flows sector (W. Welfe [2008a]).

We tried to show the suggested structure of such models using as an example the annual long-term macroeconomic models W8D built for the Polish economy. Their characteristic is provided in the following section. Next sections contain discussions of properties of alternative measures of major determinants of economic growth, the alternative approaches to their explanation showing the interdependencies within the whole economic system¹. The role of investment in fixed capital and knowledge capital is discussed in the light of multiplier analysis. The applications in scenario analyses based on model simulations are provided at the end of the paper.

THE MACROECONOMETRIC MODELS OF A KNOWLEDGE-BASED ECONOMY

The quantitative mechanisms that underlie the growth of a knowledge-based economy can be described empirically by means of adequately expanded macroeconomic models (see: Garratt et al [2006]). Such models should draw on economic growth theory which has been enjoying its renaissance, and especially on the endogenous growth theory (see Grossman, Helpman [1991], Barro, Sala-i-Martin [1995], Aghion, Howitt [1999], and more recently Nahuis [2003], Tokarski, [2001, 2007]).

The long-term macroeconomic models built along these lines, extended to include processes in which knowledge capital is generated and used, seem to be the most relevant tools of long-term economic analysis. Their structure may follow the framework of the mainstream models outlined by Klein et al. [1999].

These models specify the final demand equations along the neo-Keynesian lines, but the potential output and demand for the factors of production, as well as impacts of technological progress they generate referring to the neoclassical theory of production (Solow, [1957]). This approach draws on the early theories of growth developed by Harrod and Domar and on the concept of models of production

¹ For a comprehensive discussion see Welfe, ed. [2007].

possibility frontier that have recently been developed by Jorgenson [2000]. The stylized empirical model of growth by W. Welfe [2005] follows a similar approach.

Empirical investigations in Poland referring to the above developments draw on the concept of an empirical model of economic growth developed by Welfe [2000]. This concept gave rise to the building of the long-term macroeconometric models of the Polish economy W8D (see Welfe ed. [2001]), [2004] and recently W8D 2007 [2009a).

The new model is a complete structure. Its quantitative description is shown in table 1. It is one sectoral, medium size model. Its simulation version comprises several blocks of equations, traditionally following the familiar classification of economic activities. The blocks explain: a) final demand, including exports and imports, b) the supply side, including potential output, and the primary factors of production, c) impact of technological progress (TFP), d) prices, wages and financial flows.

Many well known economic interrelationships are established within and between the above blocks. The role of investment as a factor determining an increase in potential output as well as in final demand is especially emphasized. It offers the possibility of studying potential business cycle fluctuations.

Table 1. Major characteristics of the model W8D-2007

Characteristics	The number of variables/equations	Characteristics	The number of variables/equations
Variables total:	393	Endogenous variables:	
– excluding dummies	258	– presimultaneous	24
– exogenous (E)	157	– jointly determined	80
– excluding dummies	22	– post simultaneous	131
– endogenous	235	Equations by blocks:	
Equations – total	235	– final demand–total	64
– stochastic (B)	111	– production factors	17
– identities (I)	124	– technical progress	40
Lags, leads		– potential output	7
– maximal lag	8	– average wages and incomes, prices and deflators	48
– lags total (L)	165	– financial flows	48
Feedback variables:	7	– macrocharacteristics	11

Source: own elaboration

Hence, we shall constrain the further discussion to the specification of investment and production functions playing the central role in the functioning of the model. For the specification of remaining equations see W. Welfe [2008b].

SPECIFICATION OF INVESTMENT FUNCTION

The investment function explaining demand for investment goods covers the investment in fixed capital. The specification of private enterprises' demand for investment goods starts with an accelerator. To put it simply, producer capacities expand in the long run following expected increase in the demand for products the producers can provide. This increase creates, allowing for restitution demand the potential demand for investment goods. It has to be adjusted for a likely change of the level of utilization of the available equipment. The perception of investment risks, changes in the profitability of investment projects and substitution between labour and capital has also an effect on the effective demand.

The expected future output is typically represented by output generated in the past. As for machinery and equipment, the Koyck transformation leading to a reduced form of the investment demand function can be used, where the explanatory variables are being confined to lagged investment (J_{t-1}) and current output X_t (see W. Welfe, A. Welfe [2004]). The rate of capacity utilization (WX_t) may be obtained using several approximations. The investment risk RJ_t can be approximated using various indicators, such as government deficit, government debt service, or a rate of inflation. The profitability of investment projects is usually expressed by a ratio of producer prices (PX_t) and user costs KI_t , the latter predominantly depending on the long-term interest rate RL_t . The rate of wages (WBP_t) and investment deflator (PJ_t) stands for the effects of substitution. Hence, a typical specification of the investment demand function is as follows:

$$J_t = AJ_{t-1}^{\alpha_1} X_t^{\alpha_2} WX_t^{\alpha_3} RJ_t^{\alpha_4} (PX_t / KI_t)^{\alpha_5} (WBP_t / PJ_t)^{\alpha_6} e^{\varepsilon_t} \quad (1)$$

where: $KI_t = PJ_t(RL_t + \delta)$, PJ_t is an investment deflator, RL_t is real long-term interest rate; WX_t is capacity utilisation rate, δ is the rate of depreciation.

In the W8D model the following elasticities were obtained: with respect to the output – long-run 1, but short-run 2, and with respect to the capacity utilization: long run 1.1.

THE EXTENDED PRODUCTION FUNCTION

Analysis of the production process implies the use of the production function. Following many authors, we recommend the double-log production function, i.e. a Cobb-Douglas function with constant returns to scale in its extended version:

$$X_t^P = BA_t K_t^\alpha N_t^{(1-\alpha)} e^{\varepsilon_t} \quad (2)$$

where: X_t^P is potential output (GDP at the macroscale) in constant prices, A_t is total factor productivity (TFP), K_t is fixed capital, in constant prices, N_t is employment.

Estimating the function's parameters for a market economy is not a straightforward exercise, because observations represent effective output, i.e. the realization of the final demand X_t , and not potential output, and most frequently $X_t < X_t^P$. Therefore, changes in the rate of utilization of production potential must be addressed. In most cases, they originate from changes in final demand, i.e. from the business cycle fluctuations. The rate of capacity utilization is defined as

$$WX_t = X_t / X_t^P \quad (3)$$

Allowing for the rate of utilization, we can redefine the production function (2) into:

$$X_t = BWX_t A_t K_t^\alpha N_t^{(1-\alpha)} e^{\varepsilon_t} \quad (4)$$

To estimate the parameters of the above function, we need information on the rate of capacity utilization and total factor productivity. Several methods of estimating the rate of utilization are used: industrial surveys that ask direct questions about the level of utilization of machinery and equipment, as well as time worked (Grzęda-Latocha [2005]). Central banks use techniques involving analysis of the deviations from a GDP trend. Other procedures use the decomposition of the rate of utilization of fixed capital and employment (see W. Welfe [1992]).

THE DYNAMICS OF THE TOTAL FACTOR PRODUCTIVITY

Following the concept of Solow residual (Solow [1957]), total factor productivity is commonly used to represent the effects of knowledge capital absorption and widely applied to international comparisons. However, some measurement problems that constrain its use have not been solved yet (W. Welfe, [2002, 2007], J. Cornwall, W. Cornwall, [2002]).

Firstly, TFP is computed using effective (i.e. observable) output and not potential output. Secondly, several studies, mainly sectoral, use the concept of gross output instead of value added. Further, output elasticities with respect to fixed capital are frequently calibrated and not estimated. All these issues are discussed in the literature; their summary can be found in W. Welfe [2002].

The tradition of modelling the TFP growth factors is quite short and many questions still await their answers. A frequently suggested way of explaining TFP dynamics is a decomposition of TFP changes into the effects of the free available

capital of knowledge (\dot{A}_t^W), the impacts of expanding knowledge capital embodied in fixed capital (A_t^K) and increasing human capital (\dot{A}_t^N). Taking the production function (2), we have:

$$\dot{A}_t = \dot{A}_t^W + \alpha \dot{A}_t^K + (1 - \alpha) \dot{A}_t^N, \quad (5)$$

where ($\dot{}$) is the rate of growth

The effects of generally available knowledge capital (A_t^W) are either treated as exogenous (usually as an exponential function of time) or attributed to the growth of knowledge capital associated with improving quality of employment.

In the past the effects of expanding knowledge capital embodied in fixed capital were treated as functions of time. Recently, following the mass computerization of production and management processes, decomposition of fixed capital into computers, computer programs and tele-equipment has been suggested. This approach was used in studies exploring the growth of the US economy (Jorgenson, [2001], Jorgenson, Ho, Stiroh, [2003]), and then in research on the OECD countries (Colecchia, Schreyer, [2002]) and the Netherlands (Leeuwen, Wiel, [2003]). It means, however, the removal of computerization effects from the notion of TFP.

When the starting point is decomposition as given by (5), the impact of TFP embodied in fixed assets (A_t^K) can be related, to the anticipated effects of past R&D expenditures. The cumulative R&D expenditures (R&D capital)– both domestic (BR_t^K) and foreign, transferred from abroad (BR_t^M) – are assumed to represent the capital of technical and organizational knowledge (Coe, Helpman, [1995]).

The direct and indirect channels through which R&D is transferred from abroad are distinguished. The expanding systems of telephone lines, closer technological proximity, more frequent use of patents, etc., stimulate the direct transfer of R&D (Lee, [2005]), whereas imports of commodities, represent its indirect transfer (Xu, Wang, [1999]). The indirect transfer of knowledge can be summarized by computing the weighted sum of R&D expenditures incurred by the distinguished country j :

$$BR_t^M = \sum_j w_j BR_j^K, \quad (6)$$

where w_j represents a weight assigned to R&D expenditures of the country j .

The weights can stand for particular countries' shares in the total imports of the analyzed country, or rather for the ratios of imports from these countries (i.e. their exports) to their GDPs (Lichtenberg, van Pottelsberghe [1998]). The weights can be linked with the imports of intermediate goods (transfer of technology), with

the imports of investment goods (transfer of new machines, etc.), which seems a better approach (Bayoumi et al., [1999]), or with total imports, when their decomposition is not possible. This approach was used in earlier versions of W8D models for the Polish economy. However, in the most recent version of the model the imports were decomposed into groups of commodities that differ by the level of technological advancement (W. Welfe [2008b]).

In the last 15 years, the discussed research has been given international dimension. Its scope includes now not only the industrialized countries, but also the developing ones (Engelbrecht, [1997], Bayoumi et al. [1999]). However, the role of FDI in stimulating TFP dynamics continuous to be debatable. It seems also necessary to explore further the impact of domestic R&D expenditures and human capital on the rate of absorption of the transferred foreign capital of knowledge (Cincera, Pottelsberghe [2001]).

The above relationships are multiplicative. Hence, the first approximation we can write is:

$$\ln A_t^K = \beta_o \ln BR_t^K + \beta_i \gamma \ln BR_t^M, \quad (7)$$

where γ is a weight representing the share of imports, i.e. the degree of openness of the economy.

In the model W8D 2007 for Poland the elasticities w.r. to domestic R&D capital were close to 0,3, which is in line with the results reported by other authors for industrialized countries, whereas w.r. to foreign R&D capital they exceeded 0,6, which was found characteristic for less developed countries.

EFFECTS OF INCREASING HUMAN CAPITAL

Earlier investigations evaluating the impacts of expanding human capital on economic growth produced inconclusive results. The main reason was that they used inconsistent data on the schooling years.

In general, the scope of human capital varies in terms of its coverage. The narrow definition, most frequently used in empirical research, accentuates the differences between the levels of employees' education. The broad version allows for the impacts of learning by doing, health status, etc. (see Benabou [2002]). In either case, the measurement problems need an adequate solution. Unfortunately, many international projects use simplified measures of human capital, i.e. shares of employees with tertiary education, or with secondary and tertiary education etc. This situation has improved in the last years, as more adequate information on the number of schooling years has become available (Fuente, [2004]). Notwithstanding, only few researchers take advantage of the newly developed summary characteristics of human capital per employee.

These characteristics of human capital (H_t) are designed as the weighted sums of employees with different educational levels i :

$$H_i = \sum_i \mu_i N_{ii}, \quad (8)$$

where i is the level of education and μ_i is a weight.

The weights may represent:

a) an average number of schooling years, b) average unit costs of schooling; c) average wages per employee with educational level i . In most cases, the weights represent the number of schooling years. However, the third approach, where the weights reflect the market efficiency of the level of education, seems to be the most appropriate, treating human capital as a factor of production (W. Welfe et. al. [2002]). It was used in all W8D models of the Polish economy. It is worthwhile to note that the elasticity of TFP w.r. to human capital per employee exceeded 1.

Human capital per employee is obtained by dividing total human capital by the total number of employees:

$$h_i = H_i / N_i, \quad (9)$$

Seeking relationships between investments in human capital and educational expenditure is a difficult task. To make this search effective, a submodel describing the educational process and related expenditures has been built (see Welfe et. al. [2002]).

THE MULTIPLIERS AND SCENARIOS

The simulation version of the model was used to support the medium and long-term forecasting and scenario analyses extended to the year 2030. The properties of the model were analysed by means of multiplier analysis that revealed the major economic mechanisms. We shall concentrate here on ex-ante multipliers analysing the impact of 10% shocks in one most important

variable only: investment in fixed capital being due to foreign financing (say EU transfers, FDI) (Fig. 2.). The figures will be shown respectively for the major macro variables: household consumption (C), investment in fixed capital (J), exports (E), imports (M), GDP (X), potential GDP (XNK)².

We constructed a long-term forecast to the year 2030 first. The multipliers were calculated by means of models' simulation, using this baseline forecast. Impuls and sustained multipliers were computed.

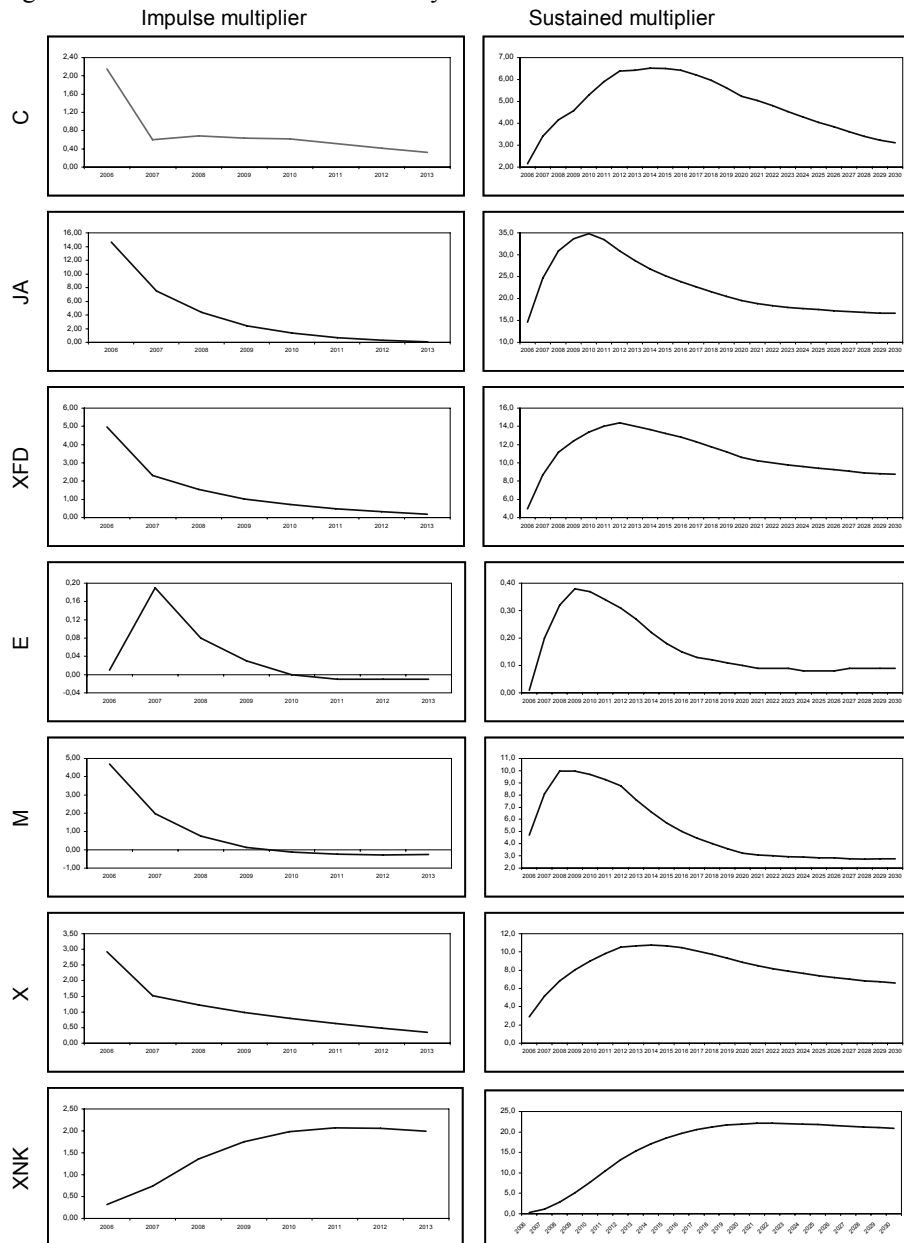
The impuls multipliers show a decline, which approaches zero in 5-8 years, except for potential output that starts declining in 5-6 years. The sustained multipliers present an interesting picture. Initially, because of the accelerator the investments grow from 15% to 35% after 6 years.

Then they decline because of declining capacity utilization and stabilize around 20%. Consumption follows this pattern with a considerable delay. As imports substantially grow the GDP increase is lagged attaining its maximum in 10

² For the detailed information see Welfe ed. [2009b].

years. For the midpoint of the period the elasticity of GDP w.r. to investment is close to 1. Notice that the potential GDP increases with a longer lag, but stabilizes at nearly 20%.

Fig. 1. 10% increase in investment outlays



Source: own investigation

Hence, there is a long-run tendency for a decline in capacity utilization, which negatively affects the investment growth. However, it would not be justified to conclude that the model predicts overinvestment, as we did not allow for any relevant increase in exports.

Table 2. Assumptions of for the scenarios of economic development up to the year 2030.

Macrovariables Scenarios	2010	2015	2020	2025	2030
Investment GDP rate In % c.p.					
optimistic	26.5	35.0	38.0	35.0	35.0
forecast	23.9	33.4	32.9	31.3	31.0
pessimistic	22.0	20.0	18.0	18.0	18.0
FDI GDP rate In % c.p.					
optimistic	4.2	4.0	4.0	3.5	3.0
forecast	3.1	2.8	2.1	1.5	1.1
pessimistic	0.4	1.4	1.3	0.8	0.7
Transfers net from UE in % GDP c.p.					
optimistic	3.5	3.2	3.0	2.8	2.8
forecast	3.5	3.2	3.0	2.8	2.8
pessimistic	3.5	3.2	3.0	2.8	2.8
Exports SNA rate of growth in %					
optimistic	10.0	12.0	9.0	9.0	9.0
forecast	7.5	11.2	5.2	7.3	4.0
pessimistic	5.0	5.0	4.5	4.0	4.0
Expenditures on education GDP ratio in % c.p.					
optimistic	5.0	4.8	4.7	4.7	4.8
forecast	4.5	4.3	4.2	4.0	4.0
pessimistic	3.8	3.5	3.1	2.7	2.5
R&D expenditures GDP ratio in % c.p.					
optimistic	1.00	2.00	2.60	2.80	3.2
forecast	0.79	1.15	2.00	2.00	2.00
pessimistic	0.76	0.75	0.65	0.60	0.50
Increase of elasticity of absorbtion of foreign R&D					
optimistic	40	60	80	100	100
forecast	20	30	40	50	60
pessimistic	0	0	0	0	0

Source: own investigation

The multiplier analysis helped to launch a new series of long-term forecasts and scenario analyses underpinned by the model-based simulations. We distinguished two scenarios only. In the optimistic variant we assumed high long-run increase in fixed capital and in domestic R&D expenditures, better absorption of transferred foreign expenditures on R&D, and growth of human capital stimulated by larger financial allocations to the tertiary and post-graduate

education. In the pessimistic variant- the deterioration of these factors was assumed. The specific assumptions are shown in table 2.

The outcome of the simulations for the period up to the year 2030 is demonstrated in figures 2–5. They show the impacts of the above assumptions on the major macrovariables³

The *optimistic* scenario generates very high rates of growth of GDP – initially up to 8% and then going down to 5–6% at the end of simulation period. The deviations from the baseline forecast are substantial: from 12% in 2013 up to above 50% in 2030.

The above high rates of growth are mainly due to assumed increase in investment expenditures (fig. 3.) They reveal a cyclical behaviour. They initially grow up to 17%, then they decline to 5% in 2025. These expenditures will be higher than forecasted by 22% already in 2013, but twice as high in 2030. Taking into account the indirect impact of investment growth the domestic final demand would exceed the baseline forecast by 14% in 2013 and by 40% in 2030.

On the supply side the high rates of growth of potential GDP are noticed, considerably exceeding the forecast (fig.4). They oscillate around 8% being higher than those for effective GDP. Hence, the capacity utilization has a declining tendency, negatively affecting the investment growth. In total, potential GDP is higher by 50% than forecasted by the end of period.

Despite the impact of rising investment the TFP growth plays a significant role (fig. 5). Its rates of growth come up from 2,8% to 3,4% at the end of period. This is due to assumed rising domestic R&D capital and growing efficiency in the absorption of foreign R&D capital. The rate of growth of employment is declining because of the high increase in labour productivity. The unemployment rates decline to 6%.

Turning now to *pessimistic* scenario, where low levels of investment in fixed capital and knowledge capital were assumed – the rates of growth would be initially negative (–2%) (fig. 2). The recession would be over in 7 years. The rates of growth would reach 2% in 2025 only, and 4% in 2030 owing to the results of recovery. Hence, the GDP level would be lower by ca 50% than the forecast by the end of simulation period.

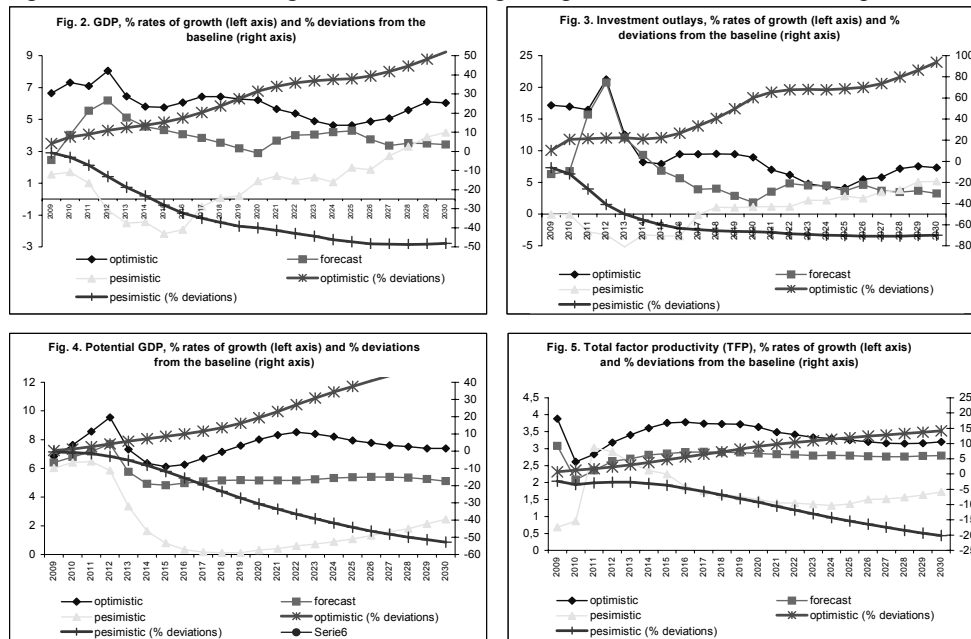
This result is mainly due to assumed decline in investment activities, initially by –5% (fig. 3). In the middle of the period they would raise to 2% and only in the last 5 years up to 5%. Nevertheless, in that period they would be lower than forecasted by 60-70%. The level of consumption stagnates. Hence, the domestic final demand after an initial decline by 2%, would show an increase by the end of period.

Potential GDP shows all the time positive rates of growth, however not exceeding 1% (fig 4). This is mainly because of the declining rates of growth of fixed capital. The labour productivity rates of growth are low, but sufficient to

³ Because of the limited space, the tables presenting the results of all simulations are available only in the monograph W. Welfe (ed.) [2009b].

sustain a systematic decline of employment. Hence, the unemployment rates rise dramatically from 10% to nearly 20% in the last years of simulation period.

Figure 5. The TFP rates of growth are declining being the results of low R&D expenditures



Source: own investigation

FINAL COMMENTS

The simulation analyses open the floor for constructing many alternative scenarios, that would take into account the factors of development other than investment in fixed capital and knowledge capital. However, the advantage the presented framework has is that it allows showing a clear distinction between the factors of growth and the results of the simulations exercise.

REFERENCES

- Aghion Ph., Howitt P. (1999) *Endogenous Growth Theory*, MIT Press, Cambridge MA. London
- Barro R.J., Sala-i-Martin X. (1995) *Economic Growth*, McGraw-Hill Inc., New York.
- Bårdsen G., Eitriheim O., Jansen E.S., Nymoen R. (2005) *The Econometrics of Macroeconometric Modelling*, Oxford University Press, Oxford.
- Bayoumi T., Coe D.T., Helpman E. (1999) R&D Spillovers and Global Growth, „*Journal of International Economics*” vol. 47, pp. 399–428.

- Benabou R. (2002) Human Capital. Technical Change and the Welfare State, „Temi di discussione del Servizio Studi”, vol. 465, Banca d'Italia, Roma.
- Cincera M., Pottelsberghe de la Potterie van B. (2001) International R&D Spillovers, A Survey, „Cahiers Economiques de Bruxelles”, vol. 169, no. I, pp. 1–20.
- Coe D.T., Helpman E. (1995) International R&D Spillovers, „European Economic Review”, vol. 39, pp. 859–887.
- Colecchia A., Schreyer P. (2002) ICT Investment and Economic Growth in the 1990s'. Is the United States a Unique Case? A Comparative Study of Nine OECD Countries, „Review of Economic Dynamics”, vol. 3, pp. 408–442.
- Cornwall J., Cornwall W. (2002) A Demand and Supply Analysis of Productivity Growth, „Structural Change and Economic Dynamics”, vol. 13, pp. 203–229.
- Engelbrecht H.J. (1997) International Spillovers. Human Capital and Productivity in OECD Economies. An Empirical Investigation, „European Economic Review”, vol. 41, pp. 1479–1488.
- Fuente de la, A. (2004) Human Capital and Growth: Some Results for the OECD. Current Issues of Economic Growth, „Proceedings of OeNB Workshops”, no. 2, Vienna.
- Garratt A., Lee K., Pesaran M.H., Shin Y. (2006) Global and National Macroeconometric Modeling: A Long-Run Structural Approach, Oxford University Press, Oxford.
- Grossman G.M., Helpman E. (1991) Innovation and Growth in the Global Economy, MIT Press, Cambridge
- Grzęda-Latocha R. (2005) Ekonometryczna analiza koniunktury gospodarczej w krajach strefy Euro (Econometric Analysis of Business Cycles in Euro-area) „Ekonomista”, nr 5, pp.621-640.
- Helpman E. (2004) The Mystery of Economic Growth, Harvard University Press, Cambridge, MASS
- Jorgenson D.W. (2000) Information Technology and the US Economy, „American Economic Review”, vol. 91. pp. 1–32
- Jorgenson D.W., Ho M.S., Strip K.J. (2002) Projecting Productivity Growth. Lessons from the Resurgence, Federal Reserve Bank of Atlanta, „Economic Review”, vol. 87, pp. 1–14
- Jorgenson D.W., Ho M.S., Stiroh K.J. (2003) Lessons for Europe from the U.S. Growth Resurgence, „CES IFO Economic Studies”, vol. 49, pp. 27–47
- Klein L.R., Welfe A., Welfe W. (1999) Principles of Macroeconometric Modelling, North Holland, Amsterdam
- Lee G. (2005) Direct versus Indirect International R&D Spillovers, „Information Economics and Policy”, vol. 17. pp. 334–348
- Leeuwen G. van, Wiel H. van der (2003) Spillovers Effects of ICT, „CBP Report”, nr 3, pp. 24–40
- Lichtenberg F.R., Van Pottelsberghe de la Potterie B. (1998) International R&D Spillovers: A Comment, „European Economic Review”, vol.42, pp. 1483–1491

- Nahuis R. (2003) *Knowledge, Inequality and Growth in the New Economy*, E. Elgar, Cheltenham UK
- Richards G.R. (2000) *Endogenous Technological Advance in Econometric Model; Implications for Productivity and Potential Output in the United States* "Economic Modelling", vol. 17 pp. 13–34
- Smith K. (2002) *What is the 'Knowledge Economy'?* Knowledge Intensity and Distributed Knowledge Bases. "Discussion Paper", series 2002–6, The UN University, INTECH, Maastricht
- Solow R., 1957, *Technical Change and Aggregate Production Function*. „Review of Economics and Statistics”, vol. 39, pp. 312–320
- Tokarski T. (2001) *Modele wzrostu endogenicznego (Models of endogenous growth)*, in: W. Welfe (ed.), *Ekonometryczny model wzrostu gospodarczego*, Wydawnictwo Uniwersytetu Łódzkiego, Łódź
- Tokarski T. (2007) *Podażowe determinanty wzrostu gospodarczego (Supply Determinants of Economic Growth)*, in: W. Welfe (ed). *Gospodarka oparta na wiedzy*, PWE, Warszawa
- Welfe W. (1992) *Ekonometryczne modele gospodarki narodowej Polski, (Econometric Models of the Polish Economy)*, PWE, Warszawa
- Welfe W., 2000, *Empiryczne modele wzrostu gospodarczego, (Empirical Models of Economic Growth)*, „*Ekonomista*”, no. 4, pp. 483–499
- Welfe W. (ed.) (2001) *Ekonometryczny model wzrostu gospodarczego, (Econometric Model of Economic Growth)*, Wydawnictwo Uniwersytetu Łódzkiego, Łódź.
- Welfe W. (2002) *Łączna produktywność czynników produkcji a postęp techniczny, (TFP and Technological Progress)*, „*Studia Ekonomiczne*”, vol. 36–37, pp. 94–115
- Welfe W., Floreczak W., Sabanty L. (2002) *Kapitał ludzki i jego endogenizacja (Human Capital and its Endogenization)*, „*Przegląd Statystyczny*”, vol. 50, nr 2, pp. 7–36
- Welfe W., Welfe A. (2004) *Ekonometria stosowana, (Applied Econometrics)*, II edition, PWE, Warsaw
- Welfe W. (ed.) (2004) *Długookresowy, makroekonometryczny model W8D-2002 gospodarki polskiej, (Long-term Macroeconometric Model of the Polish Economy W8D-2002)*, „*Acta UŁ, Folia Oeconomica*”, vol. 172, Wydawnictwo Uniwersytetu Łódzkiego, Łódź
- Welfe W. (2005) *Stylized Empirical Model of Economic Growth*, in: W. Welfe, A. Welfe (eds.), *Macromodels 2004*, „*Acta UŁ, Folia Oeconomica*” vol. 190, Wydawnictwo Uniwersytetu Łódzkiego, Łódź, pp. 109–125
- Welfe W. (2007) *Knowledge Capital and Total Factor Productivity*, „*Argumenta Oeconomica*” vol. 19 (1–2), pp. 211–235
- Welfe W. (ed.) (2007) *Gospodarka oparta na wiedzy, (Knowledge-based economy)*, PWE, Warszawa
- Welfe W. (2008a) *Knowledge Based Economy. New Directions of Macromodelling*, „*International Advances in Economic Research*”, vol. 143, pp. 167–180
- Welfe W. (2008b) *The Structure of a New Long-Term Macroeconometric Model for Poland*, „*Przegląd Statystyczny*”, vol. 55, pp. 5–23

- Welfe W. (ed.) (2009a) Makroekonometryczny model gospodarki opartej na wiedzy, (Macroeconometric Model of Knowledge based Economy, Acta UŁ, Folia Oeconomica, Wydawnictwo Uniwersytetu Łódzkiego, Łódź
- Welfe W. (ed.) (2009b) Prognozy i scenariusze długookresowego rozwoju gospodarczego Polski (Forecasts and Scenarios of Long-term Development of Polish Economy), Wydawnictwo Uniwersytetu Łódzkiego, Łódź
- Xu B., Wang, J. (1999) Capital Goods Trade and R&D Spillovers in OECD, „Canadian Journal of Economics”, vol. 32, p.1258–1274
- Zienkowski L. (ed.) (2003) Wiedza a wzrost gospodarczy, (Knowledge and Economic Growth), Wydawnictwo Naukowe SCHOLAR, Warszawa

Długookresowy model makroekonometryczny gospodarki polskiej.

Streszczenie: Rozwój współczesnych gospodarek rynkowych w kierunku gospodarek opartych na wiedzy implikuje konieczność budowy nowych długookresowych modeli makroekonometrycznych. Jako przykład pokazano strukturę i wykorzystanie długookresowego W8 modelu dla Polski. Model ten jest kompletny, objaśnia zarówno produkcję, jak i czynniki jej wzrostu, uwzględniając nakłady na B+R i kapitał ludzki, a także popyt finalny, co uniemożliwia wykazywanie potencjalnych źródeł nierównowagi. Model zamykają równania cen i płac oraz objaśniające przepływy finansowe. Symulacyjna wersja modelu zawiera 235 równań, w tym 111 stochastycznych. Model jest używany w budowie długookresowych scenariuszy rozwoju Polski do 2030r.

Słowa kluczowe: kapitał wiedzy, nakłady na B+R, kapitał ludzki, makromodelowanie, mnożniki, scenariusze

COMPARISON OF THE DETERMINISTIC AND STOCHASTIC APPROACHES FOR ESTIMATING TECHNICAL EFFICIENCY ON THE EXAMPLE OF NON-PARAMETRIC DEA AND PARAMETRIC SFA METHODS

Agnieszka Bezat

Katedra Ekonomiki Rolnictwa i Międzynarodowych Stosunków Gospodarczych
SGGW

e-mail: agnieszka.bezat@ierigz.waw.pl

Streszczenie: The author considers in the article estimation of the technical efficiency, which measures the ability of the company to obtain the maximum output from given inputs or to use the minimum input to achieve given outputs. The comparison of two approaches: deterministic (on the example of Data Envelopment Analysis) and stochastic (on the example of Stochastic Frontier Approach) has been carried out, the advantages and disadvantages of both were also described. These methods were chosen because they have become popular in polish research. In the article the possible limitations and problems, which may influence results of studies conducted by using these methods, were considered.

Keywords: efficiency measuring, Stochastic Frontier Approach, Data Envelopment Analysis, Decision Making Units

INTRODUCTION

The empirical applications of efficiency analysis were conducted in such sectors like: accounting, advertising, auditing and law firms, airports, air transport, bank branches, bankruptcy prediction, community and rural health care, dentistry, education, electricity, environment, fishing, forestry, hospitals, hotels, macroeconomics, military, rail transport, sport, tax administration, water distribution etc. [Fried et al. 2008, p 16]. The DEA (Data Envelopment Analysis) and SFA (Stochastic Frontier Approach) are the main methods commonly used to estimate efficiency of a DMU – Decision Making Units (commercial entities that produce tan-

gible goods and services that are sold in the market place, enterprise involved in delivering services or in the non-market sector, public bodies, national economic sector etc.).

The first method is a non-parametric, deterministic procedure for evaluating the frontier. Non-parametric procedures determine a frontier which “envelops” the observations. DEA employs flexible, nonparametric methods to construct the best-practice frontier and so allows the data to “speak for themselves” [Bates et al. 1996, p 1443]. For the deterministic approaches the frontier is defined by the maximum distances between input and output. Random error and characterises deviations from the frontier are interpreted as inefficiency [Coelli et al. 2005]. The DEA bases on a linear programming model which allows to build a piecewise linear frontier and assumes a convex production sets. A less constrained alternative to DEA is non-stochastic method known as free-disposal hull (FDH).

The second one widely uses stochastic procedure for parametric evaluating the frontier. The approach is stochastic – it considers additionally a random variable. The stochastic frontier approach treats deviations from production function as comprising both random error (white noise) and inefficiency [Mortimer&Peacock 2002, p 2]. This enables a distinction between a random symmetrical component which accounts for measurement errors and stochastic effects (e.g. due to weather influences) and a symmetric deviation component which represents the inefficiency. The SFA as a parametric approach requires assuming a specific function form a priori, the frontier is estimated econometrically by some variant of last squares or maximum likelihood [Coelli et al. 2005]. The SFA bases on econometric regression model, the frontier is smooth, appropriately and curved.

Both methods require all decision making units to have comparable inputs and outputs. Both methods can handle multiple input and multiple output models. These techniques should be used in conjunction with carefully compiled data on input and output quantities and prices [Coelli et al. 2005, p 133].

DETERMINISTIC, NON-PARAMETRIC MODEL (ON THE EXAMPLE OF DEA)

The piece-wise-linear convex hull approach to frontier estimation, proposed by Farrell (1957), was considered by only a few authors in the two decades following Farrell’s paper. Boles (1966), Shepard (1970) and Afriat (1972) suggested mathematical programming methods that could achieve the task, but the method had not received wide attention until the paper by Charnes, Cooper and Rhodes (1978), in which the term data envelopment analysis (DEA) was first used. Since then a large number of paper have appeared, which have extended and applied the DEA methodology.

For each DMU, to estimate efficiency, one would like to obtain a measure of the ratio of all outputs over all inputs, such as $u'q_i/v'x_i$, where u is a vector of out-

put weights and v is a vector of input weights. The optimal weights are obtained by solving the mathematical programming problem [Coelli et al. 2005, p 162]:

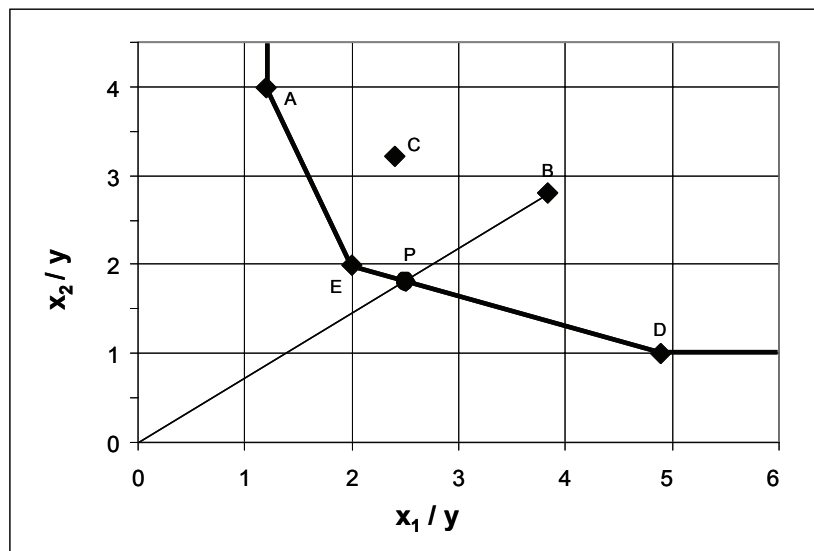
$$\max_{u,v} (u'q_i / v'x_i) \quad (1)$$

$$\text{st } u'q_i / v'x_i \leq 1 \quad (2)$$

$$u, v \geq 0 \quad (3)$$

This involves finding values for u and v , such that the efficiency measure for the i -th DMU maximised, subject to the constraints that all efficiency measures must be less or equal to one. One considers the input-output combinations that are observed in the examined sample of decisions making units. An efficiency measurement is calculated basing on the distance between the input-output combination of the respective unit and the frontier [Farrell 1957].

Figure 1. A two input and one output case



Source: Cooper et al. 2007, p 57

The technical efficiency for DMU B is calculated as a ratio (s. Figure 1):

$$TE^B = \frac{OP}{OB} \quad (4)$$

Point P indicates a virtual DMU which is similar to B, but located on the frontier. Using virtual benchmarks such as P implies that linear combinations of real DMUs (in this case E and D) are considered.

STOCHASTIC, PARAMETRIC MODEL (ON THE EXAMPLE OF SFA)

Stochastic frontier analysis (SFA) is an alternative method to frontier estimation that assumes a given functional form for the relationship between inputs and an output [Coelli et al. 2005, p 209]. The stochastic production function model was proposed independently by Aigner, Lovell, Schmidt and Meeusen, van den Broeck in 1977. The model had following form:

$$\ln q_i = x_i' \beta + v_i - u_i \quad (5)$$

where q represents the output of the i -th DMU, x_i is a $K \times 1$ vector containing the logarithms of inputs, β is a vector of unknown parameters, and u_i is a non-negative variable associated with technical inefficiency, v_i is a symmetric random error, to account for statistical noise. The statistical noise arises from the inadvertent omission of relevant variables from vector x_i , as well as from measurement errors and approximation errors with the choice of functional form.

In the equation (5) the output values are bounded by the stochastic variable:

$$\exp(x_i' \beta + v_i) \quad (6)$$

The value v_i can be positive or negative and so the stochastic frontier outputs vary about the deterministic part of the model:

$$\exp(x_i' \beta) \quad (7)$$

The SFA requires choosing a production function model: Coob-Douglas, CES, translog, generalised Leontief, normalised quadratic and its variants. The translogarithmic and the Cobb-Douglas production functions are the two most common functional forms which have been used in empirical studies on production, including frontier analyses [Battese & Broca 1997, p 397]. Using the Cobb-Douglas stochastic frontier model¹ one can illustrate the variation of random error about the deterministic part. A Cobb-Douglas stochastic frontier model takes the form:

$$\ln q_i = \beta_0 + \sum_{j=1}^k \beta_j \ln x_{ij} + v_i - u_i \quad (8)$$

so

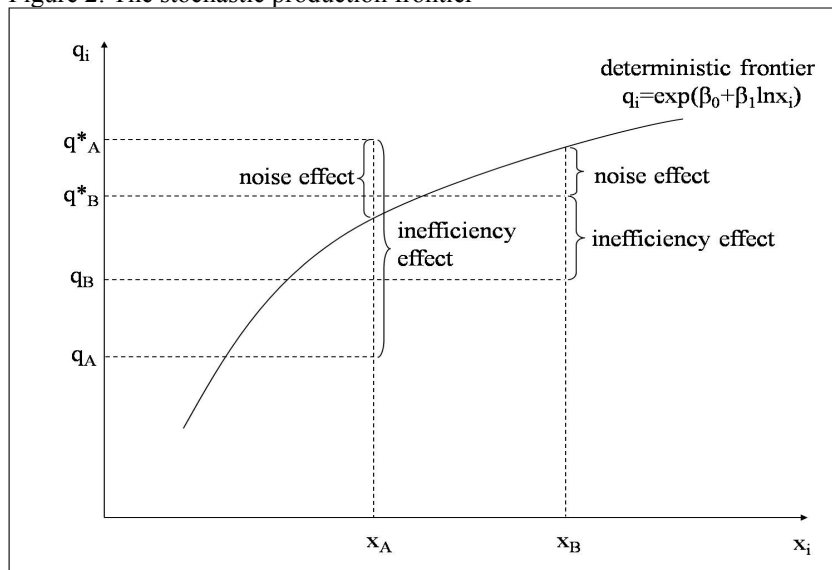
$$q_i = \exp(\beta_0 + \sum_{j=1}^k \beta_j \ln x_{ij}) + \exp(v_i) - \exp(u_i) \quad (9)$$

¹ The analysis is carried out for the DMU that produce the output q_i using only the input x_i .

The observed deviation from actual point of production to the frontier ($\exp(v_i) - \exp(u_i)$) is a composed error. Random component v refers to unsystematic deviations from the frontier – noise.

In the Figure 2 the deterministic model with the error noise and inefficiency effect was shown. The feature of frontier model was explained on the example of two DMUs: A and B, which are using the input x_A and x_B for producing the output y_A and y_B respectively.

Figure 2. The stochastic production frontier



Source: Own work on the basis of Coelli et al. 2005, p 244

In the figure the so-called frontier output is depicted. This measure assumes the case if there were no inefficiency effects (i.e., $u_A=0$ and $u_B=0$). The values q^* assume no inefficiency effects (i.e., if $u_A=0$ and $u_B=0$). For the DMU A the frontier output lies above the deterministic part of the production frontier (the value q^*_A) only because the noise effect is positive (i.e., $v_A>0$). The frontier output of the second DMU lies below the deterministic part of the frontier (the value q^*_B) because the noise effect is negative (i.e., $v_B<0$).

The frontier outputs tend to be evenly distributed above and below the deterministic part of the frontier. However, observed outputs tend to lie below the deterministic part of the frontier. Indeed, they can only lie above the deterministic part of the frontier when the noise effect is positive and larger than the inefficiency effect [Coelli et al. 2005, p 244].

Commonly using output-oriented measure of technical efficiency is the ratio of observed output to the corresponding stochastic frontier output:

$$TE_i = \frac{q_i}{\exp(x_i'\beta + v_i)} = \frac{\exp(x_i'\beta + v_i - u_i)}{\exp(x_i'\beta + v_i)} = \exp(-u_i) \quad (9)$$

Regarding to the equation (9) the TE takes a value between zero and one. It is important to understand that it measures the output of i-th DMU relative to the output that could be produced by a fully-efficient DMU using the same input vector. So using the SFA one can estimate a relative efficiency of some DMU.

In series of studies authors have explored the implications of variety of distributional assumption an estimation of efficiency. Generally it is required to assume a distribution of u from:

half-normal distribution: $u_i \sim N^+(0, \sigma_u)$,

exponential distribution: $u_i \sim EXP(\lambda)$,

truncated-normal distribution: $u_i \sim N^+(\mu_u, \sigma_u)$,

gamma distribution: $u_i \sim \Gamma(m, \sigma_u)$.

The first two distributions have just a single parameter, are empirically traceable and easy to estimate. The last two distributions have two parameters, what makes them more flexible but also more difficult to estimate. The choice of distribution of u influences quite strongly a level of TE and less rankings of DMUs. Under fairly weak assumption it is usually possible and appropriate to estimate models using the method of last squares. Slightly stronger distributional assumption allows estimating the unknown parameters using maximum likelihood² or Bayesian techniques³ [Coelli et. al 2005, p 240]. When decisions about function and distribution must be made, it is recommended to estimate a number of the alternative models and to select a preferred model using likelihood ratio test [Coelli 1996].

Commonly used method for estimation of stochastic frontier is a maximum likelihood (ML). ML estimations rest on the assumption that the distribution of the errors is actually known. Battese and Coelli (1992) propose a stochastic frontier production function which is assumed to be distributed as truncated normal random variables.

LIMITATIONS AND PROBLEMS WITH DEA APPLICATIONS

In conducting a DEA study some limitations and possible problems may encounter.

² Maximum likelihood estimators are popular because they have desirable large sample properties.

³ Bayesian estimation is becoming increasingly popular, not least because it allows obtaining exact finite-sample results concerning nonlinear functions of the parameters.

- One of the most notable characteristics of DEA is its deterministic approach to efficiency measurement. DEA does not allow for estimation or measurement error. The full distance of a brand to the efficiency frontier is interpreted as inefficiency. But a measurement error or other noise and outliers may influence the shape and position of the frontier.
- The exclusion of an important input or output can result in biased results.
- Efficiency measurements can differ depending on the model specifications (input- vs. output-oriented models) and the variable specification (e.g. the degree of aggregation and the units used to measure inputs and outputs).
- The efficiency scores are only relative to the best DMU in the sample. The inclusion of extra DMU (e.g. from other countries) may reduce efficiency scores.
- It has been pointed out that the technical efficiency of any single enterprise or decision making unit (DMU) estimated using data envelopment analysis (DEA) will tend to decrease as the number of DMUs included in the DEA application increases [Zhang & Bartels 1998, p 187]. This is because, as the number of DMUs increases, the chance of encountering enterprises close to the true production frontier increases, and therefore the frontier constructed by DEA approaches the true frontier asymptotically as the number of enterprises in an industry increases [Banker 1989].
- DEA is good at estimating "relative" efficiency, thus the measurements are only valid in a sample. Units which have not been included in the sample can produce a shift of the frontier. The method's results say nothing about the efficiency of one sample relative to the other – they only reflect the dispersion of efficiencies within each sample. One should not compare the mean efficiency scores from two studies.
- The addition of an extra DMU in a DEA analysis cannot result in an increase in the TE scores of the existing enterprises.
- The addition of an extra input or output in DEA model cannot result in a reduction in the TE scores.
- When one has a few observations and many inputs and/or outputs, many of DMUs will appear on the DEA frontier. One could reduce the sample size and increase the number of inputs and/or outputs in order to increase the TE scores.
- Not accounting for environmental differences may give misleading indications of relative managerial competence.
- Standard DEA does not account neither for multi-period optimisation nor risk in management decision making.
- Since DEA is a nonparametric technique, statistical hypothesis tests are difficult.

LIMITATIONS AND PROBLEMS WITH SFA APPLICATIONS

In conducting a SFA study some limitations and possible problems may encounter. An analyst should keep these limitations in mind when choosing whether or not to use SFA.

- In SFA studies an assumption regarding to a specific functional form of stochastic frontier is required a priori. A wrong choice of production function may influence the results.
- The simple production frontier model does not permit the prediction of the technical efficiencies of companies that produce multiples outputs.
- The maximum likelihood does not allow assessing the reliability of inferences in small samples.
- Absolute level of TE is quite sensitive to distributional assumptions, rankings are less sensitive.
- The SFA requires using of large number of DMUs.
- It should also be stressed that all of issues described concerning to DEA method are also applicable (in varying degrees) to the stochastic frontier method.

CONCLUSIONS

The traditionally econometric belief in the presence of external forces contributing to random statistical noise is continuing maintained. Thus, it is desirable for the econometric approach to be relatively more successful than others, so as to provide the basis for a subsequent investigation into determinants of variations in the efficiency. On the other hand, a researcher has to choose the functional form of the frontier and make an assumption regarding to distribution of variation in inefficiency. A wrong choice may be corrected on the basis of statistical test (e.g. likelihood ratio test or alternatively Wald's test).

In conducting a DEA study, the production frontier "envelops" the observations, thus it does not require an assumption of a functional form relating inputs to outputs. Random deviations from the frontier are interpreted as inefficiency; the statistical noise is not included. A disadvantage of DEA method is that a frontier position and efficiency scores may be strongly influenced by a measurement error or outliers. In the same way the exclusion of an important input or output results. The DEA method is useful for the relatively to SFA smaller sample. But one should be carefully because a small number of DMUs may causes that many of DMUs will appear on the DEA frontier and an average level of efficiency scores will be very high.

REFERENCES

- Afriat S.N. (1972) Efficiency Estimation of Production Functions. *International Economic Review*, 13, 568 – 598.
- Banker R.D. (1989) Econometric Estimation and Data Envelopment Analysis. *Research in Government and Non-profit Accounting* 5, 231 – 243.
- Bates J.M., Baines D., Whynes D.K. (1996) Measuring the efficiency of prescribing by general analysis. *Journal of Operational Research Society*, 47 (12), 1443 – 1451.
- Battese G.E., Broca S.S. (1997) Functional Forms of Stochastic Frontier Production Functions and Models for Technical Inefficiency Effects: A Comparative Study for Wheat Farmers in Pakistan. *Journal of Productivity Analysis*. 8, 395 – 414.
- Battese G.E., Coelli T.J (1992) Frontier Production Function, Technical Efficiency and Panel Data: With Application to Paddy Farmers in India. *Journal of Productivity Analysis*. 3, 153 – 169.
- Boles J.N. (1966) Efficiency Squared – Efficiency Computation of Efficiency Indexes. *Proceedings of the 39th Annual Meeting of the Western Farm Economics Association*, 137 – 142.
- Charnes A., Cooper W.W., Rhodes E. (1978) Measuring the Efficiency of Decisions Making Units. *European Journal of Operational Research*, 2, 429 – 444.
- Coelli T.J., Rao D.S.P, O'Donnell Ch.J., Battese G.E. (2005) *An introduction to efficiency and productivity analysis* (2nd ed.). USA: Springer.
- Coelli T.J. (1996) *A Guide to Frontier Version 4.1: A Computer Program for Stochastic Frontier Production and Cost Function Estimation*. Armidale: University of New England, CEPA Working Papers, p 6.
- Cooper W.W., Seiford L.M., Tone K. (2007) *Data Envelopment Analysis. A Comprehensive Text with Models, Applications, References and DEA-Solver Software* (2nd ed.). USA: Springer.
- Farrell M.J. (1957) The Measurement of Productive Efficiency. *Journal of the Royal Statistical Society, Series A (General)* 120 (3), 253 – 290.
- Fried H.O., Lovell C.A.K., Schmidt S.S. (2008) *The Measurement of Productive Efficiency and Productivity Growth*. USA: Oxford University Press US, 16 – 20.
- Mortimer D., Peacock S. (2002) *Hospital Efficiency Measurement: Simple Ratios vs Frontier Methods*. Australia: Centre of Health Program Evaluation. (Working Paper 135).
- Shepard R.W. (1970) *Theory of Cost and Production Functions*. Princeton: Princeton University Press.
- Zhang Y., Bartels R. (1998) The Effect of Sample Size on the Mean Efficiency in DEA with an Application to Electricity Distribution in Australia, Sweden and New Zealand. *Journal of Productivity Analysis*. 9, 187 – 204.

Porównanie deterministycznych i stochastycznych podejść w wyznaczeniu efektywności technicznej na przykładzie metod: nieparametrycznej DEA i parametrycznej SFA

Streszczenie: Autorka analizuje w artykule efektywność techniczną, która pozwala na mierzenie zdolności przedsiębiorstwa do uzyskiwania maksimum efektów przy danym poziomie nakładów lub do wykorzystywania minimum nakładów do osiągnięcia danego poziomu wyników. Przeprowadzono porównanie dwóch podejść: deterministycznego (na przykładzie Data Envelopment Analysis) oraz stochastycznego (na przykładzie Stochastic Frontier Analysis), a mianowicie przedstawiono zalety i wady tych dwóch metod. Wybrano te metody ze względu na ich coraz szersze zastosowanie w polskich badaniach. W artykule uwzględniono możliwe ograniczenia i problemy, które mogą wpływać na wynik badań przeprowadzonych za pomocą omawianych metod.

Słowa kluczowe: mierzenie efektywności, Stochastic Frontier Approach, Data Envelopment Analysis, Decision Making Units

ZALEŻNOŚĆ OCENY ZRÓŻNICOWANIA ROLNICTWA W POLSCE OD WYBRANYCH MIERNIKÓW SYNTETYCZNYCH

Agata Binderman

Katedra Ekonometrii i Statystyki SGGW
e-mail: agata_binderman@sggw.pl

Streszczenie: W pracy, za pomocą trzech różnych mierników syntetycznych zbadano poziom rozwoju rolnictwa polskich województw w 2007 roku. Podane metody wykorzystują miernik bezwzorcowy, jak również mierniki oparte o jeden lub dwa wzorce. W celu zapewnienia porównywalności otrzymanych rezultatów zastosowano standaryzację zmiennych, jako sposób normalizacji danych oraz metrykę Euklidesa, jako sposób mierzenia odległości między rozważanymi obiektami.

Słowa kluczowe: poziom rozwoju rolnictwa, mierniki syntetyczne, funkcje użyteczności, znormalizowana funkcja Hellwiga, uporządkowanie liniowe, klasyfikacja.

WSTĘP

Praca jest kontynuacją badań nad stanem rolnictwa w polskich województwach [Binderman A. 2005a,b, 2006,a,b, 2007, 2008]. W pracy [Binderman A. 2009 w druku] zbadany został poziom rozwoju polskiego rolnictwa w latach 2004 - 2007, w celu określenia tego poziomu wykorzystano metody unitaryzacji zerowanej [Kukuła K. 2000] oraz mierniki mające charakter liniowej i nieliniowej funkcji użyteczności [Binderman A. 2008]. W niniejszej pracy zbadano zróżnicowanie regionalne polskiego rolnictwa na podstawie najnowszych możliwych do uzyskania na ten czas danych GUS z roku 2007. Do uporządkowania województw wykorzystano rozważane wcześniej funkcje użyteczności oraz syntetyczny miernik rozwoju Hellwiga [Hellwig Z. 1968].

Metody porządkowania obiektów można podzielić na wzorcowe i bezwzorcowe. Metody bezwzorcowe polegają na konstrukcji miernika syntetycznego (agregatowego), na podstawie tylko znormalizowanych wartości cech. Metody wzorcowe polegają na skonstruowaniu taksonomicznych mierników rozwoju -

sztucznych punktów odniesienia, mierzeniu odległości od tych wzorców i na tej podstawie konstruowaniu miernika syntetycznego [Malina A. 2004, Młodak A. 2006, Nowak E. 1990, Zeliaś A. 2000].

W pracy, dla zapewnienia porównywalności otrzymanych wyników jako sposób normalizacji zmiennych wykorzystana została standaryzacja, natomiast jako sposób mierzenia odległości - metryka Euklidesa [Zeliaś A. 2000, Bartosiewicz S. 1976]. Wybór metryki i sposobu normalizacji związany był z metodą Hellwiga.

W celu analizy poziomu rolnictwa w Polsce, dobrym narzędziem są syntetyczne mierniki, które pozwalają opisać badane obiekty, tj. województwa, charakteryzowane przez wektory w wielowymiarowej przestrzeni cech. Metoda ta umożliwia analizę porównawczą grup województw, a także pozwala na uporządkowanie badanych obiektów według przyjętego kryterium oraz określić miejsce, jakie zajmuje dane województwo w rankingu wszystkich rozpatrywanych obiektów.

Bez straty dla ogólności rozważań przyjęto założenie, że rozważane cechy mają charakter stymulant [Borkowski B., Dudek H., Szczesny W. 2004].

METODYKA BADAŃ

Rozważmy problem polegający na klasyfikacji $m \in N$ obiektów Q_1, Q_2, \dots, Q_m badanego zjawiska za pomocą $n \in N$ zmiennych, mających charakter cech ilościowych, traktowanych, jako zmienne losowe X_1, X_2, \dots, X_n . Przyjmujemy założenie, że każdy taki obiekt daje się przedstawić za pomocą wektora należącego do przestrzeni \mathfrak{R}^n . Niech wektor $\mathbf{x}_i = (x_{i1}, x_{i2}, \dots, x_{in})$, $i=1, 2, \dots, m$, opisuje i -ty obiekt.

Jeżeli $x_{ik} > x_{jk}$ ($x_{ik} \geq x_{jk}$) dla $k = 1, 2, \dots, n$ to pisać będziemy:

$$\mathbf{x}_i > \mathbf{x}_j, \quad (\mathbf{x}_i \geq \mathbf{x}_j),$$

gdzie $i, j \in [1, m]$.

Jeżeli $\mathbf{x}_i \geq \mathbf{x}_j$ i $\mathbf{x}_i \neq \mathbf{x}_j$ to naturalnym jest nazywać obiekt \mathbf{x}_i lepszym (wyżej ocenianym) od obiektu \mathbf{x}_j . Oznacza to, że żadna ze składowych wektora \mathbf{x}_i nie jest mniejsza od odpowiednich składowych wektora \mathbf{x}_j , a przynajmniej jedna z nich ma wartość większą, tj. istnieje takie $k \in [1, n]$, że $x_{ik} > x_{jk}$.

Przyjmijmy następujące oznaczenia wektorów:

$$\mathbf{x}_0 := (x_{0,1}, x_{0,2}, \dots, x_{0,n}), \quad \text{ i } \quad \mathbf{x}_{m+1} := (x_{m+1,1}, x_{m+1,2}, \dots, x_{m+1,n}),$$

gdzie:

$$x_{0,k} := \min_{1 \leq i \leq m} x_{ik}, \quad x_{m+1,k} := \max_{1 \leq i \leq m} x_{ik}, \quad k=1, 2, \dots, n.$$

Założmy, że zdefiniowane wyżej wektory $\mathbf{x}_0, \mathbf{x}_{m+1}$ opisują obiekt Q_0 (najgorszy) i obiekt Q_{m+1} (najlepszy), odpowiednio. Obiekty Q_0 i Q_{m+1} będą pełnić w naszych rozważaniach rolę wzorców.

Założmy ponadto, że $\mathbf{x}_0 \neq \mathbf{x}_{m+1}$, oznacza to, że wektory $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m$ nie są identyczne. Oczywiście, wektory te spełniają nierówności:

$$\mathbf{x}_0 \leq \mathbf{x}_i \leq \mathbf{x}_{m+1} \quad \text{ dla } i \in [1, m].$$

Złożone zjawiska ekonomiczne, określone są najczęściej za pomocą wielu zmiennych o różnych mianach oraz rzędach wielkości. Powstaje problem przekształcenia tych cech w taki sposób, aby rozpatrywane łącznie, spełniały warunek porównywalności. Transformację zmiennych diagnostycznych, która prowadzi do porównywalności oraz addytywności wartości tych cech nazywamy normalizacją lub normowaniem zmiennych. W literaturze przedmiotu można wyróżnić cztery sposoby normalizacji (grupy metod):

1. standaryzację,
2. unitaryzację,
3. przekształcenia ilorazowe względem punktu odniesienia,
4. rangowanie wartości zmiennych.

Wyczerpująca lista sposobów normalizacji zmiennych podana jest np. w książkach [Kukuła K. 2000, Młodak A. 2006, Strahl D., Walesiak M. 1996, 1997, Walesiak M. 2006, Zeliaś A. 2000]. Przyjmijmy następującą definicję.

Definicja 1 Układ wektorów $x_0, x_1, x_2, \dots, x_m, x_{m+1}$, $m \in \mathbb{N}$ nazywać będziemy *znormalizowanym* jeżeli obiekty x_0, x_{m+1} są reprezentowane przez wektor zerowy i jednostkowy, odpowiednio, tj. $x_0=0=(0,0,\dots,0)$, $x_{m+1}=1=(1,1,\dots,1)$.

Jeżeli układ wektorów $x_0, x_1, x_2, \dots, x_m, x_{m+1}$ jest znormalizowany to $0 \leq x_{i,k} \leq 1$ dla każdego $i=0,1,\dots,m+1$; $k=1,2,\dots,n$.

Jeżeli wyjściowe zmienne diagnostyczne poddamy normalizacji za pomocą unitaryzacji zerowanej, to w wyniku tej operacji uzyskamy znormalizowany układ wektorów. Własności tej nie mają na ogół takie sposoby normalizacji, jak standaryzacja czy też przekształcenie ilorazowe.

Definicja 2. Każdą rosnącą funkcję $u: \mathfrak{R}^n \rightarrow \mathfrak{R}$ nazywać będziemy *funkcją użyteczności* [Allen R. 1964, Panek E. 2000].

Z definicji wynika, że dla dowolnej pary wektorów $x, y \in \mathfrak{R}^n$ spełniona jest implikacja: $x \geq y \wedge x \neq y \Rightarrow u(x) > u(y)$. Dlatego obiekt x uważany będzie za lepszy od obiektu y , jeżeli $u(x) > u(y)$, oznacza to, że obiekt lepszy od drugiego obiektu ma większą od niego użyteczność. Obiekty x, y uważane są za jednakowo dobre (obojętne), względem przyjętej funkcji użyteczności u , jeżeli $u(x) = u(y)$. W pierwszym przypadku mówić będziemy, że obiekt x jest *silnie preferowany* nad y , w drugim, że obiekty y i x są *indifferentne*. Jeżeli obiekty y i x są indifferentne lub obiekt x jest silnie preferowany nad y to mówić będziemy, że obiekt x jest *słabo preferowany* nad y [Panek E. 2000].

Definicja 3. Funkcję użyteczności u spełniającą warunek: $u(x_0) = 0$ i $u(x_{m+1}) = 1$, nazywać będziemy *znormalizowaną funkcją użyteczności*.

Można udowodnić następujące twierdzenia [Binderman A. 2007].

Twierdzenie 1. Niech d oznacza dowolną metrykę Minkowskiego ($1 \leq p < \infty$) [Zeliaś A. 2000], $x_0 \neq x_{m+1}$ wówczas funkcja:

$$U(x_i) := \frac{d(x_0, x_i) + d(x_0, x_{m+1}) - d(x_i, x_{m+1})}{2d(x_0, x_{m+1})}, \quad i = 0, 1, \dots, m, m+1, \quad (1)$$

jest znormalizowaną funkcją użyteczności przyjmującą wartości z przedziału $[0,1]$.

Twierdzenie 2. Funkcja

$$f(\mathbf{x}_i) = \frac{\langle \mathbf{x}_{m+1} - \mathbf{x}_0, \mathbf{x}_i - \mathbf{x}_0 \rangle}{\langle \mathbf{x}_{m+1} - \mathbf{x}_0, \mathbf{x}_{m+1} - \mathbf{x}_0 \rangle}, \quad i = 0, 1, \dots, m+1, \quad (2)$$

gdzie $\langle \mathbf{x}, \mathbf{y} \rangle$, oznacza iloczyn skalarny wektorów $\mathbf{x}, \mathbf{y} \in \mathfrak{R}^n$ jest znormalizowaną funkcją użyteczności.

Zauważmy, że zgodnie z przyjętymi określeniami funkcja U określona wzorem (1) jest nieliniowym, dwuwzorcowym miernikiem, natomiast funkcja f określona za pomocą wzoru (2) jest miernikiem bezwzorcowym.

Twierdzenie 3. Jeżeli układ wektorów $\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m, \mathbf{x}_{m+1}$ jest znormalizowany [zob. Cieślak M. 1993, Kukuła K. 2000] to:

$$f(\mathbf{x}_i) = \frac{1}{n} \sum_{k=1}^n x_{ik}, \quad \text{dla } i = 0, 1, \dots, m+1.$$

Niech będzie dany zdefiniowany wyżej układ wektorów $\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m, \mathbf{x}_{m+1}$ (przed normalizacją zmiennych). Przyjmijmy założenie, że cechy, na podstawie których określono te wektory mają charakter stymulant. W pracy do normalizacji zmiennych zastosowano standaryzację określoną za pomocą wzoru:

$$z_{ij} = \frac{x_{ij} - \bar{x}_j}{S_j}, \quad (i=1, \dots, n; j=1, \dots, m); \quad (3)$$

gdzie: $S_j = \sqrt{\frac{1}{m} \sum_{k=1}^m (x_{kj} - \bar{x}_j)^2}$ - odchylenie standardowe cechy x_j ,

$\bar{x}_j = \frac{1}{m} \sum_{k=1}^m x_{kj}$ - średnia arytmetyczna cechy $x_j, i=1, \dots, n; j=1, \dots, m$.

W wyniku normalizacji zmiennych, otrzymujemy układ wektorów:

$$\mathbf{z}_0, \mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_m, \mathbf{z}_{m+1}.$$

Syntetyczny miernik rozwoju Zdzisława Hellwiga $h(\mathbf{x}_i)$ [Hellwig Z. 1968] określany jest w następujący sposób:

$$h(\mathbf{x}_i) = h_i(\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_m) = 1 - \frac{d_i}{\bar{d} + 2\sigma(D)}, \quad (4)$$

gdzie:

$$d_i = \sqrt{\sum_{j=1}^m (z_{ij} - z_{m+1j})^2}, \quad \bar{d} = \frac{1}{m} \sum_{i=1}^m d_i, \quad \sigma(D) = \sqrt{\frac{1}{m} \sum_{i=1}^m (d_i - \bar{d})^2}, \quad i = 0, 1, \dots, m+1.$$

W powyższych wzorach wartości współrzędnych wektorów $\mathbf{z}_i, i=0, 1, 2, \dots, m+1$ są określone za pomocą wzoru (3).

Uwaga 1. Łatwo zauważyć, że jeżeli istnieją wektory x_k i x_j takie, że $x_k \geq x_j$ i $x_k \neq x_j$, $1 \leq j, k \leq m$, to po transformacji zmiennych określonych wzorem (3) wektory z_k i z_j spełniać będą warunki $z_k \geq z_j$ i $z_k \neq z_j$.

Zauważmy ponadto, że liczby \bar{d} i $\sigma(D)$ są wielkościami stałymi dla danego układu wektorów $x_1, x_2, \dots, x_m, x_{m+1}$ oraz $d_k \geq d_j$. Miernik Hellwiga wykorzystuje jeden wzorzec, najlepszy x_{m+1} .

Z powyższych rozważań wynika, że jeżeli istnieją wektory x_k i x_j takie, że $x_k \geq x_j$ i $x_k \neq x_j$, $1 \leq j, k \leq m$, to dla funkcji Hellwiga słuszna jest następująca nierówność:

$$h(x_k) > h(x_j).$$

Powyższa nierówność dowodzi słuszności następującego twierdzenia.

Twierdzenie 4. Miernik syntetyczny $h(x)$ jest funkcją użyteczności.

Miernik $h(x)$ nie jest na ogół standaryzowaną funkcją użyteczności, mimo że $h(x_{m+1})=1$. Zauważmy jednak, że funkcja

$$H(x) := \frac{h(x) - h(x_0)}{1 - h(x_0)}, \quad (5)$$

spełnia warunki normalizacyjne, tj. $H(x_0) = 0$ i $H(x_{m+1}) = 1$. Istotnie:

$$H(x_0) = \frac{h(x_0) - h(x_0)}{1 - h(x_0)} = 0, \quad H(x_{m+1}) = \frac{h(x_{m+1}) - h(x_0)}{1 - h(x_0)} = \frac{1 - h(x_0)}{1 - h(x_0)} = 1.$$

Oczywiście, jeżeli $h(x_0) = 0$, to $H(x) \equiv h(x)$.

Nietrudno zauważyć, że z faktu, iż funkcja h jest rosnąca wynika, że również funkcja H jest funkcją rosnącą. Dowiedzione zatem zostało następujące twierdzenie.

Twierdzenie 5. Jeżeli dany jest układ wektorów $x_0, x_1, x_2, \dots, x_m, x_{m+1}$, $x_0 \neq x_{m+1}$, to funkcja $H(x)$ określona wzorem (5) jest znormalizowaną funkcją użyteczności przyjmującą wartości z przedziału $[0, 1]$.

Definicja 4. Funkcję $H(x)$ nazywać będziemy *znormalizowaną funkcją Hellwiga*.

WYNIKI BADAŃ

Celem pracy było pokazanie oceny poziomu rolnictwa polskich województw, w zależności od zastosowanych mierników syntetycznych. Badań dokonano na podstawie danych empirycznych ze wszystkich polskich województw, które dotyczyły stanu polskiego rolnictwa w 2007 roku. Do opisu regionalnego zróżnicowania rolnictwa przyjęto 10 następujących zmiennych diagnostycznych [zob. Binderman A. 2007]:

- X_1 Udział użytków rolnych w % powierzchni ogółem.
- X_2 Wskaźnik waloryzacji rolniczej przestrzeni produkcyjnej (w punktach).
- X_3 Plony zbóż w tonach z 1 hektara.
- X_4 Obsada bydła w sztukach dużych na 100 hektarów użytków rolnych.

- X_5 Skup owoców z drzew w kg na 1 hektar powierzchni upraw.
- X_6 Wartość skupu produktów rolnych ogółem w zł na 1 ha użytków rolnych.
- X_7 Nakłady inwestycyjne w rolnictwie w zł na 1 hektar użytków rolnych.
- X_8 Wskaźnik zatrudnienia w rolnictwie w %.
- X_9 Średnia powierzchnia gospodarstwa rolnego w hektarach.
- X_{10} Produkt krajowy brutto w zł na 1 mieszkańca.

Wszystkie zmienne były stymulantami, oznacza to, że większe wartości cech wskazywały na wyższy poziom rozwoju badanego zjawiska.

Dla tak określonych danych otrzymano $m=16$ obiektów Q_1, Q_2, \dots, Q_{16} , każdy z nich został opisany przez $n = 10$ cech: X_1, X_2, \dots, X_{10} . Wartości przyjętych zmiennych diagnostycznych dla 16 województw w 2007 roku utworzyły macierz X o wymiarach 16×10 . Na podstawie wartości przyjętych przez cechy dla 16 województw, stworzono dwa hipotetyczne (wzorcowe) województwa, minimalne Q_0 oraz maksymalne Q_{17} , określone za pomocą najmniej i najbardziej korzystnego zestawu wartości cech (odpowiednio). Hipotetyczne obiekty reprezentowane były przez wektory x_0 oraz x_{17} o 10 składowych każdy.

Ponieważ wybrane cechy miały różne miana i rzędy wielkości, poddano je normalizacji, wykorzystując w tym celu metodę standaryzacji, według wzoru (3):

$$z_{ij} = \frac{x_{ij} - \bar{x}_j}{x_{17j} - x_{0j}}, \quad 0 \leq i \leq 17, \quad 1 \leq j \leq 10.$$

Wybór standaryzacji, jako sposobu normalizacji związany był z wymaganiami metody Hellwiga.

Do obliczenia syntetycznych mierników województw w roku 2007, dla porównania wyników, wykorzystano rozważane wcześniej trzy różne metody konstrukcji mierników syntetycznych oraz metrykę Euklidesa, tj.:

1. funkcję liniową f , określoną za pomocą wzoru (2):

$$f(z_i) = \frac{\langle z_{17} - z_0, z_i - z_0 \rangle}{\|z_{17} - z_0\|^2}, \quad i = 0, 1, 2, \dots, 17;$$

2. nieliniową znormalizowaną funkcję użyteczności U , według wzoru (1):

$$U(z_i) = \frac{\mu(z_0, z_i) + \mu(z_0, z_{17}) - \mu(z_i, z_{17})}{2\mu(z_0, z_{17})}, \quad i = 0, 1, \dots, 16, 17;$$

gdzie μ oznacza metrykę Euklidesa (metryka Minkowskiego przy $p=2$),

3. znormalizowaną funkcję Hellwiga określoną za pomocą wzoru (5):

$$H(z_i) := \frac{h(z_i) - h(z_0)}{1 - h(z_0)}, \quad i = 0, 1, \dots, 17,$$

gdzie: $h(\mathbf{z}_i) = 1 - \frac{2d_i}{2\bar{d} + \mu(\mathbf{d}, \mathbf{d})}$, $i = 0, 1, 2, \dots, 16, 17$,

$$d_i = \mu(\mathbf{z}_i, \mathbf{z}_{17}), \mathbf{d} = (d_1, d_2, \dots, d_{16}), \bar{d} = \frac{1}{16} \sum_{k=1}^{16} d_k, \mathbf{d} = (\bar{d}, \bar{d}, \dots, \bar{d}).$$

Oczywiście podane wyżej mierniki są znormalizowanymi funkcjami użyteczności, przyjmującymi wartości z przedziału $[0, 1]$, tj.:

$$f(\mathbf{x}_0) = 0, f(\mathbf{x}_{17}) = 1, U(\mathbf{x}_0) = 0, U(\mathbf{x}_{17}) = 1 \text{ i } H(\mathbf{x}_0) = 0, H(\mathbf{x}_{17}) = 1.$$

Przy użyciu podanych wyżej funkcji f , U i H , obliczone zostały mierniki syntetyczne poziomu rozwoju rolnictwa polskich województw w roku 2007. Wyznaczenie użyteczności dla województw pozwoliło następnie na dokonanie rankingu województw, tj. uporządkowania obiektów, według wartości mierników.

Tabela 1 przedstawia wartości syntetycznych mierników województw oraz ich pozycje, według trzech zastosowanych metod, jak również współczynniki korelacji między wektorami rezultatów.

Tabela 1. Wartości mierników poziomu rozwoju rolnictwa polskich województw w 2007 r. oraz uporządkowanie województw wg trzech funkcji użyteczności

Województwo	Wartości mierników			Uporządkowanie		
	Funkcje			Funkcje		
	f	U	H	f	U	H
Dolnośląskie	0,41	0,42	0,37	6	6	5
Kujawsko-pomorskie	0,46	0,46	0,41	4	4	3
Lubelskie	0,36	0,39	0,28	10	10	13
Lubuskie	0,21	0,23	0,20	15	15	15
Łódzkie	0,39	0,40	0,36	8	8	7
Małopolskie	0,36	0,37	0,32	11	11	10
Mazowieckie	0,48	0,48	0,40	3	3	4
Opolskie	0,56	0,55	0,44	2	2	2
Podkarpackie	0,20	0,23	0,17	16	16	16
Podlaskie	0,39	0,41	0,31	7	7	11
Pomorskie	0,36	0,36	0,34	12	13	8
Śląskie	0,38	0,39	0,32	9	9	9
Świętokrzyskie	0,24	0,26	0,22	14	14	14
Warmińsko-mazurskie	0,42	0,43	0,36	5	5	6
Wielkopolskie	0,61	0,60	0,53	1	1	1
Zachodnio-pomorskie	0,34	0,36	0,30	13	12	12
Współczynniki korelacji:	r_p(fU)	r_p(UH)	r_p(fH)	r_s(fU)	r_s(UH)	r_s(fH)
Pearsona (r _p)/Spearmana (r _s)	0,999	0,966	0,975	0,997	0,918	0,929

f i U – liniowa i nieliniowa funkcja użyteczności, H – znormalizowana funkcja Hellwiga.

Źródło: Obliczenia własne

W ostatnim wierszu tabeli 1 podane zostały współczynniki korelacji Pearsona i rang Spearmana (odpowiednio), pomiędzy wektorami rezultatów. Np. $r_p(\mathbf{fU})$ oznacza współczynnik korelacji Pearsona między wektorami $\mathbf{f} = (f(\mathbf{z}_1), f(\mathbf{z}_2), \dots, f(\mathbf{z}_{16}))$ i $\mathbf{U} = (U(\mathbf{z}_1), U(\mathbf{z}_2), \dots, U(\mathbf{z}_{16}))$, natomiast $r_s(\mathbf{UH})$ oznacza współczynnik rang Spearmana między wektorami $\mathbf{U} = (U(\mathbf{z}_1), U(\mathbf{z}_2), \dots, U(\mathbf{z}_{16}))$ i $\mathbf{H} = (H(\mathbf{z}_1), H(\mathbf{z}_2), \dots, H(\mathbf{z}_{16}))$.

W tabeli 2. pokazano podstawowe charakterystyki opisowe mierników syntetycznych poziomu rozwoju rolnictwa w roku 2007.

Tabela 2. Charakterystyki opisowe dla syntetycznych mierników województw w Polsce, według trzech funkcji użyteczności f , U i H .

Charakterystyki opisowe	Funkcje		
	f	U	H
Średnia arytmetyczna	0,39	0,40	0,33
Mediana	0,38	0,40	0,33
Rozstęp	0,41	0,37	0,36
Odchylenie stand.	0,11	0,10	0,09
Wsp. zmienności	28%	25%	27%
Wsp. asymetrii	0,16	0,06	0,16

Źródło: Obliczenia własne

Otrzymane wyniki wskazują, że mierniki syntetyczne, otrzymane przy pomocy różnych metod niewiele różnią się między sobą. Na ogół, mierniki (względne), otrzymane za pomocą funkcji nieliniowej U , mają odpowiednie wartości większe, niż użyteczności uzyskane za pomocą funkcji liniowej f , które z kolei przyjmują wartości większe, niż mierniki uzyskane za pomocą znormalizowanej funkcji Hellwiga. Łączy się to również z faktem, że odpowiadająca miara zmienności dla funkcji U jest najmniejsza.

Analiza miar zmienności wskazuje, że w Polsce pod względem poziomu rozwoju rolnictwa występuje znaczne zróżnicowanie województw.

Ocena asymetrii rozkładów wartości syntetycznych mierników województw pokazuje, że w badanym okresie rozkład możemy określić jako symetryczny, co oznacza skupienie badanych jednostek przy średnich wartościach cechy. Symetryczny rozkład wartości mierników wskazuje na to, że w 2007 roku stosunkowo mało województw osiągało wartości powyżej lub poniżej średniego poziomu rozwoju rolnictwa.

Na podstawie wyznaczonych użyteczności województw, w następnym kroku dokonano podziału województw na 4 klasy, charakteryzujące się zbliżonym poziomem rozwoju rolnictwa. Podstawą podziału województw na klasy były wartości ich mierników rozwoju. Wykorzystano w tym celu szeregi przedziałowe rozdzielcze, w których rozpiętości przedziałów klasowych były równe w przybliżeniu jednej czwartej rozstępu. Tabela 3 przedstawia grupowanie województw przy trzech sposobach obliczania mierników. Zestawienie w tabeli 3, podziału województw na

cztery typologiczne grupy pokazuje, że Polska w 2007 roku była krajem zróżnicowanym ze względu na poziom rolnictwa [por. Zegar J. 2003], a wybór miernika syntetycznego miał niewielki wpływ na uzyskane klasyfikacje.

Tabela 3. Klasyfikacja województw według dwóch funkcji użyteczności

Województwo	Funkcja		
	f	U	H
Wielkopolskie	I	I	I
Opolskie	I	I	II
Mazowieckie	II	II	II
Kujawsko-pomorskie	II	II	II
Warmińsko-mazurskie	II	II	II
Dolnośląskie	II	II	II
Podlaskie	III	III	III
Łódzkie	III	III	II
Śląskie	III	III	III
Lubelskie	III	III	III
Małopolskie	III	III	III
Zachodnio-pomorskie	III	III	III
Pomorskie	III	III	III
Świętokrzyskie	IV	IV	IV
Lubuskie	IV	IV	IV
Podkarpackie	IV	IV	IV

Źródło: Obliczenia własne

Analizując na podstawie tabeli 1 uporządkowanie województw według poziomu użyteczności można wnioskować, że bez względu na rodzaj funkcji, w 2007 roku na początku rankingu znajduje się województwo wielkopolskie, natomiast ostatnie miejsce województwo podkarpackie.

Do I grupy, na ogół, zaliczają się dwa województwa - wielkopolskie i opolskie, odznaczające się najwyższym poziomem rozwoju rolnictwa w Polsce. Grupa II obejmuje cztery województwa: mazowieckie kujawsko-pomorskie, warmińsko-mazurskie i dolnośląskie, które charakteryzują się stosunkowo wysokim (powyżej średniej) poziomem rozwoju rolnictwa. Grupa III obejmuje siedem województw: podlaskie, łódzkie, śląskie, lubelskie, małopolskie, zachodnio – pomorskie i pomorskie, o niskim (na ogół poniżej poziomu przeciętnego) poziomie rozwoju rolnictwa. W IV grupie znalazły się (wg wszystkich rozpatrywanych metod) trzy województwa: świętokrzyskie, lubuskie i podkarpackie, co wskazuje na to, że w 2007 roku były one regionami o najniższym poziomie rozwoju rolnictwa w Polsce. Analizując użyteczności oparte na standaryzowanym mierniku Hellwiga, do pierwszej grupy należy tylko jedno województwo - wielkopolskie, do drugiej dodatkowo opolskie i łódzkie, w trzeciej grupie znalazło się sześć województw.

Na podstawie tych wyników można wnioskować, że zmiana funkcji użyteczności liniowej na nieliniową nie zmieniła grupowania województw. Natomiast

zmiana z funkcji liniowej na standaryzowaną funkcję Hellwiga powoduje w klasyfikacji niewielkie zmiany.

Analizując grupowania województw można zauważyć, że wśród wielu metod nie ma żadnej metody, której przewaga nad innymi byłaby uznana.

WNIOSKI

Podane w pracy wyniki badań jednoznacznie świadczą, że w 2007 roku występowało duże zróżnicowanie poziomu rozwoju rolnictwa polskich województw. Ze względu na badane zjawisko, najlepszym województwem okazało się województwo wielkopolskie, najgorszym natomiast podkarpackie.

Analiza otrzymanych rezultatów pokazuje, że każdy z rozważanych mierników może równoważnie, dobrze służyć, jako narzędzie do oceny stanu poziomu rozwoju rolnictwa polskich województw. Wyniki wskazują, że występuje tu duża stabilność klasyfikacji i grupowania obiektów, ze względu na zastosowane sposoby obliczania mierników syntetycznych. Choć wyznaczenie mierników przy pomocy funkcji liniowej f są najprostsze, to nie ma to jednak większego znaczenia, jeśli wziąć pod uwagę, że obliczenia dokonywane są w obecnych czasach przy użyciu komputerów.

LITERATURA

- Allen R. G. D. (1964) *Ekonomia matematyczna*, PWN, Warszawa.
- Bartosiewicz S. (1976) Propozycja metody tworzenia zmiennych syntetycznych, *Prace Naukowe AE we Wrocławiu*, nr 84, Wrocław.
- Binderman A. (2004) Przestrzenne zróżnicowanie potencjału rolnictwa w Polsce w latach 1989-1998, *R. N. Rolniczych, Seria G., T.91, Z.1*, str. 51, „Wieś Jutra”, Warszawa.
- Binderman A. (2005a) Klasyfikacja polskich województw według poziomu rozwoju rolnictwa, *Roczniki Nauk Rolniczych, Seria G., T.92, Z.1*, str. 42, *Wieś Jutra*, Warszawa.
- Binderman A. (2005b) O problemie wyboru wzorca przy badaniu przestrzennego zróżnicowania potencjału rolnictwa w Polsce, *Metody ilościowe w badaniach ekonomicznych – V*, Warszawa, str. 46.
- Binderman A. (2006a) Wykorzystanie funkcji użyteczności do badania przestrzennego zróżnicowania rolnictwa, *R. N. SERiA, Tom VIII, Zeszyt 5*, Warszawa-Poznań, str. 5.
- Binderman A. (2006b) Klasyfikacja obiektów oparta na dwóch wzorcach, *EiOGŻ, Zeszyty Naukowe SGGW*, nr 60, Warszawa, str. 25.
- Binderman A., Krawiec M. (2006) Regionalne zróżnicowanie poziomu rozwoju rolnictwa w Polsce w latach 2002-2005, *Potencjał rozwojowy obszarów wiejskich w aspekcie wstąpienia Polski do Unii Europejskiej*, Szczecin, s. 39.

- Binderman A. (2007) Wielowymiarowa analiza regionalnego zróżnicowania rolnictwa w Polsce, praca doktorska, SGGW, Warszawa.
- Binderman A. (2008) Zastosowanie liniowej i nieliniowej funkcji użyteczności do badania poziomu rolnictwa w Polsce, *Metody ilościowe w badaniach ekonomicznych – IX*, wyd. SGGW, Warszawa, str. 29-38.
- Binderman A. (2009) Dynamika rozwoju rolnictwa w Polsce po akcesji do Unii Europejskiej, *Roczniki Nauk Rolniczych, SERiA, Tom XI, Zeszyt 3*.
- Borkowski B, Szczesny W. (2002) Metody taksonomiczne w badaniach przestrzennego zróżnicowania rolnictwa, *R. N. Rolniczych, Seria G., T.89, Z.2*.
- Borkowski B, Dudek H., Szczesny W. (2004) *Ekonometria. Wybrane zagadnienia*, PWN, Warszawa.
- Cieślak M. (1993) Ekonomiczne zastosowanie mierników syntetycznych ze zmiennym wzorcem, [w:] *Przestrzenno-czasowe modelowanie i prognozowanie zjawisk gospodarczych*, AE, Kraków.
- Hellwig Z. (1968) Zastosowanie metody taksonomicznej do typologicznego podziału krajów ze względu na poziom ich rozwoju oraz zasoby i strukturę kwalifikowanych kadr, „*Przegląd Statystyczny*”, z. 4.
- Kukuła K. (2000) *Metoda unitaryzacji zerowanej*, PWN, Warszawa.
- Malina A. (2004) Wielowymiarowa analiza przestrzennego zróżnicowania struktury gospodarki Polski według województw, AE, Seria Monografie nr 162, Kraków.
- Młodak A. (2006) *Analiza taksonomiczna w analizie regionalnej*, Difin, Warszawa.
- Nowak E. (1990) *Metody taksonomiczne w klasyfikacji obiektów społeczno-gospodarczych*, PWE, Warszawa.
- Panek E. (2000) *Ekonomia matematyczna*, Akademia Ekonomiczna, Poznań.
- Strahl D., Walesiak M. (1996) Normalizacja zmiennych w skali przedziałowej i ilorazowej w referencyjnym systemie granicznym, *Seria: Taksonomia, z. 3, Sekcja Klasyfikacji i Analizy Danych*, Wrocław – Kraków - Jelenia Góra.
- Strahl D., Walesiak M. (1997) Normalizacja zmiennych w granicznym systemie referencyjnym, „*Przegląd Statystyczny*”, z. 1.
- Walesiak M. (2006) *Uogólniona miara odległości w statystycznej analizie wielowymiarowej*, wyd. AE Wrocław.
- Zegar J. (2003) *Zróżnicowanie regionalne rolnictwa*, GUS, Warszawa.
- Zeliaś A. (2000) *Taksonomiczna analiza przestrzennego zróżnicowania poziomu życia w Polsce w ujęciu dynamicznym*, Kraków.

Dependence of development level analysis of Polish agriculture on choice of synthetic measures

Summary: In the present paper, development level analysis of Polish agriculture is investigated. The investigation is based on the 2007 year. The applied methods use three different utility functions as the preference indicators. These indicators were applied to arrangement of Polish voivodeships.

The given utility function has such property that two considered objects have an identical utility if their distances from two different fixed model objects are equal.

Keywords: agriculture development level, synthetic measures, utility functions, normalized Hellwig's function, voivodeships class division.

THE ANALYSIS OF UNEMPLOYMENT VARIATION MEASURES IN THE EU MEMBER STATES

Mariusz Borawski

Department of Information Technology, WPUT
mborawski@wi.ps.pl

Abstract: The issue of unemployment may be examined from many aspects. It is not only unemployment rate itself, but also the distribution of unemployment in a given country that are of major importance. To conduct such research, it is necessary to analyze measures of unemployment variation. The article presents unemployment variation measures together with their unique character.

Keywords: unemployment, pseudo-variance

INTRODUCTION

Unemployed represent a certain percentage in the total of population in every free-market economy. It is a phenomenon natural for every economy. It becomes a problem only when it increases considerably [Kwiatkowski 2007]. Unemployment is found both in rich and poor countries. Hence, in most countries there are institutions the aim of which is to combat and prevent the unemployment [Layard et al. 1991]. In many countries, unemployment has a regional character. There are regions that are particularly affected by unemployment [Horodelski and Sadowska-Snarska 2003]. Therefore, the research must involve analyses aimed at comparing not only unemployment rates, but also their spatial and temporal variation.

Temporal and spatial variation may be described with the use of distribution. To carry out more comprehensive analyses, it is necessary to define arithmetic operations on distribution. Convolution [Feller 2006] is the operator of addition, yet defining the operation contrary to convolution, i.e. equivalent of subtraction,

poses a major problem¹. If one adopts definition of subtraction that is in accordance with axioms of algebra, distribution may not be a result in some cases. In practice, this leads to a situation when it is not plausible to employ methods based on algebra. The only solution is to generalize the notion of distribution, i.e. introduce the notion of pseudo-distribution [Borawski 2008]. Subtracting of two distributions does not have to result in a distribution but always in a pseudo-distribution, due to which, on the basis of pseudo-distributions set, one may introduce arithmetic operations that are in keeping with axioms of algebra. Pseudo-distribution may be classified as an element of the full set, in which all the possible distributions and their inverse elements are included. Proper definition of the pseudo-distribution is very difficult because of the fact, that, for most opposite elements, their counterpart in real world cannot be found. Similar situation occurs when the arithmetic calculations on the size of the set are conducted. If the difference in sets' size is examined and it is negative, it is not possible to determine the counterpart of negative number, because the sets with negative number of elements do not exist. On the other hand, such negative numbers are indispensable. Similar problem arises while defining the actions for parameters describing the distribution, particularly those parameters that measure variation (e.g. standard deviation, variance, range). To conduct a more comprehensive analysis, it is necessary to define arithmetic operations in a proper way and make generalizations that may consist in generalizing the variance as pseudo-variance that may take negative values. For formal reasons, most methods cannot be used without making such a generalization. Due to the fact that certain axioms are not satisfied, algebra does not allow to make calculations for correlation variance, create econometric models, etc. On the contrary, it is possible in the case of pseudo-variance as it satisfies these axioms. Pseudo-variance can be defined as an element of the full set, to which all possible variances and their inverse elements are included.

Relativity measures determine the variation of a given factor. As far as unemployment is concerned, these measures determine how unevenly distributed unemployment is in a given region. According to data derived from Eurostat, unemployment rate among people aged 25 and more amounted to 8.3% in 2007 in Poland. However, this does not indicate that unemployment rate is the same in the entire Poland. Unemployment rate can be calculated individually for each administrative unit. The comparison between these rates enables one to determine the extent to which they are different, which can be defined with the use of variation measures.

Standard deviation, describing the distribution, is one of the most frequently used variation measures. In the case of unemployment, the distribution of rates is the case. It determines the broadness of distribution, i.e. how much the rate fluctuates around its mean value. High and low rates may be considered favourable de-

¹ Operation opposite to convolution has been discussed in the following publications: Mikusiński [1953], Mareš [1989], Mareš [1994].

pending on the rate under consideration. As for unemployment, high standard deviation can be considered unfavourable as it indicates that unemployment distribution is extremely uneven in the area investigated.

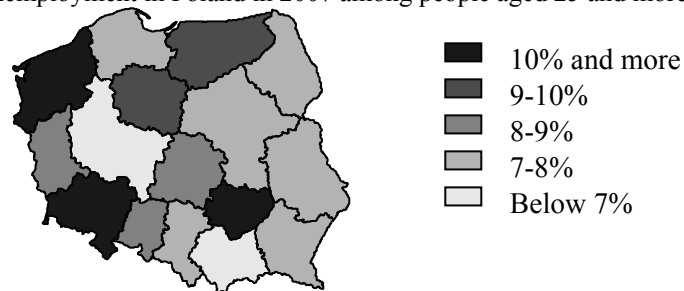
Furthermore, high standard deviation will translate into uneven supply and demand for labour in different regions of a particular administrative unit. On the one hand, this will lead to migration of population from regions characterized by higher unemployment to regions where it is lower. On the other hand, working places will “migrate” to regions where unemployment is higher in order to find employees whose pay demands are low.

Actions aimed at reducing the disproportion in unemployment may have different character than actions aimed at bringing down the unemployment itself. The disproportion may be reduced via improving transport infrastructure. The construction of motorways and dual clearways enables people to commute to work even from far away places. Motorways increase this distance nearly twofold. Everyday journey from place situated even 150 km away from working place becomes real, which can considerably reduce disproportion in employment in communes and counties.

EVENNESS OF SPATIAL DISTRIBUTION OF UNEMPLOYMENT

The fact that unemployment rate amounts to 8.3% in Poland does not indicate that unemployment is distributed evenly in the entire country. Figure 1 shows large disproportion among particular voivodships. In eastern Poland, unemployment amounts to about 7-8%, yet unemployment rates in western Poland are characterized by considerable disproportion. Wielkopolskie voivodship, where unemployment rate is low, neighbours Zachodniopomorskie voivodship and Dolnośląskie voivodship where the rates are extremely high.

Figure 1. Unemployment in Poland in 2007 among people aged 25 and more



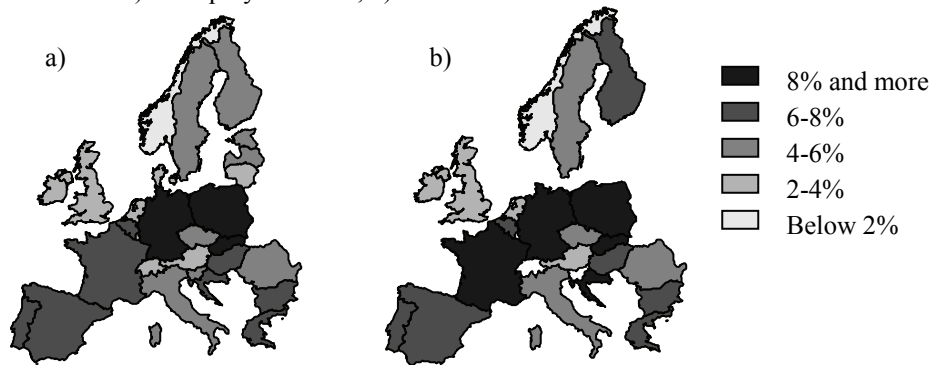
Source: own elaboration based on data derived from Eurostat

Unemployment rate for a particular country is calculated for the entire area without taking rates for particular administrative units into account. To determine spatial distribution of unemployment, one must use parameters describing the distribution of unemployment. Hence, one must calculate unemployment rates for

particular administrative units and then calculate parameters describing the distribution of unemployment in the entire country. Mean value and standard deviation are the most important parameters describing the distribution.

Mean value is an equivalent of unemployment rate in the entire country, yet it is not the same (which is shown in Figure 2). It can be noticed that mean unemployment rates calculated for particular countries at NUTS2 level and unemployment rates recorded in these countries are very much similar. The only difference lies in their interpretation. Mean value defines the mean value of the rate for NUTS2 administrative units and hence it refers directly to the value of the rate. Difference in values of both parameters results from the fact that mean value refers to administrative units as objects without paying attention to their scale (in the case of unemployment – number of residents).

Figure 2. Unemployment in Europe in 2007 among people aged 25 and more: a) unemployment rate, b) mean values of rates for NUTS2 administrative units

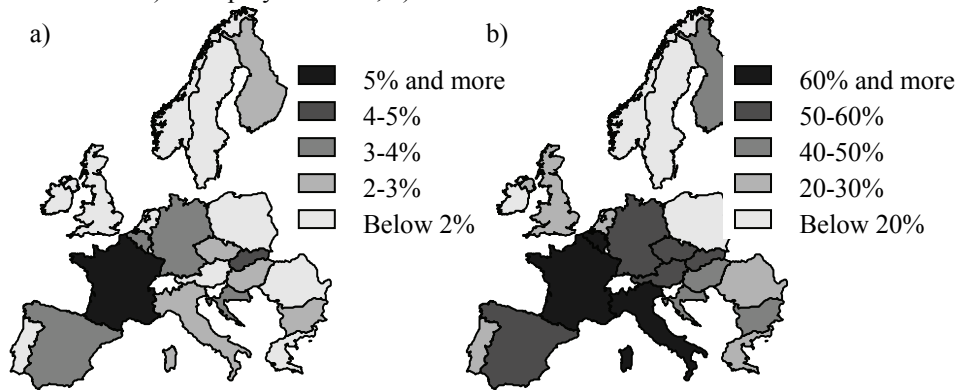


Source: own elaboration based on data derived from Eurostat

Standard deviation is the other important parameter. It defines the extent to which value fluctuates around mean value. The higher the deviation, the greater the fluctuation. Standard deviation is expressed in units of values that it describes. Standard deviation of unemployment rates in particular countries may be calculated when unemployment rates for administrative units at different levels are known (e.g. NUTS2, NUTS3). Hence, it defines the variation of unemployment rate at a given level.

Figure 3a shows the variation of unemployment rates in countries determined for NUTS2 level, which corresponds to Polish voivodships. Low values of standard deviation are the most favourable. Like Ireland, Norway or Sweden, Poland is among countries with low standard deviation of unemployment rate, which implies that the unemployed are distributed quite evenly in particular voivodships (compared to other countries).

Figure 3. Unemployment in Europe in 2007 among people aged 25 and more:
a) unemployment rate, b) mean values of rates for NUTS2 administrative units



Source: own elaboration based on data derived from Eurostat

Standard deviation is expressed in units of rate for which it has been calculated. Such a way of expressing standard deviation does not always illustrate the situation well. Deviation with two units should be interpreted in a different way when mean value of the rate has four units, and in a different way when it has two hundred units. In the former case, variation of value should be considered extremely high, whereas in the latter case – very low. Therefore, value of standard deviation should always be related to mean value. To make the interpretation of standard deviation independent of mean value one can multiply them. The product of such an operation is variation coefficient, usually expressed as the percentage of mean value.

Figure 3b shows variation coefficient values for particular countries. It can be noticed that many affluent countries are characterized by high scatter coefficients exceeding 50%, which indicates that there is large disproportion among particular administrative units as far as unemployment rates are concerned. This state of affairs often stems from regional differences, just like in Italy (North-South) or Germany (in the past, the Federal Republic of Germany and the German Democratic Republic).

CORRELATION AMONG MEASURES OF RATES' VARIATION

Correlation formula may be derived in many ways [Borawski 2007]. Vector calculus is one of options. However, this requires defining vector space for variation measures. In the case of variance and standard deviation, it is necessary to make a generalization that allows negative values. The coordinates of vectors are ordered pairs that include mean value and generalized standard deviation (or mean value and variance generalization). These pairs can be added and multiplied by real value in accordance with universally accepted rules underlying arithmetic opera-

tions on mean values, variance and standard deviation [Jaworski 1979]. Thus, vectors defined for these pairs can be added and multiplied by scalar. On the basis of vector space defined in such a way, one may derive variance correlation and standard deviation formula which is analogical to mean value formula.

As far as arithmetic operations on variation measures are concerned, operations on standard deviation and variance can be distinguished. Arithmetic operations on standard deviation refer to random variables entirely dependent on one another, and operations on variance – to entirely independent ones. In practice, real result is somewhere in between values calculated for standard deviation and variance, which stems from the fact that random variables are usually partially dependent.

Table 1. Correlation between unemployment rates and other rates

	Unemployment rates by sex and age, at NUTS levels 1, 2 and 3 (%)			Long-term unemployment (12 months and more), at NUTS levels 1 and 2 (1000; %)		
	Mean value	Standard deviation	Variance	Mean value	Standard deviation	Variance
Unemployment rates by sex and age, at NUTS levels 1, 2 and 3 (%)	1	1	1	0.59	0.04	0.21
Long-term unemployment (12 months and more), at NUTS levels 1 and 2 (1000; %)	0.59	0.04	0.21	1	1	1
Economic activity rates by sex and age, at NUTS levels 1 and 2 (%)	-0.56	0.46	0.50	-0.53	-0.10	0.03
Average number of usual weekly hours of work in main job (full-time), at NUTS levels 1 and 2 (hours)	0.75	-0.04	-0.09	0.59	0.18	0.21

Source: own elaboration based on data derived from Eurostat

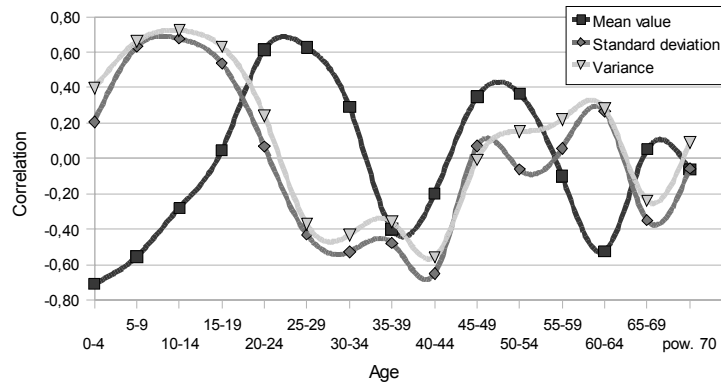
Table 1 shows correlations determined. It can be noticed that unemployment rate is strongly correlated with the average number of working hours, which results from the fact that the more hours employees spend at work, the fewer workers are needed for accomplishing a given task. In consequence, employers offer employment to smaller number of people.

Unemployment rate is negatively correlated with people's economic activity. Thus, the greater this activity, the lower the unemployment. At the same time, the variation of this activity is positively correlated with unemployment variation. Hence, it may be concluded that in regions characterized by great economic activ-

ity, unemployment is either very high or very low. In the case of variation measures, positive and negative values of correlation cannot be interpreted explicitly. It can only be stated that correlation is found, but it is impossible to determine its character.

Figure 4 shows the correlation between unemployment rate and percentage share of people from particular age groups in the total of population. It can be noticed that mean value is strongly and negatively correlated with the number of people aged 0-4, which can be justified on the grounds of the fact that parents decide have children when their financial situation is stable, i.e. when one of them has regular work. On the other hand, one parent can take maternity leave and hence is not considered unemployed.

Figure 4. Correlation between unemployment rate and percentage share of people from particular age groups in the total of population



Source: own elaboration based on data derived from Eurostat

The maximum is reached for people aged 20-29, which stems from the fact that they are just entering labour market. As they have no professional experience, their situation on the market is worse, they find it difficult to find a job and remain unemployed much longer. Having gained certain experience, at the age of 35-40 they are sought by employers and thus unemployment rate is the lowest among them. As they grow old, they are less and less efficient and able, and so are less attractive as potential workers, which is reflected in higher unemployment rate. Finally, people aged 60-64 are in a way protected, due to which employers cannot dismiss them so frequently. On the other hand, they can take early retirement instead of becoming unemployed. As a result, unemployment rate is lower in this age group.

CONCLUSION

The analysis of variation may provide additional information concerning the spatial distribution of unemployment, which allows to determine how evenly it is distributed among administrative units under consideration. Furthermore, it is plausible to determine the correlation between the variation measures of different rates and unemployment. On the basis of the research conducted, it can be stated that spatial distribution of unemployment is correlated with economic activity of unemployment and the number of people aged 5-15.

LITERATURE

- Borawski M. (2007) Rachunek wektorowy w przetwarzaniu obrazów, Wydawnictwo Uczelniane Politechniki Szczecińskiej, Szczecin
- Borawski M. (2008), Pseudorozkład jako uogólnienie pojęcia rozkładu, Przegląd Statystyczny, Vol. 55, No. 3, str. 71-85
- Feller W., (2006) Wstęp do rachunku prawdopodobieństwa, PWN, Tom I, Warszawa
- Horodelski R., Sadowska-Snarska C. (2003) Red. Rynek pracy w Polsce na progu XXI wieku. Aspekty makroekonomiczne i regionalne. IPiSS Warszawa
- Jaworski J. (1979) Matematyczne podstawy metrologii, WNT, Warszawa
- Kwiatkowski E., (2007) Bezrobocie Podstawy teoretyczne, PWN, Warszawa
- Layard R., Nickell S., Jackman, (1991) Unemployment. Macroeconomic Performance and the Labour Market, Oxford University Press, Oxford
- Mareš M., (1989) Addition of rational fuzzy quantities: Convolution approach, Kybernetika nr 25,
- Mareš M., (1994) Computation over Fuzzy Quantities, CRC Press, Boca Raton
- Mikusiński J., (1953) Rachunek operatorów, Polskie Towarzystwo Matematyczne, Warszawa

Analiza miar zmienności bezrobocia w krajach Unii Europejskiej

Streszczenie: Problem bezrobocia można rozpatrywać w wielu aspektach. Interesującym czynnikiem jest nie tylko sam wskaźnik bezrobocia, ale także rozłożenie liczby bezrobotnych na obszarze danego kraju. Aby wykonywać tego rodzaju badania konieczne jest przeprowadzanie analiz miar zmienności bezrobocia. W artykule przedstawiono badanie miar zmienności bezrobocia z uwzględnieniem specyfiki tych miar.

Słowa kluczowe: bezrobocie, pseudowariancja.

TAXONOMIC ANALYSIS OF AREAL STRUCTURE OF FARMS IN POLAND WITHIN 2000 - 2007

Jadwiga Bożek, Karol Kukula, Danuta Bogocz

Katedra Statystyki Matematycznej, Uniwersytet Rolniczy w Krakowie
e-mail: rrbozek@cyf-kr.edu.pl

Abstract: The paper presents the results of the grouping procedure applied to voivodships in Poland with respect to areal structure of farms in 2000 and 2007. The investigation was based on EU classification of farms and the results compared with areal structure presented in reports of the Central Statistical Office. Dynamics of changes in individual groups was also determined.

Keywords: areal structure of farms, voivodship grouping, level of structural changes.

INTRODUCTION

The paper is the continuation of earlier investigations concerning areal structure of farms in Poland [Bożek 1994, Bożek 2007,], and in particular – its spatial differentiation within the period of 1976-2006 [Bożek 1994, Bożek 2000, Bożek 2008, Kukula 2005]. The results show that there is a significant differentiation among regions within country area and this seems to be a permanent situation. In 2006 there were 5 groups of voivodships with similar areal structure of farms and high dispersion between groups. [Bożek 2008]. At the same time, one must remember that the research was based on the structure consisting of the following areal intervals: 1-2 h, 2-5 h, 5-10 h, 10-15 h, 15-20 h, 20-30 h, 30-50 h, 50-100 h and 100 h and more. Classification of farms in EU terminology is as follows: very small:– up to 5 h, small – 5-10 h, average – small – 10-20 h, average large – 20-30 h, large 30-50 h and very large – above 50 h. One can easily notice that the two structures are different. Because of the fact that the change of structure elements will produce new grouping results [Bożek 1994, Wysocki 1989], the question arises what spatial differentiation of agrarian structure in Poland is like if we take into account farm classification according to the one of EU.

The grouping procedure carried out in the paper concern the data covering the period of 2000 and 2007. The results, obtained as the consequence of application of some chosen taxonomic methods, show regional differentiation of areal structure of Polish agriculture within the period of 2000 – 2006 with its spatial changes and allow for determination of dynamics and direction of changes. They also form the basis for estimation whether the regional differences in areal structure increase or decrease.

INVESTIGATION METHOD

The analysis was conducted on the basis of the data obtained from the Central Statistical Office, concerning the number of farms of the area more than one hectare of arable land in Poland in 2000 and 2007 for voivodships.

The classification of objects (voivodships) into groups of similar structure was carried out by vector elimination method [Chomałowski, Sokołowski, 1978] and then best choice method [Wysocki, Wagner, 1989]. The measure of structures differentiation between the two objects i, j was adopted from Kukuła [Kukuła 1989]:

$$\varepsilon_{ij} = \frac{\sum_{l=1}^k |a_{il} - a_{jl}|}{2} \quad (i, j=1, \dots, k); \quad (1)$$

where a_{il} - share of number of farms from areal group l in the total number of farms in voivodship i .

In order to establish the level of differentiation between typological groups average differentiation measures between groups were calculated [Nowak 1990].

The formula applied to estimate the degree of changes that had taken place within the period from $t - \tau$ to t was as follows: [Kukuła 1989]:

$$v_{t,t-\tau} = \frac{\sum_{k=1}^r |\alpha_{tk} - \alpha_{(t-\tau)k}|}{2} \quad (2)$$

where α is the structure of shares investigated in time $t=0,1,\dots,n$, consisting from r elements, i.e.. the matrix is given $[\alpha_{tk}]_{(t=0,1,\dots,n; k=1,\dots,r)}$.

The measure (2) takes values from the interval $[0;1]$. Its high value indicates that the changes that had taken place within the period $t - \tau$ to t are significant. In particular, v_{n0} can be applied to compare the structure from the zero period ($t = 0$) with the structure from the end period ($t = n$).

RESULTS

The average differentiation of areal structure between voivodships in 2000 reached the level of $\beta = 0,2504$ and in 2007 decreased to $\beta = 0,2310$. Application of methods described in previous section to the data from 2007 resulted in classification into four groups of voivodships (table 1), which at the same time means that in Poland in 2007 there were four types of areal structure of farms with significant differences between each two of them. The important fact is that the fourth group consists only from one voivodship - lubelskie.

Table 1. Groups of voivodships with similar areal structure of farms in Poland in 2007.

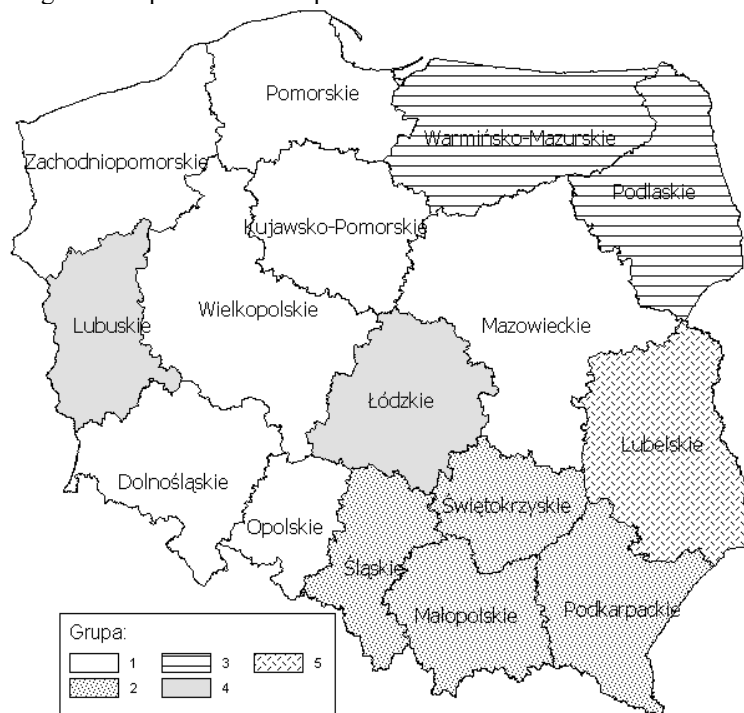
Voivodship groups	Farms with arable land of the area [in hectares]					
	1-5	5-10	10-20	20-30	30-50	50 and more
	in %					
I łódzkie mazowieckie lubuskie wielkopolskie dolnośląskie opolskie kujawsko-pomorskie pomorskie	47,6	23,9	17,8	5,0	3,3	2,3
II małopolskie śląskie podkarpackie świętokrzyskie	78,5	15,8	4,3	0,8	0,4	0,3
III podlaskie zachodniopomorskie warmińsko-mazurskie	34,5	19,8	24,7	10,0	6,6	4,3
IV Lubelskie	55,6	27,7	12,4	2,5	1,3	0,5
POLAND	57,4	22,2	13,5	3,6	2,1	1,2

Source: Own calculations

This grouping differs from the one carried out for nine element structure, the results of which are shown in table 2 and figure 1. The difference is in group size. However, if we compare composition of groups it turns out that the two grouping have very much in common. In both cases voivodships of south east Poland form one group (II). In both grouping lubelskie voivodship forms one group. The differences concern only three voivodships: zachodnio-pomorskie,

lubuskie and łódzkie. It is a consequence of the fact that these voivodships differ significantly from other voivodships from the groups they belong in grouping of six element structure regarding shares of farms of the area up to 10 hectares and therefore in grouping of nine element structure they are assigned to different groups. Lubuskie and łódzkie voivodships constitute a separate group (IV) while zachodnio-pomorskie voivodship belongs to the largest – first – group.

Fig. 1. Groups of voivodships of similar areal structure of farms in 2006



Source: own investigations.

Table 2. Areal structure of farms in distinguished groups of voivodships in Poland in 2006

Voivodship groups	Farms of the area of arable land in hectares								
	1-2	2-5	5-10	10-15	15-20	20-30	30-50	50-100	>100
	W %								
Group I	20,1	25,8	22,6	12,5	6,4	5,7	3,6	2,1	1,2
Group II	34,4	44,1	16,1	3,1	0,9	0,7	0,4	0,2	0,1
Group III	13,0	17,5	21,3	17,2	11,4	10,2	5,9	2,4	0,9
Group IV	14,2	26,3	34,2	14,2	4,6	3,1	1,6	1,3	0,8
Group V	17,7	38,1	28,1	9,2	3,1	2,4	1,0	0,4	0,1

Source: [Bożek 2008]

Another task that should be taken into account is to determine whether the classification is stable in time. In order to answer this question the grouping procedure was carried out for the data from 2000. It turns out that composition of groups is almost the same as in 2007. The exception is zachodniopomorskie voivodship, that in 2000 was in group I and in 2007 – in group III.

This was the result of significant changes that had taken place in the structure of this voivodship, due to which it became closer the structure of voivodships from group III. Zachodniopomorskie voivodship is the second – with respect to the degree of changes within the period being investigated, which is reflected in the level of measure of structural changes (table 3). The largest changes were observed in lubuskie voivodship (0,113), then zachodniopomorskie (0,088), warmińsko-mazurskie (0,083), podlaskie (0,082) and opolskie (0,076).

Tab.3. Measure of structural changes $v_{2007,2000}$ in individual voivodships

Voivodship	$v_{2007,2000}$
Lubuskie	0,113
Zachodniopomorskie	0,088
Warmińsko-mazurskie	0,083
podlaskie	0,082
Opolskie	0,076
mazowieckie	0,056
Pomorskie	0,054
dolnośląskie	0,052
kujawsko-pomorskie	0,040
łódzkie	0,040
wielkopolskie	0,027
Lubelskie	0,026
świętokrzyskie	0,022
śląskie	0,018
Małopolskie	0,012
podkarpackie	0,011
POLSKA	0,024

Source: [Bożek 2008] and own calculations

According to the definition (2) the coefficients show that the total changes (differences in absolute values) in the structure of these voivodships are oscillating from 22,6 (lubuskie) to 15,2 pp. (opolskie). The lowest rate is observed in case of podkarpackie, małopolskie, śląskie and świętokrzyskie voivodship, where the total changes do not exceed 5 pp. Comparison of structure of groups obtained in

grouping procedure in 2000 and 2007 (table 4) show some trends that are characteristic for agrarian structure in Poland.

Table 4. Areal structure in distinguished groups in 2000 and 2007.

Group		1-5 h	5-10 h	10-20 h	20-30 h	30-50 h	More than 50 h
I	2000	44,6	24,4	20,3	5,8	2,8	2,1
	2007	47,6	23,9	17,8	5	3,3	2,3
II	2000	79,2	16,2	3,5	0,4	0,5	0,2
	2007	78,5	15,8	4,3	0,8	0,4	0,3
III	2000	26,3	22,4	32,5	11,1	5	2,6
	2007	34,5	19,8	24,7	10	6,6	4,3
IV	2000	54,1	29,9	12,8	2,3	0,7	0,2
	2007	55,6	27,7	12,4	2,5	1,3	0,5

Source: own investigations

In group I the share of very small farms increased (from 44,6 to 47,6%) as well as large and very large (in total from 4,9 to 5,6%), while the share of small farms and average farms decreased (from 24,4 to 23,9% in total from 26,1 to 22,8% respectively).

In group II the changes were negligible – small drop of shares of very small and small farms and increase of shares of other areal classes. The largest changes had taken place in group III – significant increase of shares of very small farms (from 26,3 to 34,5%) and large and very large (in total from 7,6 to 10,9%) and decrease in other areal classes.

CONCLUSIONS

1. In Poland in 2007 there were four groups of voivodships with respect to similarity of areal structure of farms (according to EU classification of farms). Such division maintains from 2000 (except for one voivodship - zachodniopomorskie, which within the period being investigated due to significant changes in structure changed its membership). Therefore, regional differentiation has a permanent character.
2. The largest number of small and very small farms can be observed in the group of south east vovodships, i.e.: małopolskie, śląskie, podkarpackie and świętokrzyskie, where very small farms constitute 78,5% of total number and small farms – 15,8%. Farms of the area above 10 hectares constitute a very small percentage (5,8%). In this group the structure changed very little – there had been very slow decrease of shares of small and very small farms and increase in other areal classes..
3. The lowest number of very small land small farms occurs in podlaskie, zachodnio-pomorskie and warmińsko-mazurskie voivodship, where they

constitute 34,5% and 19,8% respectively. Average farms compose 34,7% Chile large and very large farms hale shares equal to 10,9%. In this group the rate of changes is at the highest level – significant increase of shares of very small farms (from 26,3 to 34,5%) as well as large and very large farms (in total from 7,6 to 10,9%) and decrease in other areal classes.

4. Other voivodships (except for lubelskie) compose the largest group. These are: łódzkie, mazowieckie, lubuskie, wielkopolskie, dolnośląskie, opolskie, kujawsko-pomorskie i pomorskie. Unlike other groups the percentages of very small and small farms achieve 47,6% and 23,9% respectively. In his group the percentage of average farms is higher than in other groups – 22,8%, while large and very large farms constitute 5,6%. The changes that take place in this group have the same direction as in group III, but the rate of changes is lower.

LITERATURE

- Bożek J. (1994a) Przestrzenne zróżnicowanie struktury obszarowej rolnictwa indywidualnego, *Wiadomości Statystyczne*, nr 2.
- Bożek J. (1994b) Łańcuchy Markowa w badaniach nad strukturą agrarną w Polsce, materiały na XXIV Colloquium Metodologiczne z Agrobiometrii.
- Bożek J. (2000). Zmiany struktury obszarowej gospodarstw rolnych (według metod taksonomicznych), *Wiadomości Statystyczne*, 2000, nr 9, str.13-20.
- Bożek J. Prognoza zmian w strukturze obszarowej gospodarstw rolnych w Polsce do roku 2015 , *Więś i Rolnictwo*, 2007 nr 2, str.106-116.
- Bożek przestrzenno.Porównanie zróżnicowania przestrzennego struktury agrarnej Polski w latach 2000 i 2006 w *Metody ilościowe w badaniach ekonomicznych*, IX, Wydawnictwo SGGW, Warszawa 2008, str.49-56 .
- Chomątowski S., Sokołowski A. (1978) Taksonomia struktur. *Przegląd Statystyczny*, 2.
- Halamska M. (1988) Struktura agrarna Polski – ciągłość czy zmiana?, *Więś i Rolnictwo*, nr 3.
- Kukuła K. (1989) Statystyczna analiza strukturalna i jej zastosowanie w sferze usług produkcyjnych dla rolnictwa. *Zeszyty Naukowe AE w Krakowie*, Seria specjalna: Monografie, 89, Kraków.
- Kukuła K. (2005) Statystyczna analiza struktury obszarowej gospodarstw rolnych w Polsce. *Agrobiznes 2005*, *Prace Naukowe AE we Wrocławiu*, Wrocław.
- Nowak E., 1990. *Metody taksonomiczne w badaniach społeczno-ekonomicznych*, PWE, Warszawa.
- Wysocki F., Wagner W. (1989) O ustalaniu wartości progowej zróżnicowania struktur z danych empirycznych. *Wiadomości Statystyczne*, 9.

**Taksonomiczna analiza struktury obszarowej gospodarstw rolnych
w Polsce w okresie 2000-2007**

Streszczenie: W pracy przedstawiono wyniki grupowania województw ze względu na podobieństwo struktury obszarowej gospodarstw rolnych w roku 2000 i 2007. Przyjęto strukturę zgodną z klasyfikacją gospodarstw w UE i porównano z wynikami dla struktury podawanej w publikacjach GUS. Określono tendencje przemian w wyodrębnionych grupach.

Słowa kluczowe: struktura obszarowa gospodarstw rolnych, grupowanie województw, stopień zmian strukturalnych

SZACOWANIE I MODELOWANIE TFP W PRZEMYSŁE POLSKIM NA PODSTAWIE DANYCH PANELOWYCH¹

Barbara Dańska-Borsiak

Katedra Ekonometrii Przestrzennej UŁ
e-mail: danska@uni.lodz.pl

Streszczenie: W artykule przedstawiono próbę oszacowania łącznej produktywności czynników produkcji (TFP) według działów sekcji „przetwórstwo”, oraz określenia czynników determinujących jej kształtowanie. Do oszacowania wartości TFP zastosowano dwie alternatywne metody, bazujące na funkcji produkcji Cobba–Douglasa. Następnie skonstruowano i oszacowano dynamiczny model panelowy, opisujący kształtowanie się TFP w działach. Zmienną objaśnianą były wartości oszacowane w pierwszym etapie analizy. Do estymacji zastosowano metody bazujące na GMM.

Słowa kluczowe: dynamiczny model panelowy, funkcja produkcji, łączna produktywność czynników produkcji (TFP), systemowy estymator GMM.

WSTĘP

Łączna produktywność czynników produkcji (TFP) jest definiowana jako produkt, który może być wytworzony z jednostkowych nakładów czynników produkcji. Wzrost TFP jest sposobem oceny zmian efektywności procesu produkcyjnego, wynikających z postępu technicznego.

Zasadniczym celem referatu jest określenie czynników, które determinują kształtowanie się TFP w działach sekcji D „przetwórstwo” w Polsce. Ze względu na niemierzalność TFP, wstępnym zadaniem jest oszacowanie wartości tej zmiennej. W prezentowanym badaniu zastosowano dwie metody, bazujące na funkcji produkcji Cobba – Douglasa, różniące się sposobem oszacowania elastyczności produkcji względem nakładów pracy. Jako narzędzia analizy zastosowano

¹ Praca naukowa finansowana ze środków na naukę w latach 2007-2009, jako projekt badawczy nr N111 0938 33

modele panelowe: model z heteroskedastycznością grupową, oraz model dynamiczny.

METODOLOGIA

Modele panelowe są szacowane na podstawie danych przekrojowo – czasowych, w których liczba obserwowanych obiektów N przekracza, niekiedy znacznie, liczbę punktów w czasie T . Cechą charakterystyczną ich konstrukcji jest wyróżnienie efektu grupowego, stałego w czasie, specyficznego dla danego obiektu². Dynamiczny model panelowy ma postać:

$$y_{it} = \alpha_0 + \gamma y_{i,t-1} + \mathbf{x}_{it}^T \boldsymbol{\beta} + (\alpha_i + \varepsilon_{it}), \quad (1)$$

dla $i = 1, \dots, N$, $t = 1, \dots, T$. Opóźniona zmienna objaśniana jest skorelowana ze stałymi w czasie efektami grupowymi α_i , co powoduje, że metody stosowane do estymacji statycznych modeli panelowych nie mogą być stosowane do estymacji modelu (1), gdyż estymatory te byłyby niezgodne i obciążone. W literaturze proponuje się alternatywne metody estymacji panelowych modeli dynamicznych. Ich przegląd znaleźć można np. w [Baltagi 2008], [Hsiao 2003]. Metody te bazują na Metodzie Zmiennych Instrumentalnych, Metodzie Największej Wiarygodności lub Uogólnionej Metodzie Momentów (GMM). Zaletą GMM jest między innymi możliwość uwzględnienia alternatywnych założeń odnośnie korelacji zmiennych objaśniających (elementów wektora \mathbf{x}_{it} modelu (1)) ze składnikiem losowym ε_{it} . W szczególności dopuszczona może być endogeniczność zmiennych \mathbf{x}_{it} (wszystkich, lub części z nich), to znaczy ich skorelowanie z wartościami składnika losowego: bieżącą ε_{it} i opóźnionymi $\varepsilon_{i,t-s}$, lub słaba egzogeniczność zmiennych \mathbf{x}_{it} , to znaczy ich korelacja z wartościami $\varepsilon_{i,t-s}$. Pozostałe metody zakładają ścisłą egzogeniczność zmiennych \mathbf{x}_{it} , to znaczy brak korelacji tych zmiennych z ε_{it} , $\varepsilon_{i,t-s}$, i $\varepsilon_{i,t+s}$.

Do estymacji modelu kształtowania się łącznej produktywności czynników produkcji (TFP), prezentowanego w niniejszym referacie zastosowano dwie, alternatywne metody: GMM pierwszych różnic (FDGMM) przedstawioną w [Arellano, Bond 1991], oraz systemową GMM (SGMM) [Blundell, Bond 1998]. Poniżej przedstawiona jest zasadnicza idea tych metod.

Zastosowanie FDGMM wymaga przyjęcia założenia, że składnik losowy ε_{it} w równaniu (1) nie wykazuje autokorelacji. W celu usunięcia efektów grupowych α_i oblicza się pierwsze różnice modelu (1). W tak powstałym modelu, postaci:

$$y_{it} - y_{i,t-1} = \gamma(y_{i,t-1} - y_{i,t-2}) + (\mathbf{x}_{it}^T - \mathbf{x}_{i,t-1}^T) \boldsymbol{\beta} + (\varepsilon_{it} - \varepsilon_{i,t-1}), \quad (2)$$

² Można też wyodrębnić dodatkowo trzecią składową, stałą względem obiektów, zwaną efektem czasowym.

zmienne objaśniające zastępuje się instrumentami, którymi są opóźnione poziomy zmienne słabo egzogeniczne i endogeniczne, oraz pierwsze różnice zmiennych egzogenicznych. Estymator, uzyskany przy zastosowaniu GMM z wykorzystaniem wskazanych powyżej instrumentów zaproponowano w [Arellano, Bond 1991]. Nosi on nazwę estymatora GMM pierwszych różnic (ang. *first-differenced GMM*, FDGMM).

Estymatory FDGMM mogą być silnie obciążone w przypadku, kiedy opóźnione poziomy zmienne są słabymi instrumentami dla zmiennych zróżnicowanych. W takich przypadkach lepsze rezultaty daje stosowanie systemowego estymatora GMM (ang. *system GMM*, SGMM) [Blundell, Bond 1998]. Zasadnicza idea SGMM polega na oszacowaniu systemu równań: modelu (2) i modelu (1), a więc przyrostów i poziomów tego samego modelu. Dla równań na przyrostach postępowanie jest identyczne, jak w przypadku FDGMM. Natomiast w równaniach na poziomach, instrumentami dla z góry ustalonych i endogenicznych zmiennych objaśniających są opóźnione pierwsze różnice odpowiednich zmiennych. Instrumenty te są właściwe, przy założeniu, że ε_{it} nie wykazuje autokorelacji oraz że prawdziwe są warunki początkowe, postaci: $E(\alpha_i \Delta y_{i2}) = 0$ dla $i = 1, \dots, N$.

Oceny jakości modelu oszacowanego FDGMM lub SGMM dokonać można na podstawie testu autokorelacji Arellano - Bonda, lub testu Sargana (por. [Arellano, Bond 1991]). W części empirycznej niniejszego referatu stosowany jest jedynie pierwszy z nich, gdyż obliczane są odporne błędy szacunku parametrów, a wówczas rozkład statystyki testu Sargana nie jest znany. Hipoteza zerowa testu Arellano - Bonda głosi, że w modelu pierwszych różnic (2) nie występuje autokorelacja składnika losowego drugiego rzędu³, co oznacza, że warunki momentów są spełnione, a więc zastosowane instrumenty są właściwe.

Dodatkową możliwość sprawdzenia, czy oceny parametrów uzyskane na podstawie FDGMM lub SGMM są nieobciążone daje porównanie ich z ocenami wyznaczonymi na podstawie estymatora wewnątrzgrupowego (WG) i estymatora KMNK modelu *pooled*⁴. [Nickell 1981] wykazał, że estymator WG parametru autoregresyjnego γ jest, przy ustalonym T , obciążony w dół, a estymator KMNK jest obciążony w górę (np. [Hsiao 2003]). Wartość zgodnego estymatora parametru γ powinna zawierać się zatem pomiędzy wartościami tych dwóch estymatorów.

³ Występowanie w modelu (2) autokorelacji pierwszego rzędu jest zjawiskiem spodziewanym, gdyż jeśli ε_{it} są niezależne, to ich pierwsze różnice są skorelowane rzędu 1.

⁴ Estymator wewnątrzgrupowy jest stosowany do szacowania modeli statycznych, w których efekty grupowe α_i są nielosowe (modeli FEM). Model *pooled* to model szacowany na podstawie danych panelowych, w którym jednak nie wyróżnia się efektów grupowych ani czasowych i zakłada się, że macierz wariancji-kowariancji składnika losowego jest sferyczna.

OSZACOWANIE ŁĄCZNEJ PRODUKTYWNOŚCI CZYNNIKÓW PRODUKCJI (TFP)

Jak wspomniano we wstępie, łączna produktywność czynników produkcji (TFP) jest zmienną nieobserwowalną. Jednym ze sposobów oszacowania TFP jest wykorzystanie funkcji produkcji. Poniżej zaprezentowane zostaną wyniki, uzyskane dwiema alternatywnymi metodami. Pierwsza z nich opiera się na propozycji zawartej w [Tokarski 2008], i wymaga oszacowania parametrów funkcji wydajności. Druga metoda opiera się na założeniu o doskonałej konkurencji na rynkach czynników produkcji i założeniu, że elastyczność produkcji względem nakładu pracy jest stała w czasie, a może się różnić między działami. Została ona zaproponowana w [Ascari, Di Cosmo 2004].

Opierając się na propozycji Tokarskiego [2008] wartości TFP dla poszczególnych działów wyznaczono na podstawie zależności:

$$\ln\left(\frac{Y_{it}}{L_{it}}\right) = \left(\sum_{i=15}^{37} \beta_i d_i + gt\right) + \alpha \ln\left(\frac{K_{it}}{L_{it}}\right) + \varepsilon_{it}. \quad (3)$$

Jest to zlogarytmowana funkcja wydajności pracy Cobba-Douglasa gdzie: Y – wartość produkcji sprzedanej w mln. zł.⁵, L – nakłady pracy (w tys. pracujących), K – nakłady kapitału rzeczowego (wartość środków trwałych brutto w mln. zł.), α – elastyczność zmiennej Y względem zmiennej K , i – numer działu, t – numer okresu.

Łączna produktywność czynników produkcji (TFP) jest równa $e^{\sum \beta_i d_i + gt}$, gdzie g – stopa postępu technicznego w sensie Hicksa, d_i – zmienne zero-jedynkowe specyficzne dla działów (i – numer działu). Celem ich wprowadzenia jest dopuszczenie zróżnicowania TFP według działów, a procedura taka nosi nazwę procedury dywersyfikacji stałej⁶.

Model (3) oszacowano na podstawie danych z lat 1998 – 2007, pochodzących z Roczników Statystycznych Przemysłu GUS. Dotyczą one 22 działów sekcji D – „przetwórstwo”, oznaczonych w klasyfikacji PKD numerami 15 – 37. Do estymacji parametrów modelu (3) zastosowano Uogólnioną Metodę Najmniejszych Kwadratów (UMNK), gdyż ze względu na duże zróżnicowanie wydajności i technicznego uzbrojenia pracy między działami dopuszczono występowanie heteroskedastyczności grupowej.

Szczegółowe wyniki estymacji nie są tu prezentowane z braku miejsca; wydają się one jednak zadowalające. Oszacowana stopa postępu technicznego w sensie Hicksa wynosi ok. 4% ($\hat{g} = 0,0439$), a elastyczność wydajności pracy względem

⁵ Oszacowano również model, w którym za Y przyjęto wartość dodaną brutto w mln. zł., ale wyniki były znacznie gorsze, dlatego nie są tu prezentowane.

⁶ por. Pindyck, Rubinfeld [1991].

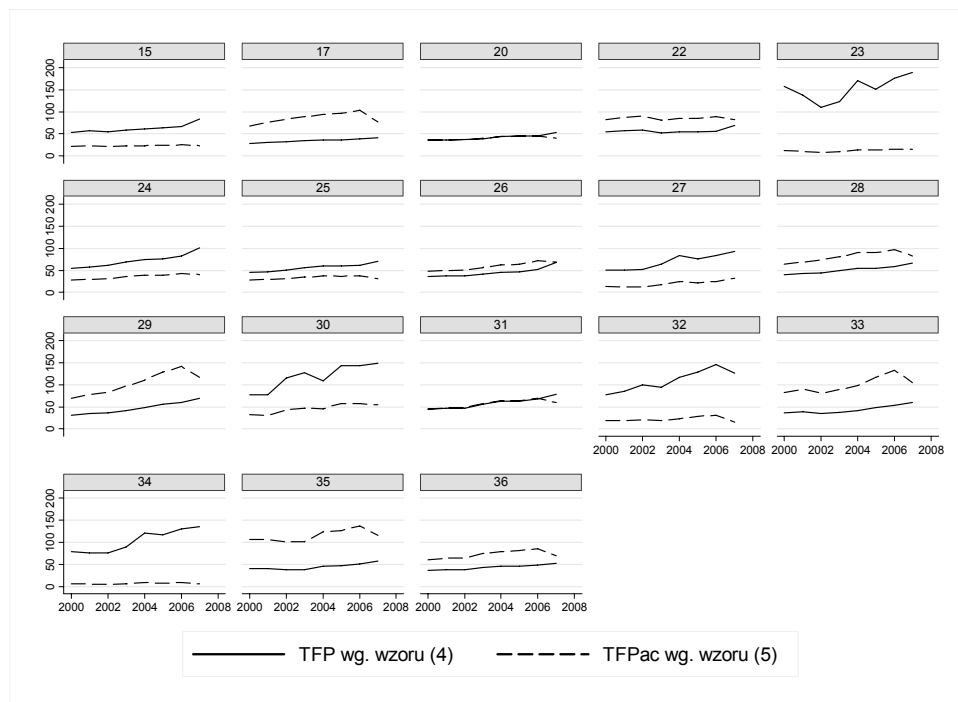
technicznego uzbrojenia pracy jest równa 0,26 ($\hat{\alpha}=0,2632$). Ponadto wszystkie zmienne są statystycznie istotne na każdym poziomie istotności, a bardzo dobre dopasowanie modelu ($R^2=0,9744$) pozwala uznać, że wartości TFP wyznaczone na podstawie modelu (3) są wiarygodne.

Na podstawie wyników estymacji modelu (3) oszacowano następnie wartości TFP w dziale i w roku t według wzoru:

$$TFP_{it} = \frac{(Y_{it} / L_{it})}{(K_{it} / L_{it})^{0,26321}}, \quad (4)$$

gdzie wartość wykładnika w mianowniku jest oceną parametru α modelu (3).

Rysunek 1. Porównanie wartości TFP (wzór (4)) i TFP_{pac} (wzór 5))



Źródło: opracowanie własne

Kolejnym etapem badań była próba ponownego oszacowania TFP przy zastosowaniu podejścia [Ascari, Di Cosmo 2004]. Polega ono na wyznaczeniu oceny parametru α_i , będącego elastycznością produkcji względem nakładu pracy na podstawie zależności $\alpha_i = w_i \cdot L_i / Y_i$, gdzie w_i jest przeciętnym wynagrodzeniem, Y_i – wartością produkcji sprzedanej, L_i – zatrudnieniem. Wszystkie wartości są specyficzne dla działów oraz stałe w czasie i reprezentują średnie dla poszczególnych

działów po czasie. Jest to zgodne z przyjętymi przez autorów założeniami, że α_i jest parametrem technologicznym, który nie powinien się zmienić w okresie objętym próbą, zaś rynki czynników produkcji charakteryzują się konkurencją doskonałą. Następnie wyznacza się łączną produktywność czynników produkcji korzystając z zależności:

$$TFPac_{it} = Y_{it} / (K_{it}^{1-\hat{\alpha}_i} L_{it}^{\hat{\alpha}_i}), \quad (5)$$

gdzie $\hat{\alpha}_i$ jest oszacowaną wartością parametru α_i . Metoda zawarta w [Ascari, Di Cosmo 2004], podobnie jak w [Tokarski 2008], wykorzystuje więc funkcję produkcji Cobba – Douglasa ze stałymi efektami skali.

Wyniki otrzymane na podstawie wzoru (5) różnią się istotnie w porównaniu z wynikami uzyskanymi na podstawie wzoru (4). Porównanie wartości TFP wyznaczonych na podstawie wzorów (4) i (5) przedstawiono na rysunku 1.

Wydaje się, że oszacowania TFP w poszczególnych działach, uzyskane na podstawie wzoru (5) nie są poprawne. Znacznie bardziej prawdopodobne jest, że średnie wartości TFP w działach rozkładają się tak, jak uzyskane na podstawie wzoru (4). Intuicyjnie, efekty postępu technicznego powinny być bardziej widoczne w dynamicznie rozwijających się działach związanych z produkcją sprzętu komputerowego (30), telekomunikacyjnego (32) czy samochodowego (34), niż w działach związanych z włókiennictwem (17), optyką (33) czy pozostałym sprzętem transportowym (35). Dlatego zmienną objaśnianą w modelu kształtowania się łącznej produktywności czynników produkcji będzie zmienna *TFP* wyznaczona ze wzoru (4).

MODEL KSZTAŁTOWANIA SIĘ ŁĄCZNEJ PRODUKTYWNOŚCI CZYNNIKÓW PRODUKCJI

Po oszacowaniu wartości TFP podjęto próbę skonstruowania i oszacowania modelu ekonometrycznego, opisującego kształtowanie się tej zmiennej według działów. W związku z niedostępnością części danych, dotyczących działalności badawczo – rozwojowej (brak danych o nakładach i zatrudnieniu dla działów 16, 18, 19, 21, 37) oraz danych o imporcie i eksporcie dla wszystkich działów w latach 1998 – 1999, zdecydowano się zawęzić próbę. Analiza kształtowania się TFP przeprowadzona została na podstawie danych dla 18 działów ($i=15, 17, 20, 22, \dots, 36$) z lat 2000–2007.

Na podstawie opracowań o charakterze teoretycznym i empirycznym, znanych z literatury światowej⁷, wytypowano czynniki, mogące wpływać na poziom TFP. Są to: kapitał ludzki, kapitał fizyczny, nakłady na działalność badawczo – rozwojową (BiR), transfer technologii, dokonujący się poprzez handel międzyna-

⁷ por. Cameron [2006], Griffith i in. [2003], Coe i in. [1995], Acharya i in. [2007]

rodowy. Oszacowano kilkanaście modeli, w których zmiennymi objaśniającymi były następujące, alternatywne miary wymienionych powyżej czynników:

- dla określenia kapitału ludzkiego: hk – stosunek liczby zatrudnionych na stanowiskach nierobotniczych do liczby zatrudnionych ogółem, $zbrlp$ – stosunek liczby zatrudnionych w działalności BiR do liczby pracujących ogółem;
- dla określenia kapitału fizycznego: $stbh$ – wartość środków trwałych brutto na jedną przepracowaną godzinę w zł/godz.;
- dla określenia wpływu transferu technologii: $imps$ – stosunek wartości importu do wartości produkcji sprzedanej, $ekps$ – stosunek wartości eksportu do wartości produkcji sprzedanej;
- dla określenia wpływu działalności badawczo – rozwojowej: $brinw$ – stosunek nakładów na BiR do nakładów inwestycyjnych.

Wymienione powyżej zmienne wprowadzane były do modelu albo jako wartości bieżące, albo opóźnione o 1 lub 2 okresy. Dodatkowo konieczne okazało się wprowadzenie zmiennych zero-jedynkowych, które dla kolejnych lat próby oznaczono symbolami $r00, r01, \dots, r07$. Ostateczna postać modelu, który charakteryzował się najlepszymi własnościami merytoryczno – statystycznymi jest następująca:

$$\begin{aligned} \ln(TFP_{it}) = & \alpha_0 + \gamma \ln(TFP_{i,t-1}) + \beta_1 \ln(hk_{it}) + \beta_2 \ln(brinw_{i,t-2}) + \\ & + \beta_3 \ln(stbh_{it}) + \beta_4 \ln(imps_{i,t-1}) + \\ & + \beta_5 r02 + \beta_6 r03 + \beta_7 r05 + \varepsilon_{it} \end{aligned} \quad (6)$$

Do estymacji modelu (6) stosowane były FDGMM i SGMM, przy czym przyjmowane były alternatywne założenia o korelacji zmiennych objaśniających ze składnikiem losowym. Ostatecznie uznano, że najlepszą metodą estymacji modelu (6) jest SGMM z założeniem endogeniczności zmiennych $stbh$ i hk . Właściwość instrumentów stwierdzono testem autokorelacji Arellano-Bonda. Ponadto ocena parametru autoregresyjnego (0,673) zawiera się pomiędzy oceną WG (0,505) a oceną KMNK (0,9342), co dodatkowo świadczy o zgodności estymatora.

Najsilniejszy wpływ na wzrost TFP wywiera wzrost kapitału ludzkiego (elastyczność 0,22%). Bardzo znaczący jest również wpływ transferu technologii na TFP, dokonujący się poprzez wymianę międzynarodową. Elastyczność zmiennej $imps$ względem TFP wynosi 0,12%. Nieco mniejsza jest siła wpływu kapitału fizycznego (elastyczność 0,05%) i nakładów na BiR (elastyczność 0,03%). Oddziaływanie transferu technologii i nakładów na BiR jest odroczone w czasie, przy czym długość tego opóźnienia wynosi jeden rok w przypadku zmiennej $imps$, a w przypadku zmiennej $brinw$ – nawet 2 lata. Wysoka wartość oceny parametru autoregresyjnego modelu (6), równa 0,673, świadczy o stabilności łącznej produktywności czynników produkcji w działach.

ZAKOŃCZENIE

Uzyskane wyniki wskazują na duże zróżnicowanie TFP między działami, przy czym nie stwierdzono tendencji do wyrównywania się tych różnic. Wykazano, że istotny wpływ na kształtowanie się TFP wywiera kapitał fizyczny, kapitał ludzki, transfer technologii i nakłady na działalność badawczo-rozwojową.

LITERATURA

- Acharya, R. C., Keller, W. (2007), Technology Transfer through Imports, *NBER-Working Paper* 13086.
- Arellano, M., Bond, S. (1991), Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations, *Review of Economic Studies* 58, 277–297.
- Ascari, G., Di Cosmo, V. (2004), Determinants of Total Factor Productivity in the Italian Regions, *University of Pavia, Department of Economics, Working Paper # 170 (12-04)*, <http://129.3.20.41/eps/mac/papers/0511/0511009.pdf> (8.03.2009r.).
- Baltagi, B. H. (2008), *Econometric Analysis of Panel Data*, Wiley&Sons, Chichester
- Blundell, R., Bond, S. (1998), Initial Conditions and Moment Restrictions in Dynamic Panel Data Models, *Journal of Econometrics* 87(1), 115–143.
- Cameron, G. (2006), Openness, R&D, and Growth at the Industry Level, str. 137–158, w: Finley, L. A. (ed) *Perspectives on Economic Growth*, Nova Publishers: Hauppauge NY.
- Coe, D., Helpman, E. (1995), International R&D Spillovers, *European Economic Review* 39, 859–887.
- Griffith, R., Redding, S., Van Reen, J. (2003), R&D and Absorptive Capacity: Theory and Empirical Evidence, *Scandinavian Journal of Economics* 105, 99–118.
- Hsiao, C. (2003), *Analysis of Panel Data*, 2nd edn., Cambridge University Press, Cambridge.
- Nickell, S. (1981), Biases in Dynamic Models with Fixed Effects, *Econometrica* 49, 1417–1426.
- Pindyck, R.S., Rubinfeld, D.L. (1991), *Econometric Models and Economic Forecasts*, McGraw-Hills, New York.
- Tokarski, T. (2008), Oszacowanie regionalnych funkcji produkcji, *Wiadomości Statystyczne* 10, 38–53.

Estimating and modeling of TFP in the Polish industry. Panel data analysis

Summary: The paper attempts to estimate total factor productivity (TFP) for sectors included in section „manufacturing”, and then to determine factors influencing it. Two alternative methods, based on the Cobb-Douglas production function were applied. The final step was construction and estimation of dynamic panel data model, describing TFP formation by sector. The explained variable was TFP, which values were estimated in the first step. GMM-based methods were used for estimation of the model.

Keywords: dynamic panel data model, production function, total factor productivity (TFP), system GMM estimator

FORECASTING THE END-OF-THE-DAY REALIZED VARIANCE^{*}

Małgorzata Doman

Katedra Matematyki Stosowanej, UE w Poznaniu
e-mail: malgorzata.doman@ue.poznan.pl

Ryszard Doman

Wydział Matematyki i Informatyki, AMU w Poznaniu
e-mail: ryd1oman@amu.edu.pl

Abstract. A large package of information is being reflected in stock prices during a short period after opening. Moreover, the start-of-the-day (morning) volatility has a strong impact on the price variability during all the day. In this connection, the question is whether the morning realized variance calculated as the sum of morning squared intraday returns can be useful in forecasting the daily realized variance (end-of-the-day volatility). In the paper, we apply three different methods of forecasting the daily realized variance for stocks quoted on the Warsaw Stock Exchange. Our findings show that the morning realized variance provides valuable information that can be used in forecasting the daily realized variance.

Keywords: realized volatility, forecasts, ARFIMA, unobserved component model

INTRODUCTION

The accurate measurement, estimation and forecasting of stock market volatility is a crucial task in portfolio management, asset pricing (especially, derivatives pricing) and risk management. The most popular definition used in quantitative finance is that the price volatility is the variance of return conditional on the information available a period earlier. Thus, volatility is an unobservable variable and its measurement depends on a model. The most common approach to model the stock market volatility is to use a model from the GARCH family [Tsay 2002]. Another possible method is to calculate the so-called implied volatility [Poon

^{*} This work was financed from the Polish science budget resources in the years 2007-2010 as the research project NN 111 1256 33.

2005]. The daily realized volatility or the realized variance, proposed by Andersen and Bollerslev [1998], is calculated as a sum of squared intraday returns from the day.

The intraday returns coming from different part of the day give different input into the daily realized variance. A very large package of information is being reflected in prices during a short period after opening the stock. Therefore, the start-of-the-day (morning) volatility has a strong impact on the return dynamics during all the day. In this connection, the question arises of whether the morning realized variance can be useful in forecasting the daily realized variance (end-of-the-day volatility).

The aim of the presented research is to assess to what extent the intraday data from the beginning of the day can explain and predict the end-of-the-day volatility. We use the realized variance as a volatility measure and show that the realized variance calculated for first few hours of the trading day (MORNING realized variance) is a useful predictor of the daily volatility. The idea of our approach is slightly similar to the one presented in [Frijns and Margaritis 2008] though the applied way of describing the realized variance dynamics is different. We apply 3 methods of forecasting the daily realized variance of the stock returns. They include a direct modeling as an ARFIMA process, as an ARFIMA process with MORNING realized variance as an additional explanatory variable, and by using an unobserved component model. The analysis is performed for four chosen stocks from the Warsaw Stock Exchange (AGORA, TVN, BZWBK and PEKAO). Our results show that the daily volatility forecasts based on the morning realized variance can be a useful tool in short term investments.

REALIZED VARIANCE AND MARKET MICROSTRUCTURE EFFECTS

In this paper we consider logarithmic percentage returns

$$R_t = 100(\ln P_t - \ln P_{t-1}), \quad (1)$$

where P_t is the closing price on trading day t .

The daily volatility σ_t^2 of a financial instrument is defined as the conditional variance of its daily return given the set of information Ω_{t-1} available on day $t-1$, i.e.

$$\sigma_t^2 = E((R_t - E(R_t | \Omega_{t-1}))^2 | \Omega_{t-1}) \quad (2)$$

Thus the volatility is an unobservable variable and (assuming that the conditional mean of a daily return equals zero) the square of a daily return is an unbiased

estimator of it. However, there exist many empirical findings showing that this estimator is very noisy [Andersen and Bollerslev 1998].

In their paper devoted to forecasting abilities of GARCH models, Andersen and Bollerslev [1998] proposed another estimator of daily volatility: the sum of squared intraday returns, and called it the realized volatility (variance). We denote the daily realized variance by $\tilde{\sigma}_{[DAY]t}^2$. Thus when we consider D+1 intraday quotations, $p_{t,d}$, $d = 0, 1, \dots, D$, we have

$$\tilde{\sigma}_{[DAY]t}^2 = \sum_{d=0}^D r_{t,d}^2, \quad (3)$$

where $r_{t,d} = 100(\ln p_{t,d} - \ln p_{t,d-1})$, $p_{t,0}$ is the opening price, and $p_{t,-1} = p_{t-1,D}$ is the closing price on the day before.

THE MODELS

The realized volatility exhibits long memory and this phenomenon suggests a possibility of describing the realized volatility dynamics by using ARFIMA models [Granger and Joyeux 1980, Hosking 1981].

An ARFIMA process x_t may be defined by

$$\alpha(L)(1-L)^d(x_t - \mu) = \beta(L)\varepsilon_t, \quad (4)$$

where ε_t is a white noise process, $\alpha(L)$ and $\beta(L)$ are the lag polynomials of order p (autoregressive) and q (moving average), respectively, and $(1-L)^d$ is a fractional differencing operator defined by the binomial expansion. If $|d| < 0.5$ and the roots of $\alpha(L)$ and $\beta(L)$ lie outside the unit circle, the process x_t is stationary and invertible. For $0 < d < 0.5$, x_t displays long memory property. ARFIMA models are the most common way of modeling and forecasting the daily realized volatility [Koopman et al. 2005, Doman 2006]. It is possible to extend the ARFIMA model by introducing additional explanatory variables into the model equation. In this case, the constant μ in (4) is replaced by

$$\mu_t = \mu + \sum_{i=1}^n c_i x_{i,t}. \quad (5)$$

The idea of modeling the realized volatility by unobserved component (UC) models comes from the results of Barndorff-Nielsen and Shephard [2002]. They studied the statistical properties of the volatility estimate error, $\sigma_t^2 - \tilde{\sigma}_t^2$, where

$\tilde{\sigma}_t^2$ is the estimator of the realized volatility based on intraday returns, and suggested modeling the spot volatility as some continuous-time stochastic process, more precisely, as a sum of independent Ornstein-Uhlenbeck processes. According to their results, in such a situation the actual volatility corresponding to day intervals can be modeled as the sum of ARMA(1,1) processes.

Following Koopman et al. [2005], we use only one ARMA(1,1) process and the specification of a UC model given by the equations

$$\tilde{\sigma}_t^2 = \sigma_t^2 + u_t, \quad u_t \sim N(0, \sigma_u^2), \quad (6)$$

$$\sigma_t^2 = \mu + \phi(\sigma_{t-1}^2 - \mu) + \theta\eta_{t-1} + \eta_t, \quad \eta_t \sim N(0, \sigma_\eta^2) \quad (7)$$

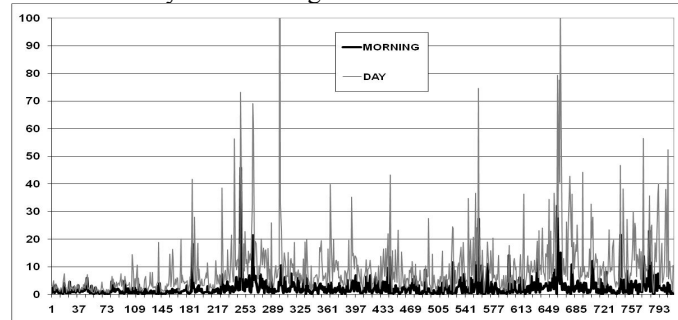
The estimated parameters are μ , ϕ , θ and σ_u^2 . The variation of the error η_t is fixed. The model equations (6) and (7) can be represented in the state-space form, and parameter estimates as well as one-day-ahead volatility forecasts can be obtained by application of the Kalman filter [Durbin and Koopman 2002].

THE DATA

Our data set includes the intraday quotations of four companies listed on the Warsaw Stock Exchange. These are two media companies: AGORA and TVN, and two banks: BZWBK and PEKAO. The period under scrutiny is from June 1, 2005 to August 22, 2008. It is divided into the in-sample period (761 days) and the out-of-sample period (50 days) for which the realized variance forecasts are calculated.

The realized variance is based on 5-minute returns. We consider the daily realized variance (DAY) and the MORNING realized variance corresponding to the period from the opening to 11:00. The MORNING realized variance is in average about 0.3 of the daily realized variance in each case. The plots in Figure 1 show the general impression of the dynamics of the considered realized variances.

Figure 1. AGORA. The daily and morning realized variances.



Source: Own calculation.

EMPIRICAL RESULTS

To answer the question about the usefulness of the morning volatility in prediction of all the day volatility, we applied 3 methods of modeling and forecasting the daily realized variance for the investigated stock returns. Besides the most common approach of modeling the daily realized variance as an ARFIMA process, we propose also to consider the ARFIMA process with the MORNING realized variance as an additional explanatory variable. The third of the considered possibilities is to use unobserved component models described in section 3.

It is a known phenomenon that the realized variance exhibits long memory [Andersen et al. 2001]. This effect is confirmed by our results. When considering ARFIMA models, the estimates of parameter d have values about 0.2 and all are significant (Table 2). The interesting observation is that, except the case of BZWBK, there is no ARMA part in the fitted models.

Table 1. Parameters estimates of fitted ARFIMA model

Parameters	AGORA	BZWBK	PEKAO	TVN
a_0	1.8420 (0.0795)	7.9670 (1.1218)	4.1132 (0.3405)	1.9007 (0.4540)
d	0.1673 (0.01485)	0.1346 (0.0574)	0.1939 (0.0295)	0.2009 (0.0162)
a_1		0.6099 (0.1812)		
b_1		0.4875 (0.1937)		

Source: Own calculation.

Table 2. Parameters estimates of fitted ARFIMA model with additional explanatory variable MORNING

Parameters	AGORA	BZWBK	PEKAO	TVN
a_0	0.6533 (0.1299)	3.3245 (0.8809)	2.2905 (0.2455)	1.8763 (0.3270)
d	0.1461 (0.0133)	0.2091 (0.0536)	0.1719 (0.0226)	0.1204 (0.0189)
a_1		0.8820 (0.1069)	-0.0051 (0.0424)	0.0797 (0.0469)
b_1		0.9081 (0.0789)		
c_1	1.1522 (0.1012)	1.2967 (0.0788)	1.1839 (0.0623)	1.3159 (0.0554)

Source: Own calculation.

Introducing into the ARFIMA model equation the additional explanatory variable MORNING, results in lowering the estimate of d (except BZWBK). The

estimates of parameter c_1 are significant implying the importance of information contained in MORNING for the explanation of the daily realized variance dynamics (Table 3).

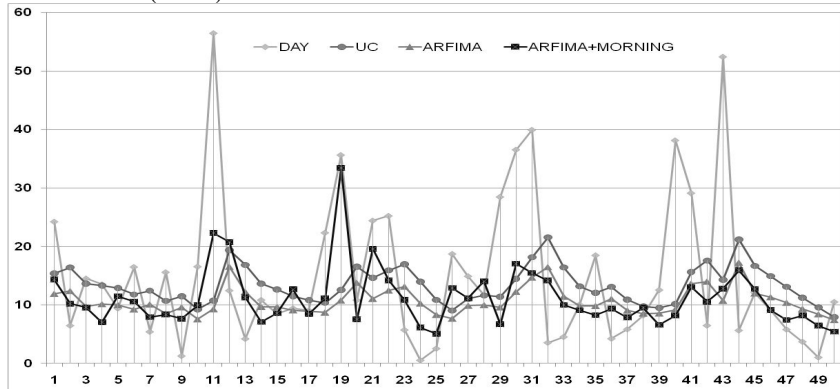
Although our earlier results [Doman and Doman 2004] show that UC models can be very useful in predicting the daily realized variance, the estimation results presented in Table 4 are not very impressive. It seems from the estimates of the variance parameters that this time the UC model turned out to not be so successful in capturing the realized variance dynamics.

Table 3. Parameters estimates of fitted UC models.

Parameters	AGORA	BZWBK	PEKAO	TVN
μ	9.4329 (1.1134)	13.0234 (0.7403)	7.9647 (0.5654)	10.7288 (1.2803)
ϕ	0.9912 (0.0523)	0.7703 (0.0582)	0.8392 (0.0487)	0.9000 (0.0305)
θ	-0.6685 (0.0903)	-0.4057 (0.0933)	-0.5013 (0.0885)	-0.6466 (0.0576)
σ_u^2	15.834 7.0362 35.634	18.676 12.676 27.517	5.6360 3.4562 9.1908	22.126 13.2819 36.8590
σ_η	11.0	8.0	5.0	10.0

Source: Own elaboration.

Figure 2. Forecasts obtained by means of the considered models and the daily realized variance (DAY). AGORA.

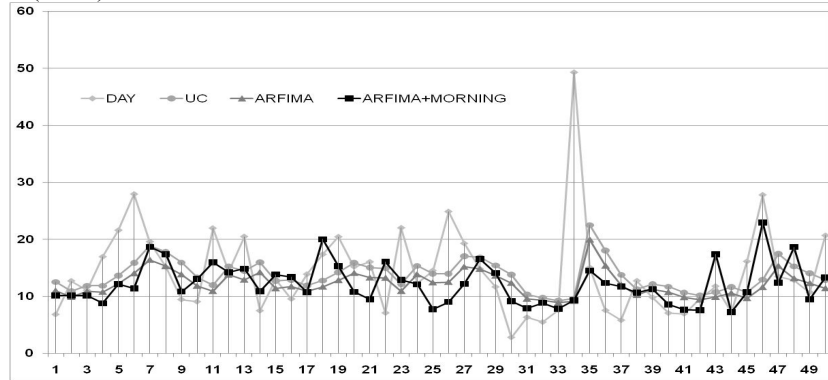


Source: Own elaboration.

The next step of the presented analysis is to compare the forecasting performance of the considered models. In each case we calculate 50 one-day-ahead forecasts. The most representative results are shown in Figures 3-4. Formal evaluation was carried out with a set of forecast quality measures. It turned out that for all

the considered stocks the best results were obtained mostly by means of AR-FIMA+MORNING models.

Figure 3. Forecasts obtained by means of the considered models and the daily realized variance (DAY). BZWBK.



Source: Own elaboration.

The final question we aim to address is how informative each method of forecasting is in relation to the other. To make these comparisons we run so-called encompassing regressions that can test the performance of one forecasting method against the other. These regressions take the form

$$\hat{\sigma}_{[DAY]_t}^2 = \alpha + \beta_1 \hat{\sigma}_{1,t|t-1}^2 + \beta_2 \hat{\sigma}_{2,t|t-1}^2 + u_t \quad (8)$$

where $\hat{\sigma}_{1,t|t-1}^2$ and $\hat{\sigma}_{2,t|t-1}^2$ are the forecasts based on two different forecasting approaches. The results of this investigation are as follows. In the case of AGORA and TVN, the forecasts from the ARFIMA+MORNING model are more informative than that from ARFIMA. For the two analyzed banks, the results are different. Though the variable MORNING was significant in the ARFIMA+MORNING equation, the forecasts from this model are not more informative. The only case where the information from ARFIMA and UC models is complementary (both betas significant) is AGORA. For the other companies, the information contain of the forecasts obtained by means of these two models is similar (both betas insignificant). Taking into account UC and ARFIMA+MORNING models, we find that the results are ambiguous. For the two media companies, the information apparently is complementary. In the case of BZWBK, the forecast informativeness of AR-FIMA+MORNING is higher but in the case of PEKAO both betas are insignificant.

CONCLUSIONS

The presented research addresses the question of to what extent intraday data FROM THE BEGINNING OF THE DAY CAN EXPLAIN AND PREDICT THE END-OF-THE-DAY VOLATILITY. Using the realized variance as volatility measure we show that the realized variance calculated for first few hours of the trading day (the MORNING realized variance) is a useful predictor of the daily volatility. To get possibility of some comparison, we apply 3 methods of forecasting the daily realized variance of stock returns. These are: a commonly used approach by a direct modeling as an ARFIMA process, by modeling as an ARFIMA process with the MORNING realized variance as an additional explanatory variable, and a more sophisticated way, by a direct modeling using an unobserved component model. The analysis is performed for four chosen stocks from the Warsaw Stock Exchange (AGORA, TVN, BZWBK and PEKAO). Our results show that the models taking into account the morning realized variance provide the best forecasts of the daily volatility and thus can be a useful tool in short term investments.

REFERENCES

- Andersen T.G., Bollerslev T. (1998), Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts, *International Economic Review* 39, 885- 905
- Andersen T.G., Bollerslev T., Diebold F.X., Ebens H. (2001) The Distribution of Realized Stock Return Volatility, *Journal of Financial Economics*, 61, 43-76
- Barndorff-Nielsen O.E., Shephard N. (2002) Econometric analysis of realized volatility and its use in estimating stochastic volatility models, *Journal of the Royal Statistical Society, Series B*, 63, 253-280
- Doman M. (2006), Modeling the Realized Volatility with ARFIMA and Unobserved Component Models: Results from the Polish Financial markets, in: Milo W., Wdowiński P. (eds), *Financial Markets. Principal of Modeling, Forecasting and Decision-Making*, Lodz University Press, Lodz, 123-137
- Doman M., Doman R. (2004) *Ekonometryczne modelowanie dynamiki polskiego rynku finansowego*, Wydawnictwo Akademii Ekonomicznej w Poznaniu, Poznań
- Durbin J., Koopman S.J. (2002), *Time Series Analysis by State Space Methods*, Oxford University Press, Oxford
- Frijns B., Margaritis D. (2008), Forecasting daily volatility with intraday data, *The European Journal of Finance* 14, 523-540
- Granger C. W. J. and R. Joyeux (1980) An Introduction to Long-Memory Time Series Models and Fractional integration, *Journal of Time Series Analysis* 1, 15-29
- Hosking J. R. M. (1981) Fractional Differencing, *Biometrika* 68, 165-176

- Koopman S.J., Jungbacker B., Hol E. (2005) Forecasting daily variability of the S&P 100 stock index using historical, realized and implied volatility measurement, *Journal of Empirical Finance* 12, 445-475
- Poon S.-H. (2005) *A Practical Guide to Forecasting Financial Market Volatility*, John Wiley, Chichester
- Tsay R.S. (2002), *Analysis of Financial Time Series*, John Wiley, New York

Prognozowanie zmienności zrealizowanej na koniec dnia

Streszczenie: W krótkim okresie po otwarciu giełdy ceny akcji kształtowane są przez duży pakiet nagromadzonych informacji, a zmienność zrealizowana z początku dnia ma silny wpływ na dzienną zmienność zrealizowaną. W związku z tym powstaje pytanie, czy poranna zmienność zrealizowana, wyliczana jako suma kwadratów śróddziennych stóp zwrotu, z kilku pierwszych godzin dnia giełdowego może być użyteczna w prognozowaniu dziennej zmienności zrealizowanej (zmienności na koniec dnia). W pracy stosowane są trzy różne metody prognozowania dziennej zmienności zrealizowanej akcji notowanych na Giełdzie Papierów Wartościowych w Warszawie. Uzyskane wyniki pokazują, że poranna zmienność zrealizowana może dostarczać informacji zwiększających skuteczność prognozowania dziennej zmienności zrealizowanej.

Słowa kluczowe: zmienność zrealizowana, prognozy, ARFIMA, modele składowej nieobserwowalnej

QUANTITATIVE METHODS IN ACCOUNTING RESEARCH

Marek Gruszczyński

Instytut Ekonometrii, SGH

e-mail: marek.gruszczyński@sgh.waw.pl

Abstract: Quantitative methods are in frequent use in modern accounting research. The evidence may be found e.g. in the journals like “Journal of Accounting Research”, “European Accounting Review”, “Review of Quantitative Finance and Accounting” or in the *Accounting Research Network* in SSRN base. Paper presents a brief survey of research areas and statistical-econometric approaches in accounting research. Particular reference goes to research on corporate disclosure. Methodological component of the paper includes remarks on the use of binary response models with choice-based and matched samples as well as comments on the sample selection approaches.

Keywords: accounting research, corporate disclosure, binary response, choice-based samples, matched samples, sample selection.

INTRODUCTION

Paper presents the brief survey as well as some methodological remarks on quantitative methods used in accounting research. Let us begin with the question of “how probable is probable in accounting?”. Documents and practice of accounting principles like IAS (International Accounting Standards) or IFRS (International Financial Reporting Standards)¹ quite distinctly call for the use of probability in delivering statements like e.g. the auditors’ opinion on some possible outcomes of company’s activities.

Table 1 lists the various uses of the language of probability in accounting. There are 27 shades of probability listed here: from “virtually certain” to “not genuine”. How do we decide which of the 27 categories should be used in the statement? It is evident that some quantitative advice on how to find the relevant

¹ IAS are set by International Accounting Standards Board while IFRS – by Financial Accounting Standards Board. In Poland IAS are known as MSR (Międzynarodowe Standardy Rachunkowości) and IFRS – as MSSF (Międzynarodowe Standardy Sprawozdawczości Finansowej).

probability level in a given situation might be in order. This is the example of how the accounting research may help the practitioners. One could imagine that, given a large number of situation of a kind, the researcher could be in the position to estimate such probability with some degree of confidence.

Table 1. Various categories of probability in accounting

Description of the probability level of business event	In accounting standards
Virtually certain	IAS 37.33
No realistic alternative	IAS 37.10
Highly probable	IFRS 5 BC82
Reasonably certain	IAS 17.4
Substantially all	IAS 17.8
Substantially enacted	IAS 12.46
Highly effective	IAS 39.88
Principally	IFRS 5.6
Significant	IAS 18.14 (a)
Major part	IAS 17.10 (c)
Probable - more likely than not	IAS 37.14 (b)
More likely	IAS 39.22
Likely	IAS 39 AG 40
May, but probably will not	IAS 37 app. A
Reasonably possible	IAS 32.92
Possible	IAS 37.10
Unlikely	IAS 39 AG 44
Highly unlikely	IAS 40.31
Extremely unlikely	IFRS 4 app. B 23
Minimal probability	IFRS 4 app. B 25
Sufficiently lower	IAS 17.10 B
Insignificant	IAS 39.9
Remote	IAS 37.28
Extremely rare	IAS 1.17
Virtually none	IAS 34 IN 6
Not genuine	IAS 32.25

Source: Aktualności MSSF, No. 44/2006, PriceWaterhouseCoopers

So far the “real” accounting research does not necessarily follow this path of reasoning. The question of how to estimate the level of probability is skipped in accounting research to more general situations like audit failure or company’s bankruptcy. As of now, answering questions posed by Table 1 remains the challenge for the future.

Section 2 presents the scope of accounting research as evident from the literature. In Section 3 tackles on research methods. As the example of research area Section 4 presents the subject of corporate disclosure. In Section 5 some methodological questions are raised, following a few recent papers on the use of qualitative variables on accounting research. Section 6 concludes.

SCOPE OF ACCOUNTING RESEARCH

Countless papers and books on accounting research make impossible to reasonably survey the entire area. In this section we begin with indicating the major journals covering modern accounting research. These are:

- Journal of Accounting Research
- European Accounting Review
- Review of Quantitative Finance and Accounting
- The Accounting Review
- Journal of Accounting and Economics
- Journal of Applied Corporate Finance
- Accounting, Auditing & Accountability Journal
- Accounting and Finance
- Accounting, Organizations and Society
- Journal of Management Accounting Research

The journals serve as the dissemination tool of new research findings and as the guide for the research directions.

As the representatives of the books on the topic let us mention just three items:

- Handbook of Management Accounting Research (2007),
- Methodological Issues in Accounting Research (2006),
- Advances of Quantitative Analysis of Finance and Accounting (in 6 volumes, 2004-2006).

Referring to the scope of accounting research we just quote the list of topics summarized in “Journal of Accounting Research”. These are:

- Cost allocation
- Capital investment decisions
- Performance measurement
- Transfer pricing
- Auditor changes
- Audit failures
- Auditor independence
- Internal control assessment
- Corporate tax planning
- Impact of changes in tax laws
- Differences across jurisdictions
- Tax vs. financial reporting
- Financial disclosure
- Financial reporting quality
- Impact of accounting standards
- Impact of securities laws
- International financial reporting

- Financial reporting of bankers
- Valuation use of accounting
- Contracting use of accounting

As we see, this is a very comprehensive list. In a sense it covers also most of the issues considered by the research on corporate finance.

Another source showing the scope of accounting research are international congresses of major scientific accounting societies. The evidence from annual congresses of European Accounting Association, as presented in Table 2 shows the topical distributions of papers delivered to those meetings.

Table 2. Papers presented to EAA Annual Congresses (2003-2008)

Subject area	Percentage of papers
Accounting and Capital Markets	4%
Accounting Education and Research	3%
Accounting History	3%
Accounting and Information Systems	2%
Accounting and Strategy	1%
Accounting Theory	3%
Auditing	6%
Critical Perspectives on Accounting	3%
Economic Analytical Accounting	2%
Financial Analysis	3%
Finance/Financial Management	1%
Financial Accounting (capital markets)	8%
Financial Reporting	11%
Corporate Governance	8%
International Accounting	8%
Management Accounting	16%
Behavioural Aspects of Accounting	5%
Public Sector and Non-profit Accounting	7%
Social and Environmental Accounting	5%
Accounting and Taxation	2%

Source: Fülbier and Sellhorn (2008)

The analysis of Table 2 reveals that accounting researchers in Europe major in Management Accounting or in the Financial Reporting areas as well as in Corporate Governance, International Accounting and Financial Accounting. Public Sector and Non-profit Accounting follow the list of most popular subjects.

To summarize – research scope in accounting is broad. It seems that the common questions in accounting research are directed to general problems, ready to tackle as a kind of average of the group of events or companies. What is more difficult remains unresolved. Typical accountant or auditor makes his/her own decision more on legal grounds than on the statistical evidence.

RESEARCH METHODS

Table 2 presents the list of research methods in accounting as compiled by Fülbier and Sellhorn (2008) on the basis of papers presented to the annual congresses of European Accounting Association in 2000 and 2005. It turns out that most popular are research methods classified as “empirical archival methods”. About 70% of papers of EAA congress in 2005 were classified into this category.

Table 3. Research methods. Papers presented to EAA Annual Congresses: 2000 and 2005

Methods	2000	2005
1) Empirical archival – database or archive Research based on historical documents, texts, journal articles, corporate annual reports, company disclosures etc. Research mainly applies statistical techniques to data drawn from commercial databases.	51%	70%
2) Empirical experiment The defining feature of an experiment is that the researcher manipulates one or more variables with subjects who are assigned randomly to various groups.	2%	0%
3) Empirical field or case study Field study and case study research is preoccupied with studying the role and function of accounting in its natural context. The term ‘case study’ usually implies research confined to a single unit of analysis. ‘Fieldwork’ encompasses more general studies of social activity.	4%	1%
4) Empirical survey Researchers usually ask a random sample of individuals, companies etc. to respond to a set of questions on a given subject.	7%	6%
5) Non-empirical – analytical Analytical research comprises theory construction and evaluation using formalized, mathematical models. This includes financial modelling, formal game theory, agency models etc. in, among others, the areas of auditing, financial reporting, and disclosure.	0%	1%
6) Non-empirical – theory Every form of scientific reasoning which is non-analytical and non-empirical. It is more “a set of tentative explanations”, which “provides acceptable answers to interesting questions”.	7%	6%
7) Other This category captures authors adopting multiple methods. It also includes research methods not easily assigned to the previous categories, as well as abstracts without any clue about the method.	21%	10%
8) Ambiguous Research method not made clear in the abstract	9%	6%
Total	100%	100%

Source: European Accounting Association, Fülbier and Sellhorn (2008)

The detailed explanation of methods classified in Table 3 is shown inside of the table. Methods relevant to the topic of this paper clearly belong to most popular group 1. These are based on statistical-econometric methodology as applied to the issues covering practically all research areas mentioned above in Section 2.

RESEARCH EXAMPLE: CORPORATE DISCLOSURE ISSUES

As an example of research founded on statistical methodology one may take the issues of corporate disclosure. This research field in accounting is strongly connected with corporate governance – currently being readily investigated in various disciplines, including corporate finance, law and economics. Below we introduce two research papers showing how quantitative approach may help in explaining corporate disclosure issues.

COUNTRY-WIDE CORPORATE TRANSPARENCY

Bushman, Piotroski and Smith (2004) in their paper on corporate transparency explore its determinants on the basis of cross-country data. Corporate transparency is understood as availability of firm-specific information to those being outside publicly traded firms. The investigation concentrates on analyzing this transparency on a country level as the system producing, collecting, validating and disseminating information to market participants outside the firm.

For each of 40+ countries examined the range of measures capturing firm-specific information environments i.e. measures of transparency has been collected. Major index taken into account is CIFAR i.e. country-specific index created by means of examining companies' annual reports on the inclusion or omission of 90 items. CIFAR measures intensity of disclosure. Along with CIFAR other variables on corporate reporting were included: DISCL measuring financial disclosures, GOVERN for governance disclosures, MEASURE for accounting principles, TIME for timeliness of disclosure, AUDIT for credibility of disclosures. The private information, acquisition and communication were represented by measures like NANALYST for average number of analysts following large firms, POOL_INV for the average ratio of the value of pooled investment schemes to GDP as well as the dummy variable IT_ENF equal to 1 if the country enforced insider trading laws before 1995. Information dissemination was shown by MEDIA variable representing penetration of media channels in the economy. MEDIA was the average rank of countries' per capita number of newspapers and televisions as reported by World Development Indicators.

This set of variables has been factor analysed for the data on 45 countries. The results reveal that of the two factors isolated the first may be regarded as representing financial transparency since it captures intensity and timeliness of financial disclosures, their interpretation and dissemination. The second factor represents governance transparency, capturing the intensity of governance disclosures used by outside investors to hold management accountable.

Consequent result of the Bushman, Piotroski and Smith (2004) research shows that governance transparency factor is related to country's legal/ judicial regime: this transparency is higher in countries with common law legal origin and high judicial efficiency. Financial transparency factor primarily is related to political regime: this transparency is higher in countries with low state ownership of enterprises, banks and low risk of state expropriation of firms' wealth.

BRAZILIAN CORPORATE DISCLOSURE INDEX

Main issue of paper by Lopes and de Alencar (2008) is the construction of customized Brazilian Corporate Disclosure Index (BCDI) and the investigation of its relevance to the cost of equity. The proposed index measures disclosure across following six dimensions:

- (i) general information about the firm its market and major events over the last year,
- (ii) relations to employees and managers regarding compensation and policies,
- (iii) non-financial information about markets, sales, and products,
- (iv) information about forecasts of sales, cash flows and earnings,
- (v) discussion and analysis of financial data including time series information about performance and explanations of past behavior,
- (vi) other information.

The final score is measured over 47 questions with binary answers being 1 for answers considered to be good disclosure and 0 otherwise. Data is composed by the panel of 50 shares with highest liquidity on São Paulo Stock Exchange (BOVESPA) on December 2005. The panel consists of firm-specific observations from the years 1998, 2000, 2002, 2004 and 2005.

Sample means of BCDI are as follows: 0.24 for 1998, 0.32 for 2000, 0.42 for 2002, 0.46 for 2004 and 0.54 for 2005. Thus, it is worth to note the increase of BCDI over time. Authors state that "firms in Latin America are immersed in a very poor institutional environment (...) Some firms in the region decide to voluntarily adopt better corporate governance practices in order to differentiate themselves and access foreign sources of capital. This movement causes governance and disclosure practices to improve on average in the region but with significant dispersion."

In their paper Lopes and de Alencar (2008) look into the relation of BCDI and the cost of equity capital (ks) measured as a function of firm's earnings and the growth in earnings. Panel data regression of ks against BCDI and control variables shows significant negative association between disclosure and cost of equity capital. In addition, this relation is more pronounced for firms which receive less attention from analysts and have dispersed ownership structures.

RESEARCH ON CD INDEX FOR POLAND

Research on corporate disclosure index for Poland is being conducted in Warsaw School of Economics (Chair of Managerial Accounting and the Institute of Econometrics). We aim at constructing the index covering country-specific disclosure issues pertinent to current accounting standards and principles in Poland. Major dimensions of the disclosure index are rooted in annual financial statement, in the company report and in the report on external relations. These are the following 9 areas of reporting:

- Financial statement: (1) Non-material assets, (2) Financial risk, (3) Fair value, (4) Accounting for derivatives, (5) Leasing, (6) Segments of activity, (7) Reserves
- (8) Company report
- (9) Report on external relations.

In each area the disclosures concerned the following issues are examined:

- Accounting policy
- Additional information and explanations – mandatory disclosure
- Additional information and explanations – voluntary disclosure.

The index has been constructed for a sample of 48 companies listed on Warsaw Stock Exchange, a panel covering companies' reports for 2005, 2006 and 2007. Major elaboration went into the detailed investigation of numerous items covered in reports. We believe that the Polish Corporate Disclosure Index (PCDI) constructed in our research correctly grasps the quality of disclosure by Polish companies.

At current stage the research is directed into quantitative analysis of associations between PCDI and the variables representing companies' performance and investors' protection.

METHODOLOGICAL ISSUES: BINARY RESPONSE MODELS AND SAMPLE SELECTION

Quantitative research issues in accounting are customarily connected with sampling problems as well as with questions pertaining to modelling qualitative variables. As in other fields, also in accounting researchers frequently do not bother to apply all the necessary methodological rigour. Some problems pertaining to methodology are described below.

BINARY RESPONSE MODELS

Since the seminal article by Maddala (1991) on limited-dependent and qualitative variables models in accounting research there have been numerous attempts to use such models in accounting. In terms of research quality the results are mixed.

Ge and Whitmore (2009) reviewed more than 30 articles published in accounting journals. The classic Maddala's (1991) paper on dangers in modelling qualitative variables in accounting research is still valid. Most of the articles researched by Ge and Whitmore (2009) have ambiguities and errors in the presentation of the logistic regression model. As they point out, incorrect presentations of the model, even in conjunction with a correct analysis, may lead to a serious misinterpretation of research findings.

CHOICE-BASED AND MATCHED SAMPLES

In their paper entitled "Three threats to validity of choice-based and matched sample studies in accounting research" Cram, Karan and Stuart (2007) discuss the dangers connected with the use of some popular non-random sampling schemes. The authors have reviewed 73 audit research papers and pointed out that most of them suffer from one or more errors mentioned in the article's title.

The choice-based samples occur when groups making different choices are sampled at different rates. In binary choice models it is typical to oversample units with $Y=1$ (e.g. bankrupt companies) as opposed to those with $Y=0$ (non-bankrupt).

The matched samples are two samples in which the members are clearly paired or are matched explicitly by the researcher. Another scheme is named semi-matched – in this case there does not need to be an equal number of each outcome collected in each stratum (e.g. industry). In accounting research we may also encounter the scheme called within-subjects design, e.g. study of firms' audit fees compared before and after some event.

Cram, Karan and Stuart (2007) indicate that since choice-based and matched samples are not random samples, it is obvious that the statistical analyses based on them shall not be the same as for the random samples. The three threats to validity of the research based on choice-based and matched samples are (1,2 and 3 below contain excerpts from the article):

1) *Use of unconditional analysis, when analysis conditional upon effects of matching variables is needed.* Researchers believe that the selection of a matched sample already controls for the matching variables. An unconditional analysis is performed, rather than the conditional one that is justified. Failure to account for industry, size, and other matching variables may have driven incorrect findings in many research studies. The authors' guidance to researchers is to either avoid use of matching, or to take the matching into account when analyzing the data. If matching is not taken into account, by either evaluating pairwise differences, or by including dummy variables for each matched set, then the research should not be accepted.

2) *Failure to control for effect of imperfectly matched variables.* Where matching is by "closest" size or other continuous measure, the matching is imperfect, and there remains the possibility that case vs. control differences in this matching variable could be the cause of differences in outcome, so researchers must evaluate that

possibility and perhaps control for it. The authors advise researchers either to avoid imperfect matching, or to perform and report sensitivity analyses on how imperfection in the matching might have influenced outcomes. A closest-matched variable such as size can still have influence. It might be controlled for by including a linear term. But, as size or another variable's contribution might be non-linear, in general, there is no fully satisfactory resolution. The researcher must make some effort to examine the possibility that all results are driven by the omitted effect. Sensitivity analyses including linear and quadratic terms, for example, might be performed and discussed. Otherwise, the researcher has not established that other reported effects are not merely the result of an omitted variable problem.

3) *Failure to re-weight observations according to differing sampling rates.* The disproportionate sampling for different population strata (implicit in the choice-based and matched sample selection) would usually necessitate weighting data by the sampling rates in each strata, but re-weighting or other appropriate adjustment to the analysis is often not implemented. Authors suggest that choice-based and matched sampling should be avoided unless explicit sampling rate information can be obtained (allowing for explicit reweighting) or unless logit regression will suffice to analyze the research questions (taking advantage of the logit exemption to the need for reweighting).

SAMPLE SELECTION APPROACHES

Final example of methods readily applied in accounting studies are selection models. The following is based on the paper by Francis and Lennox (2008). *Selectivity* occurs whenever observations self-select into discrete groups, resulting in potential coefficient bias. Estimation procedures such as OLS ignore the nonrandomness of the sample (Maddala 1991). The traditional approach to controlling for self-selection bias is the two-step procedure developed by Heckman (1979). In the first step the *selection equation* is estimated with the use of "instruments" and used to predict inverse Mills' ratios (IMRs) for each observation. In the second step the IMRs are included in the *outcome equation* (primary model of interest) as a control for the effects of selection.

Francis and Lennox (2008) have identified 30 papers that use selection models out of 545 empirical papers published from 2000 through 2007 in *The Accounting Review*, *Journal Accounting and Economics* and *Journal of Accounting Research*.

To successfully identify selectivity, the researcher should include in the first stage choice model at least one exogenous independent variable that has no direct effect on the dependent variable in the second stage regression (can be validly excluded from it). Francis and Lennox (2008) point out that these "exclusion restrictions" are rarely recognized in the accounting literature. An absence of exclusion restrictions can cause: (a) severe multicollinearity problems and (b) provide infer-

ences that are extremely fragile. The economics literature argues that researchers need to find credible and convincing exclusion restrictions.

Selection model is very sensitive to small changes in model specification whereas OLS models are robust. Moreover, the selection model is fragile even if the researcher imposes ad hoc exclusion restrictions that substantially reduce multicollinearity. In addition, the selection model is very sensitive to minor changes in sample composition with results that easily flip signs across individual sample years, whereas the OLS models are robust to such perturbations.

The empirical analysis by Francis and Lennox (2008) demonstrates how easy it is for the selection model to provide different sets of results, all of which lack robustness. Robustness is an attractive characteristic of an estimation procedure because, in practice, empirical researchers do not typically know the correct model specification (e.g., researchers may lack a theory for whether company size should be controlled for using assets, sales, or both). Not only are the main results of the selection model fragile to minor changes in specification, the signs and levels of significance for the IMR coefficients are also highly sensitive. This is important because researchers often use the significance of the IMR coefficients to judge whether or not it is necessary to control for selectivity bias. Results illustrate the dangers of relying on the IMR coefficients to infer whether it is necessary to estimate a selection model instead of OLS.

SUMMARY

Accounting research is still to be fully discovered in the countries like Poland. This paper shows the topical possibilities of the accounting research. We have indicated the research areas in accounting as well as some dangers of using statistical and econometric methodology in the studies founded on large samples.

In terms of quantitative methodology the accounting field is suitable to applying the methods of multivariate statistical analysis and microeconometrics. Paper points out several questions pertaining to correct application of qualitative and limited dependent variables models for accounting research problems.

It should be noted that accounting research is in many instances common with the research in corporate finance. Empirical corporate finance and empirical accounting research have joint roots. Both areas make use of financial reports, both produce results being pertinent to accounting as well as to corporate finance.

REFERENCES

- Advances of Quantitative Analysis of Finance and Accounting, ed. by C.-F. Lee vol. 1-6 (2004, 2005, 2006, 2006, 2007, 2008), World Scientific.
- Bushman R.M., Piotroski J.D., Smith A.J. (2004), What determines corporate transparency, *Journal of Accounting Research*, 42, no. 2, May 2004.

- Cram D.P., Karan V., Stuart I. (2007), Three threats to validity of choice-based and matched sample studies in accounting research, available at SSRN: <http://ssrn.com/abstract=955031>
- Francis and Lennox (2008), Selection models in accounting research, available at SSRN: <http://ssrn.com/abstract=1120796>
- Fülbier R.U., Sellhorn T. (2008), Approaches to accounting research – evidence from EAA Annual Congresses, available at SSRN: <http://ssrn.com/abstract=985119>
- Ge W., Whitmore G.A. (2009), Binary response and logistic regression in recent accounting research publications: a methodological note, *Review of Quantitative Finance and Accounting*, Vol. 33.
- Handbook of Management Accounting Research (2007), ed. by C.S. Chapman, Hopwood A.G., Shields M.D., Elsevier.
- Heckman (1979) The sample selection bias as a specification error, *Econometrica*, Vol. 47.
- Lopes A.B., de Alencar R.C. (2008), Disclosure and cost of equity capital in emerging markets: the Brazilian Case, available at SSRN: <http://ssrn.com/abstract=1099900>
- Maddala G.S. (1991), A perspective on the use of limited-dependent variables models in accounting research, *The Accounting Review*, vol. 66, No. 4 October 1991.
- Methodological Issues in Accounting Research, ed. by Z. Hoque (2006), Spiramus Press.

Metody ilościowe w badaniach w zakresie rachunkowości

Streszczenie: Współczesne badania naukowe w rachunkowości często korzystają z podejścia ilościowego, co można potwierdzić w wiodących czasopiśmie, takich jak “Journal of Accounting Research”, “European Accounting Review”, “Review of Quantitative Finance and Accounting” lub w bazie *Accounting Research Network* w serwisie SSRN. Artykuł przedstawia krótki przegląd dziedzin zastosowań metod ilościowych w badaniach dotyczących rachunkowości. W szczególności omówione są przykłady dotyczące tematyki ujawnień. Uwagi natury metodycznej dotyczą wykorzystania modeli zmiennej dwumianowej przy użyciu prób dobieranych a także modeli doboru próby (*sample selection*).

Słowa kluczowe: rachunkowość, ujawnienia korporacyjne, zmienne dwumianowe, próby dobierane, modele doboru próby.

AGMEMOD MODEL – STRUCTURE AND APPLICATION FOR ANALYSIS AND SIMULATION OF POLISH AGRICULTURAL SECTOR

Mariusz Hamulczuk

Katedra Ekonomiki Rolnictwa i Międzynarodowych Stosunków Gospodarczych
SGGW, e-mail: mariusz_hamulczuk@sggw.pl

Katarzyna Hertel

NBP, email: katarzyna.hertel@nbp.pl

Abstract: The aim of this paper was to present the AGMEMOD¹ model as an example of partial equilibrium model applied in the field of agricultural economics and its capacity to evaluate changes in the Common Agricultural Policy (CAP). The AGMEMOD model is an econometric, dynamic, multi-product, partial-equilibrium modelling system constructed within the 5th and 6th EU Framework Project. As an example of application of the model a quantitative assessment of the impact of the milk quotas abolition on the Polish dairy sector was displayed.

Keywords: partial equilibrium modeling, simulation, agricultural policy analysis, milk quota

INTRODUCTION

The agricultural sector is a part of economy which is strongly regulated, so in the last decades the use of economic models in relation to agricultural policy issues has increased substantially. The results of liberalisation of agricultural market and of planned agricultural policy reforms (for example ‘Health Check’ of the

¹ AGMEMOD is an acronym of the name of the project: Agriculture Member States Modeling. This research was supported by EU FP6 research funding, contract SSPE-CT-2005-021543, by contributions from the partners’ institutes throughout the EU and through associated projects for the Institute for Prospective and Technological Studies (IPTS). The authors would like to acknowledge the work of the AGMEMOD Partners in the development of the model used for this study (<http://www.tnet.teagasc.ie/agmemod/>).

CAP) need to be assessed. For a better understanding of the behavioural relationships underlying agricultural commodity markets and for forecasting in the agricultural sector a wide variety of methods are applied. Partial equilibrium (PE) modelling is one of the methods employed for the purpose of agricultural policy analysis. It allows to capture the most significant relations inside the sector, including the effects of applying different specific policy instruments.

The objective of the paper was to present the application of partial equilibrium models in the field of agri-food sector. As an example the AGMEMOD model was chosen. One of the most important reforms carried on currently by the European Commission in this sector is the dairy market reform. Therefore a simulation of milk quota abolition for the Polish dairy sector was performed and displayed as an example of the AGMEMOD model application.

The article is organised as follows: in the first part the general assumptions and structure of the AGMEMOD model are presented. In the second section the construction of dairy sub-model is considered with more details. Finally, in the last section the results of the simulation of dairy quota abolition are presented.

AGMEMOD MODEL STRUCTURE

The AGMEMOD model is an example of partial equilibrium models (PE) applied in the field of agriculture. PE models' main characteristic is that they concentrate on the isolated market, considering the rest of the economy as an unchanging background for the analysis of supply and demand in this particular market. When compared to general equilibrium models (GE), PE models are more limited in the sense that they cannot include the effects of dispersion of changes in the analysed market to other sectors of the economy. However they can be quite useful for a detailed examination of many particular factors affecting the behaviour of agents in this market, especially when the influence of these factors does not seriously affect sectors other than the one in question. Thanks to their detailed structure, they often include diversified policy variables and serve to study the effects of their changes. More about equilibrium models and their applications can be found in the paper by Tongeren and van Meijl [1999] as well as Tongeren et. al [2001]. The main features of PE models make them particularly suitable for modelling the agricultural markets, as this sector is subject to the influence of a large set of complicated policy instruments [Conforti 2001; Garforth & Rehman 2006]. Moreover, for many agricultural commodities the data in the form of complete and detailed trade balance sheets is accessible, which helps in the analysis of the equilibrium between supply and demand.

The AGMEMOD is an econometric, dynamic, multi-product, partial-equilibrium modelling system which was built in the aim to undertake a model-based economic analysis of the potential impact of policy or other changes in the agri-food sector of each EU Member State and the EU as a whole. It was constructed within the 5th and 6th EU Framework Project assuming as a main principle

that the participating institutions in each Member State would develop a part of the model describing the local agricultural sector. The country models are based on a structure common for all the EU countries and with common procedures for data collection, estimation and validation, but take into account local conditions and rely on the local experts' knowledge. Consequently, the AGMEMOD model works as a system of aggregated local models and is able to produce forecasts and scenario analyse of various policy and external conditions' changes for the Member States separately as well as for the entire EU [Donnellan et. all 2002; Chantreuil Hanrahan 2007]. The work in the project was inaugurated in 2001 by the institutions representing the Old Member States (OMS) and resulted in a coherent system of models able to produce the aggregated forecasts for the EU-15 [Hanrahan 2001; Chantreuil, Hanrahan 2007]. In the subsequent years, the project was extended to consecutive countries² willing to enter the EU (NMS), Poland among them³. The enlarged AGMEMOD model was then used to estimate the consequences of their accession to the local agricultural sectors⁴.

Each particular country model consists of a set of sub-models of the main agricultural products: grains, oilseeds and the derived products, industrial plants, milk and dairy products, livestock and meat as well as some other, of lesser importance and more locally grown products. The variables entering in each sub-model represent consecutive positions in the balance sheet of each market. On the supply side the beginning stocks, production and imports are being considered and on the demand side the domestic use, exports and ending stock are modelled. For each product in each country also the respective domestic prices (market-clearing prices) are modelled. The equilibrium in each market is reached in the model also on the level of the whole EU. This implies that the EU net export variable is used as the closing variable at the EU level. The necessary condition for the model to be solved is that the equality between supply and demand in each market in each country must hold. The general structure of the model is presented in Figure 1.

Such a disaggregation of the agricultural sector to the specific markets as well as the decentralisation of the model allowed to introduce into the model the instruments of the CAP in a very detailed manner. The instruments range from market-specific (quotas, subsidies, production levies, coupled payments) to more general ones, such as decoupled payments. Their influence can be traced in each market separately and the specificity of the agricultural policy in each member

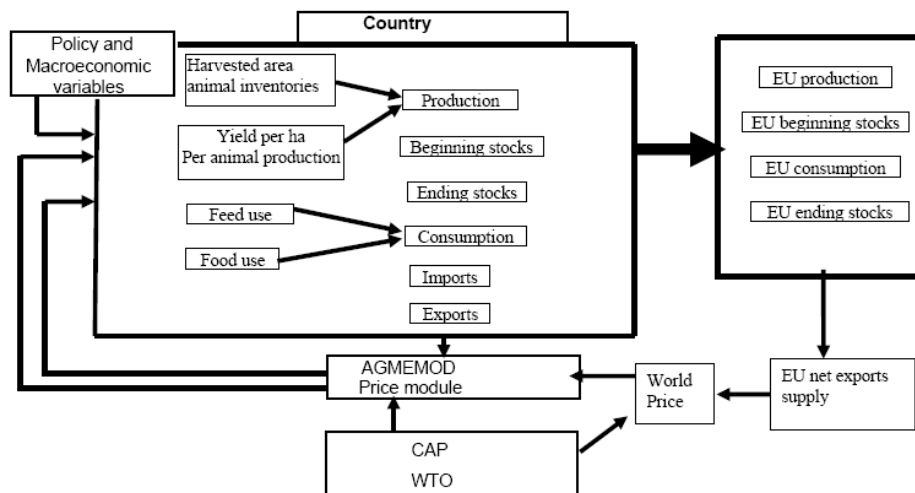
² At the moment, all the EU countries except Malta and Cyprus are represented in the model and the work is carried on to include also some Balkan countries, Russia, Ukraine and Turkey.

³ The work on the Polish part of the AGMEMOD model entertained in the Warsaw School of Economics in 2002-2006 and since 2006 was carried on in the Warsaw University of Life Sciences.

⁴ The influence of accession of Poland to the EU on the main supply and demand variables in the dairy sector was presented by Syczewska [2004].

state is taken into account. This huge advantage of partial equilibrium models such as AGMEMOD makes possible the analysis and simulations of a large spectrum of policy changes [AGMEMOD ... 2005].

Figure 1. Structure of the AGMEMOD model



Source: Chantreuil F., Tabeau A., Van Leeuwen M. (2008)

In practice, the model contains endogenous and exogenous variables. The endogenous variables are mostly prices and the variables determining the supply and demand in the market of each product in every country. The exogenous variables include a set of variables describing the general macroeconomic conditions for the functioning of the agricultural market (GDP, inflation, exchange rates, population), world agricultural prices as well as CAP instruments. A change in exogenous variables may determine the assumptions of scenarios simulated by the model.

In the AGMEMOD model the equilibrium is reached in each market of each country independently. The characteristic feature of the model is that the price does not serve as a variable which would lead to the equality between the supply and demand in the separate market at a given moment of time, but is exogenous for the supply and demand variables at a given moment of time. Therefore one of the positions of the trade balance sheet, in most cases imports or exports, is treated as a *closing variable*.

On the country level, however, there exist interdependencies between markets of particular products. They result from the assumptions of substitution or complementarities between the products. The arable land allocation between different kinds of crops according to the relative profitability of their production can be given as an example of substitution. At the same time the use of different grains,

oilseeds and the derived products as fodder in the livestock production determines the relation of complementarity between crops and livestock.

As it was mentioned before, country market models are solved independently, nevertheless the behaviour of supply and demand variables in each member state market model is driven by a common factor, so called *key price* of this particular product. For each market included in the AGMEMOD model, the *key price* is the price of the product in the country which is its most important producer in the EU. In most cases, the country price of a given commodity depends on a simultaneous development of key price, lagged domestic and EU (or key country) self-sufficiency rates and other variables.

It is assumed that the EU is not a closed economy. The determinants of key prices include the respective world prices (which are exogenous in the model), the EU self sufficiency rate, the EU intervention prices and other variables important for the behaviour of key market prices (e.g. exchange rates, tariff rates, quota limits and subsidized export limits). To increase the economic integration between OMS and NMS, the process of price level convergence in the NMS to the EU (key price) levels is assumed [Chantreuil, Tabeau, van Leeuwen 2008; Esposti & Camaioni 2007].

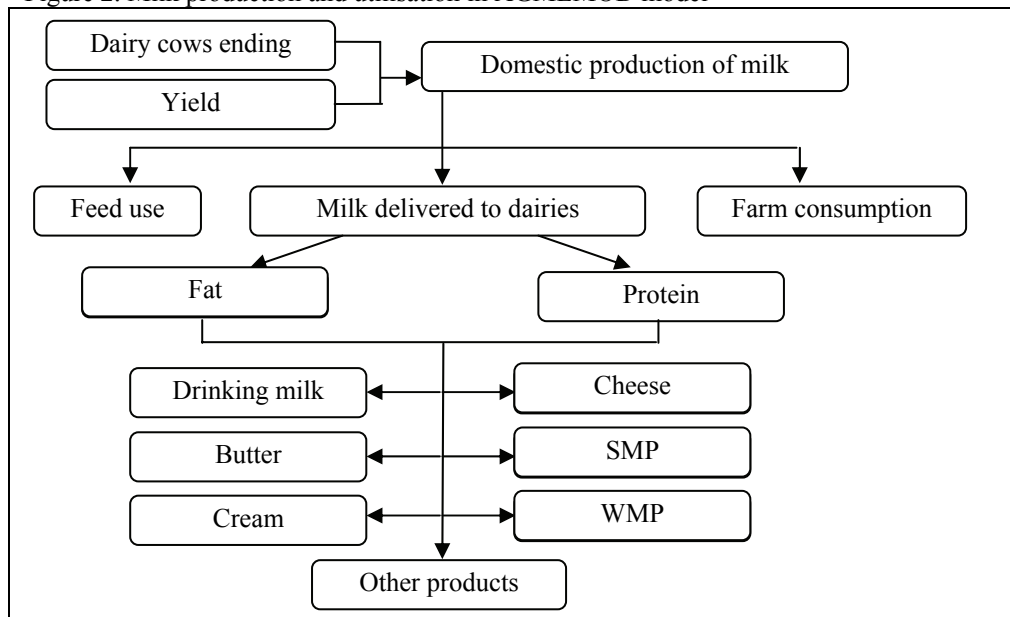
The behavioral equations of the model are mostly individually estimated with econometric techniques (generalized least squares). However, in situations where too short time series were accessible, the quality of data was unsatisfactory, important structural breaks (e.g. policy regime change) were observed or the estimation resulted in the values of parameters inconsistent with the economic theory (such as positive price elasticity of demand or negative reaction of domestic price to the key price), calibration techniques were used [AGMEMOD ... 2005]. This last technique was applied especially in case of NMS allowing for reliable parameters estimation and consequently for a long term forecasting even though relatively short time series. After the values of parameters were established, the model generated the baseline results. The validation played an important role in the construction of the model. Apart from general econometric tests on parameters and residuals, baseline results were analysed by national experts from the point of view of their feasibility. To check the reliability of the dynamics of the model a shock analysis were carried out.

DAIRY MARKET SUB-MODEL

Among the AGMEMOD market models, the dairy model is one of the most complicated. It comprises several components (Figure 2). The first component determines the production, import and export of raw milk (the last two are insignificant in Polish case), the second component allocates milk to different utilisations and the third component models milk use for further processing into dairy products [AGMEMOD ... 2005].

A particular feature of the dairy model is its emphasis on the allocation of milk fat and milk protein to the production of the various dairy commodities modelled (third stage). These products include drinking milk, butter, cream, cheese, skimmed milk powder (SMP), whole milk powder (WMP) and “other dairy products”. For each dairy commodity modelled, the supply and utilisation is projected as well as the wholesale price.

Figure 2. Milk production and utilisation in AGMEMOD model



Source: own elaboration

The key variable of dairy sub-model is milk production. Generally it is modelled as a function of quota level and ratio between milk price and milk production cost. To model the milk production under the quota regime and non quota regime different equations are used because relations between price and level of milk output are dissimilar in both periods. Moreover, the long experience of many EU countries with milk quota made it impossible to assess econometrically the reaction of milk production to milk price changes (because of the distortionary role of milk production limiting). Thus some parameters concerning country-specific quota rents (marginal costs of production) were computed outside the model. Detailed description of methodology applied for estimation of milk production (AGMEMOD version 2.0) can be found in the paper by Chantreuil et. al [2008].

The milk yield per cow can be expressed as a function of technological trends, real milk prices, quota level and other exogenous variables that could have an impact on the milk yield per cow. The number of dairy cows at the end of a period can be derived as an identity involving the milk production and milk pro-

duction per cow. The milk price received by farmers is modelled as a function of prices of dairy products like cheese, butter, SMP, WMP.

As noted before, the AGMEMOD model allocates the fat and protein components of raw milk. The amount of fat and protein in the raw milk produced that is used in the manufacturing sector is first calculated. This calculation involves a number of assumptions concerning the fat and protein content of the raw milk and the fat and protein content of the dairy commodities produced out of milk. Once the available supplies of milk protein and fat have been calculated, the next step is to allocate the protein and fat components [AGMEMOD ... 2005].

Generally fat and protein allocation to different dairy products depends on the historical allocation, prices of a given commodity and prices of substituting commodities and other exogenous variables that affect their allocation. The total protein and fat available is allocated to n dairy commodities. The milk protein and fat allocation equations are estimated for $n - 1$ products, with the milk protein and fats allocation to the n^{th} product derived as a balancing residual allocation. It is worth noticing that for a given commodity only one behavioral equation is applied (to estimate fat or to estimate protein). For example, fat for butter is behaviorally estimated while proteins are computed as an identity.

The production of a given dairy commodity is derived as the total milk protein or fats allocation divided by a technical protein and fat content in the product. The domestic consumption of a specified commodity is derived from a multiplication of the population number by consumption per capita. The consumption per capita of a given commodity is a function of GDP/capita, price of a given commodity, prices of substituting products and other exogenous variables affecting domestic consumption.

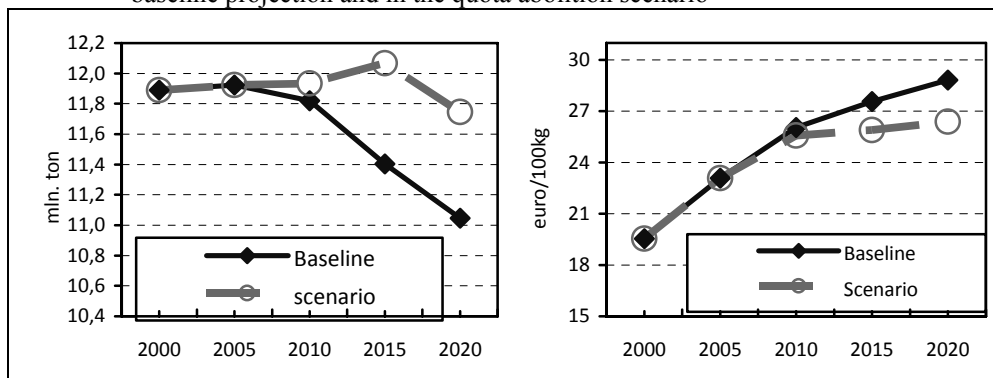
EXAMPLE OF AGMEMOD MODEL APPLICATION: MILK QUOTA ABOLITION IMPACT ON DAIRY SECTOR

Milk quotas were introduced in the European Union in 1984 in order to try to stop the over-production. In the so-called CAP "Health Check" the European Commission has recently proposed gradual transitional procedures to allow for a "soft landing" of the milk sector to quota expiry in 2014/15 [CAP... 2008, Réquilart et. al 2008, Seremak-Bulge 2008]. The objective of this study is to perform an ex ante quantitative assessment of the impact of abolition of the EU milk quotas on the Polish dairy sector. The time period covered by the analysis is up to 2020. The impact of policy scenario was assessed and compared to a baseline scenario (that is considered to be continuation of the current quota system). The scenario assumes a gradual increase of milk quota by 1% a year between 2009 and 2014 and the removal of quotas since 2015.

It is commonly expected that the abolition of milk quota system would result in an increase of milk production and a decrease of prices of raw milk and prices of

dairy products [Seremak-Bulge 2008, Jongeneel, Tonini 2008]. The simulations carried with the use of the AGMEMOD model confirm those expectations. The rise of total production of raw milk as a result of milk quota abolition is evaluated at 1.0% in 2010, 5.8% in 2015 and 6.3% in 2020 (Figure 3).

Figure 3. Milk production (left) and milk prices (right) in Poland and their forecasts in the baseline projection and in the quota abolition scenario



Source: Own calculations on the base of Agmemod model

Going into details, diverse changes of milk use are noticeable as a consequence of the potential reform. The abolition of milk quota results in a greater production capability for the dairy industry. By 2020 there can be observed an 8% increase of milk delivered to dairies which is a base for industrial production of dairy products. The level of milk consumption on farms could become lower by 5% and the feed use can boost by nearly 6% at the same time.

The baseline scenario indicates further increase of milk prices paid to farmers. Firstly, this growth would be explained by the convergence of Polish milk price to the prices observed in the OMS. The second reason would be a decrease of self-sufficiency rates in individual markets of dairy products in the situation of a relatively fixed milk supply. In the quota abolition scenario the price of raw milk in 2020 is lower by 8.4% than the price in the baseline scenario (Figure 3). As a consequence of lower prices paid to farmers the production profitability would decrease. Going further, a faster concentration of milk production should be expected as a result of the deregulation of the dairy sector.

The changes in production, consumption and prices of in particular dairy markets are presented in the Table 1. The highest decrease of the wholesale price of dairy products is observed in the case of GOUDA cheese (10%) and the least decrease can be noticed in the case of whole milk powder (WMP) (7.4%). The changes in prices of the remaining products are situated between those two values.

The growth of 6.3% (in comparison to the baseline scenario) in the quantity of milk delivered to dairies results in a similar increase of the quantity of fat and

protein, which would also stimulate growth of production of dairy commodities (Table 1). The highest increase of production is expected in case of cheese (8.7%), butter (7.2%), and SMP (7.1%). The most modest growth of production is to be observed in case of drinking milk and WMP, respectively: 2.0% and 3.4%.

Table 1. Influence of milk quota abolition on dairy market in 2020 (changes with respect to the baseline projection)

Product	Changes in:		
	Production	Consumption	Prices
Raw milk	6.3%	-	-8.4%
Drinking milk	2.0%	1.3%	-
Butter	7.2%	2.7%	-8.4%
Cheese	8.7%	2.1%	-10.0%
SMP	7.1%	3.5%	-7.8%
WMP	3.4%	3.6%	-7.4%

Source: Own calculations on the base of Agmemod model

The expansion of supply and the reduction of wholesale prices of individual dairy products would have a positive effect on the domestic consumption of those commodities. The forecasted growth of consumption caused by the abolition of milk quota system varies between 1.3% in the case of drinking milk and 3.5-3.6% in the case of both sorts of milk powder. The adjustment of production and domestic use after a dairy market reform results also in the changes of the self-sufficiency rates of products considered. Generally, the milk quota abolition would cause a rise of those indicators.

CONCLUSIONS

1. Partial equilibrium models represent a suitable tool for analyzing the changes in agricultural sector due to the fact that agriculture is influenced by large set of policy instruments. PE models allow for a detailed analysis of their impact on particular elements of this market.
2. The AGMEMOD model allows for simulation and forecasting of a variety of policy scenarios for Poland as well as for the whole EU. Further development of the model should tend to endogeneisation of the world prices in the Rest of World sub-model.
3. The application of AGMEMOD model version 2.0 for an analysis of impact of dairy quota reform indicate an increase in production and consumption of dairy products and a decrease of their prices in Poland. The direction of changes are consistent with the a priori expectations.

REFERENCE LIST

- AGMEMOD Partnership (2005) Impact analysis of the CAP reform on main agricultural commodities. Final Report. European Commission Directorate General JRC – IPTS.
- CAP HEALTH CHECK – IMPACT ASSESSMENT NOTE N° 6 (2008) European Commission, Directorate-General for Agriculture and Rural Development, Brussels.
- Conforti P. (2001) The Common Agricultural Policy in main Partial Equilibrium models. Osservatorio sulle Politiche Agricole dell'UE, INEA, Working Paper No. 8.
- Chantreuil F., Hanrahan K. (2007) AGMEMOD EU Agricultural Market Outlook. Društvo Agrarnih Ekonomistov Slovenije, DAES.
- Chantreuil F., Donnellan T., van Leeuwen M., Salamon P., Tabeau A., Bartova L. (2008) EU dairy quota reform – AGMEMOD scenario analysis. XIIth Congress of the European Association of Agricultural Economists, Ghent, Belgium.
- Chantreuil F., Tabeau A., Van Leeuwen M. (2008) Estimation of impact of EU agricultural policies on the world market prices. Modeling of Agricultural and Rural Development Policies. Sevilla, 107th EAAE Seminar "Modelling of Agricultural and Rural Development Policies". Sevilla, Spain.
- Donnellan T., Hanrahan K., McQuinn K., Riordan B. (2002) Policy analysis with AG-MEMOD Model: Dealing with diversity in the EU Agri-Food Sector. RERC, Paper prepared for presentation at the EAAE Congress 2002.
- Esposti R., Camaioni B. (2007). Technical Report on the Modeling Structure, AGMEMOD WP2 D2, Project no. SSPE-CT-2005-021543.
- Garforth C., Rehman T. (2006) Review of Models for Agricultural Policy Analysis Research Project EPES 0405/17. Project Report No. 5 The University of Reading.
- Hanrahan K. (2001) The GOLD Model 2.1: An introductory manual. Mimeo, Rural Economy Research Centre, Dublin.
- Jongeneel R., Tonini A. (2008) The 'Milk Quotas Rent Puzzle' In The EU: Economic Significance, Review, And Policy Relevance. Paper prepared for the 109th EAAE Seminar, The CAP after the Fischler. reform: national implementations, impact assessment and the agenda for future reforms, Viterbo, Italy, 2008.
- Réquillart V., Bouamra-Mechemache Z., Jongeneel R., and Penel C. (2008) Economic analysis of the effects of the expiry of the EU milk quota system. Report prepared for the European Commission, IDEI.
- Seremak-Bulge (2008) Wpływ kwotowania na funkcjonowanie rynku mleka. In: Rozwój sektora rolno-spożywczego w Polsce na tle tendencji światowych. IE-RIGŻ-PIB, Warszawa.
- Syczewska E (2004) Model AGMEMOD-CEEC+PL: Struktura i projekcje. Metody ilościowe w badaniach ekonomicznych V, SGGW, Warszawa.

Tongeren F.W. and van Meijl H. (eds.) (1999) Review of applied models of international trade in agricultural and related resource and environmental modeling. Agricultural Economics Research Institute, Report 5.99. (EU-Fair VI-CT 98-4148, interim Report 1).

Tongeren F.W., van Meijl H., van Surry Y. (2001) Global models applied to agricultural and trade policies: A review and assessment. *Agricultural Economics*, 26:149–172.

Model AGMEMOD – struktura i zastosowanie w analizie i symulacji polskiego sektora rolnego

Streszczenie: Celem opracowania było przedstawienie modelu równowagi cząstkowej AGMEMOD w kontekście modelowania sektora rolnego i możliwości ewaluacji zmian Wspólnej Polityki Rolnej (WPR). Model AGMEMOD zbudowany w ramach 5 i 6 Projektu Ramowego UE należy do ekonometrycznych, dynamicznych, wieloproduktowych modeli równowagi cząstkowej. Jako przykład zastosowania modelu i jego możliwości przedstawiono analizę wpływu likwidacji kwot mlecznych na zachowanie polskiego sektora mleczarskiego.

Słowa kluczowe: równowaga cząstkowa, modelowanie, symulacja, polityka rolna, kwoty mleczne

METODA AHP W OCENIE STOSOWALNOŚCI METOD OCENY ZDOLNOŚCI KREDYTOWEJ W SYSTEMIE EKSPERTOWYM

Władysław Hoffmann

Instytut Informatyki w Zarządzaniu US
e-mail: woodieh@gmail.com

Streszczenie: Podstawową działalnością większości banków jest udzielanie kredytów. Wiąże się to z ryzykiem utraty części lub całości długu. Banki mają obowiązek dokonywania oceny zdolności kredytowej, jednak metoda jest dowolna. Dobrym rozwiązaniem wydaje się zastosowanie systemu ekspertowego wspomagającego pracę inspektora kredytowego. Największym problemem w zbudowaniu takiego systemu jest określenie poprawnej i przejrzystej bazy reguł. W pracy ocenia się stosowalność poszczególnych metod oceny zdolności kredytowej do zastosowania w SE. Do zbudowania rankingu posłużyła metoda AHP.

Słowa kluczowe: AHP, zdolność kredytowa

WSTĘP

Udzielając kredytu, bank za każdym razem ponosi pewne ryzyko. Banki starają się zminimalizować niebezpieczeństwo, że kredytobiorca nie będzie w stanie z jakichkolwiek powodów spłacić zaciągniętego zobowiązania ani oddać pożyczonego produktu. Jednym z czynników krytycznych oceny tego ryzyka jest rzetelne zbadanie zdolności kredytowej rozumianej jako zdolność do zgromadzenia środków pieniężnych w kwocie pokrywającej wszelkie zobowiązania. Dokonanie oceny zdolności kredytowej jest obowiązkiem każdego banku zapisanym w prawie bankowym. Ważny jest fakt, że ustawodawca nie narzuca metody oceny zdolności kredytowej. Banki mają więc swobodę w opracowaniu i stosowaniu metod, którymi będą się posługiwać. Zawsze jednak, niezależnie od stosowanych procedur i metod, banki udzielają kredytu wtedy, kiedy oceniają, że będą w stanie odzyskać zaangażowane środki. Warto dodać, że ze względu na istniejącą różnorodność metod i procedur w poszczególnych bankach konkretny wniosek kredytowy może być rozpatrzony pozytywnie lub negatywnie.

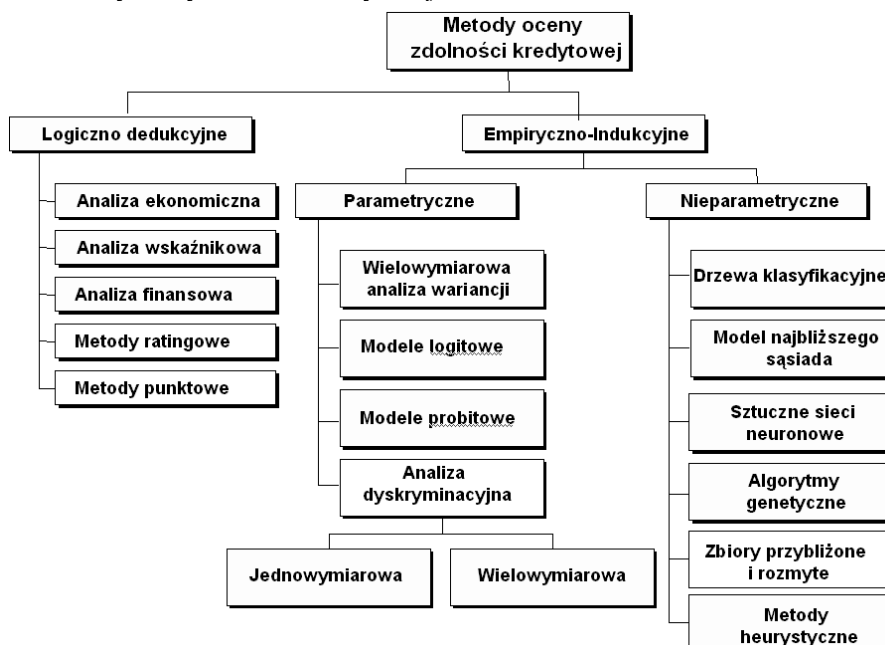
Problemy z dokładnym zbadaniem zdolności kredytowej są zazwyczaj spowodowane różnorodnością rynku, złożonością lub specyfiką prowadzonej przez kredytobiorcę działalności, wpływu czynników zewnętrznych, niezależnych. Nie bez znaczenia pozostaje problem braku porównywalności danych, niedostatku informacji oraz zafalszowania rzeczywistości przez kredytobiorcę. Jeszcze trudniej jest oszacować dane ryzyko w przypadku kredytowania rolnictwa, gdzie gospodarstwa zazwyczaj nie prowadzą księgowości.

Ilość tzw. kredytów straconych w polskich bankach waha się od kilku do nawet kilkunastu procent wolumenu udzielonych kredytów ogółem. Zależy to od banku, rodzaju kredytu itp. Wszelkie opóźnienia spłat i utracenie kredytów powodują poważne straty dla banków. Badanie zdolności kredytowej poprzez poszukiwanie reguł jakimi kierują się kredytobiorcy mogłoby ograniczyć ilość złych kredytów. Takie reguły mogłyby zasilić system ekspertowy wspomagający prace inspektora kredytowego. Problemem jest jednak dobór metody, która generowałaby sensowne i przejrzyste reguły mogące zasilać system ekspertowy.

METODY OCENY ZDOLNOŚCI KREDYTOWEJ

W trakcie wieloletniej praktyki, banki wypracowały wiele różnych metod oceny zdolności kredytowej swoich klientów. To, jakie dane bank będzie chciał pozyskać od kredytobiorcy zależy przede wszystkim od typu, wielkości i okresu spłaty kredytu a także od charakteru klienta (osoba fizyczna lub prawna). Zastosowanie tych metod ma na celu określenie czy sytuacja ekonomiczna klienta pozwoli na terminowe wywiązanie się ze zobowiązań wobec banku, czy poziom ryzyka nie jest zbyt wysoki. Niestety większość z nich nie polega na budowaniu reguł a jedynie określa wartości wskaźników, współczynników i zmiennych, przy których klient może uzyskać kredyt. Klasyfikacja metod oceny zdolności kredytowej w najogólniejszej postaci przedstawia się tak, jak przedstawiono na rys. 1. Metody opisowe (logiczno-dedukcyjne), opierają się na informacjach o kredytobiorcy, np. o jego sytuacji ekonomiczno- finansowej. W tym podejściu ocenia się szereg wskaźników uzyskanych na podstawie analizy ekonomicznej, w tym bilansu, cash flow, stanu zadłużenia, zabezpieczeń, itp. Również tutaj bierze się pod uwagę ocenę personalną kredytobiorcy. [Migut 2003] Opierają się na założeniu, że poziom ryzyka zależy od obszarów oceny związanych z konkretnym kredytobiorcą i przedmiotem kredytowania.[Staniec 2002]. W metodach statystyczno- matematycznych zdolność kredytową ocenia się na podstawie cech i zachowań wcześniejszych kredytobiorców. Do tej grupy od pewnego czasu należą też pewne metody nieklasyczne, które oprócz danych ilościowych potrafią wnioskować z informacji miękkich. Wnioski na temat ryzyka kredytowego wyciągane są na podstawie statystycznych danych innych kredytobiorców. Metody te są również szeroko stosowane do prognozowania bankructw.

Rysunek 1 Metody oceny zdolności kredytowej



Źródło: Migut G. „Modelowanie ryzyka kredytowego”, materiały StatSoft Polska www.statsoft.pl

AHP W OCENIE METOD DO ZASTOWOWANIA W S.E.

W pracy rozważane mogą być jedynie metody, które mogą znaleźć zastosowanie w systemach ekspertowych, inaczej mówiąc takie, które na wyjściu potrafią wygenerować reguły mogące służyć budowie bazy wiedzy SE. Ogranicza to ilość możliwych do zastosowania metod do pięciu- Scoring, Drzewa Klasyfikacyjne, Metoda Najbliższego Sąsiada (KNN), Zbiory Przybliżone, Modelowanie Rozmyte. Celem niniejszego artykułu jest wybór trzech najlepiej pasujących do budowy systemu ekspertowego metod do dalszych badań. Ocenie pod względem stosowalności w systemie ekspertowym poddano również aktualnie stosowaną metodę przez badany bank.

Do określenia najlepiej dopasowanych metod posłużono się metodą AHP (Analytic Hierarchy Process). Jest to jedna z wielokryterialnych metod hierarchicznej analizy problemów decyzyjnych. Pozwala na dekompozycję złożonego problemu decyzyjnego i stworzenie rankingów finalnego dla zdefiniowanego, skończonego zbioru wariantów. Metoda AHP potrafi łączyć w sobie subiektywne oceny ekspertów z obiektywną statystyczną analizą decyzji dotyczących hierarchicznie usytuowanych częściowych zależności i powiązań pomiędzy czynnikami opisującymi analizowany problem. Wykorzystuje się ją w wielu dziedzinach takich jak:

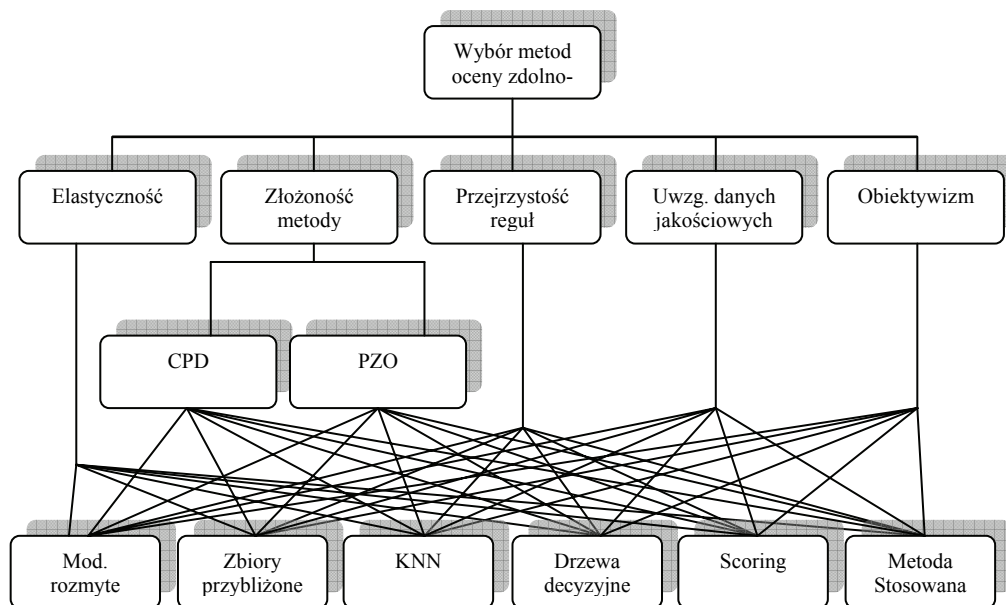
zarządzanie, politologia, socjologia, wytwarzanie, czy transport. [Omkarprasad 2006]. Algorytm składa się z czterech faz:

1. Utworzenia hierarchicznej struktury procesu decyzyjnego
2. Zdefiniowania preferencji decydenta oraz obliczenie ocen ważności dla wszystkich elementów hierarchii
3. Badania spójności macierzy preferencji
4. Tworzenia rankingu końcowego.

Pierwsza faza obejmuje kilka poziomów: określenia celu, kryteriów, podkryteriów i wariantów, przy czym poziomy podkryteriów nie są obowiązkowe. W badanym przykładzie celem było określenie trzech najlepszych metod do zastosowania w systemie ekspertowym oceny kredytobiorców. Analiza literatury przedmiotu i konsultacje z ekspertami wyłoniły pięć kryteriów, przy czym jedno podzielono na dwa podkryteria. Obiektywizm został uznany za czynnik dominujący; zdecydowano, że system powinien być w jak najmniejszym stopniu wrażliwy na subiektywną ocenę inspektora kredytowego. Zasadność stosowania systemów ekspertowych opiera się m.in. na założeniu, że są one, w przeciwieństwie do ekspertów odporne na stres, zmęczenie choroby, uprzedzenia, zapominalstwo lub złą wolę. [Niderliński 2006]. Subiektywna ocena eksperta miałaby znaczenie tylko w przypadku kredytów zakwalifikowanych jako ryzykowne lub wtedy, gdy w systemie nie znalazłaby się żadna reguła określająca dany przypadek. Obiektywizm w tym przypadku jest rozumiany jako brak wrażliwości metody na subiektywne oceny eksperta i uzyskuje wagę 0,42. Elastyczność to druga pod kątem istotności cecha wskazana przez ekspertów. Jest związana z możliwością prostej modyfikacji bazy wiedzy przez użytkownika bez konieczności naruszania integralności systemu wnioskującego. Jest to jedna z najbardziej istotnych cech systemów ekspertowych, gdyż w dziedzinach takich jak bankowość, lub medycyna sytuacja zmienia się bardzo dynamicznie i wymaga ciągłej aktualizacji zasad działania (aktualizacji reguł) [Niderliński 2006] Cecha ta związana jest również z możliwością samouczenia się systemu ekspertowego. Badane metody powinny generować reguły tak, aby bez problemu mogły zasilać bazę wiedzy. Wagę elastyczności metody wyznaczono na poziomie 0,25. Kolejna cecha to uwzględnianie danych jakościowych – odpowiada na pytanie w jakim stopniu w danej metodzie możliwe jest ujęcie danych jakościowych. Eksperti wskazali tę cechę jako trzecią w hierarchii istotności- waga 0,14. Poziom złożoności metody zdekomponowano na dwa podkryteria i przypisano wagę 0,05. Im większa złożoność metody tym bardziej skomplikowany może być proces pozyskiwania danych i trudniejsza jest aktualizacja bazy wiedzy. Wyróżnione podkryteria to czas przygotowania danych (CPD) i poziom złożoności obliczeń (PZO) koniecznych do uzyskania wyników. W tym przypadku paradoksalnie wysoka ocena poziomu złożoności oznacza że metoda jest prosta obliczeniowo a przygotowanie danych trwa krótko. Ostatnie oceniane kryterium to przejrzystość reguł z wagą 0,14 Przejrzystość reguł oznacza na ile logicznie reguły generowane przez metodę wyjaśniają decyzję. Kryteria i wagi określone zostały przy pomocy ekspertów zajmujących się tworzeniem systemów informatycznych, w

tym również systemów ekspertowych. Odpowiedzi udzielali pod kątem systemu ekspertowego oceny zdolności kredytowej. Podsumowując wybrano pięć kryteriów głównych a dla kryterium poziomu złożoności wyodrębniono dwa podkryteria. Struktura procesu decyzyjnego została przedstawiona na rysunku 2.

Rysunek 2. Struktura hierarchiczna procesu decyzyjnego



Źródło: Opracowanie własne

W drugim kroku preferencje decydenta określone są przy pomocy względnych ocen ważności kryteriów, podkryteriów i wariantów. Oceny te powstają poprzez porównywanie parami wszystkich obiektów znajdujących się na danym poziomie hierarchii. Warianty są porównywane przy pomocy stopnia intensywności preferencji. Określa się je dla każdego poziomu hierarchii. Najczęściej stosowana jest dziewięciostopniowa skala porównawcza, gdzie 1 oznacza równowagę wariantów a 9 ekstremalną przewagę. Macierze te cechuje spójność parami tzn.:

- dany element macierzy jest równoważny względem samego siebie: $a_{i,i} = 1$
- wartość oceny elementu b względem elementu a jest odwrotnością oceny a względem b $a_{i,j} = 1 / a_{j,i}$

Metoda dopuszcza istnienie pewnej niespójności globalnej macierzy - dopuszczalna jest sytuacja, w której element A jest preferowany względem B, B względem C, natomiast A nie jest preferowany względem C. [Bodin 2003]. Następnym krokiem jest wyznaczenie preferencji kryteriów i oceny wariantów ze

względem na dane kryterium. W tym celu dokonuje się normalizacji macierzy według wzoru

$$w_{ij} = \frac{a_{ij}}{\sum_{i=1}^n a_{ij}} \quad (1)$$

Lokalną wagą jest średnia wiersza znormalizowanej macierzy. Wagi (priorytety) globalne wyznacza się jako iloczyn lokalnej wagi tego elementu i wagi każdego z elementów poziomu bezpośrednio wyższego. Wagi kryteriów głównych przedstawia tabela 1.

Tabela 1. Wagi cech głównych

	O.	E.	U.C.J.	P.R.	Z.M.
Obiektywizm (O)	1	2	6	6	7
Elastyczność (E)	5	1	4	4	5
Uwzg. Cech Jakość. (UCJ)	0,167	0,25	1	1	2
Przejrzystość Reguł (PR)	0,167	0,25	1	1	2
Złożoność metody (ZM)	0,143	0,2	0,5	0,5	1

Źródło Opracowanie własne

W badanym przykładzie preferencje oceny dokonano posiłkując się literaturą opisującą poszczególne metody. Wybór preferencji konsultowano z ekspertami-pracownikami naukowymi znającymi wybrane do badań metody. Wagi kryteriów lokalnych przedstawia tabela 2.

Trzeci krok obejmuje sprawdzenie spójności ocen (czyli sprawdzenie logiki ocen – np. gdy $A > B$ i $B > C \Rightarrow A > C$) oraz obliczenia dwóch wskaźników: i.c (index of consistency) i c.r (consistency ratio).

Tabela 2. Podstawowe wskaźniki dla cech głównych

	Waga	Suma kolumn	Współczynnik λ	Wyniki
Obiektywizm	0,42	2,31	0,97	$\Lambda_{\max}=5,026$
Elastyczność	0,25	4,20	1,06	C.I. =0,07
Uwzg. Cech Jakość.	0,14	7,33	1,02	R.I.=1,12
Przejrzystość Reguł	0,14	7,33	1,02	C.R.=0,07
Złożoność metody	0,05	19	0,96	

Źródło Opracowanie własne

C.I. (wskaźnik zgodności) Jest ilorazem różnicy wartości sumy iloczynów uśrednionych wierszowo wartości znormalizowanych wag i kolumnowych sum (λ_{\max}), odpowiadających poszczególnym kryteriom i ilości elementów na danym poziomie do ilości elementów na danym poziomie pomniejszonych o 1

$$C.I. = \frac{\lambda_{\max} - n}{n - 1} \quad (2)$$

C.R. (stosunek zgodności) to iloraz C.I. przez R.I., gdzie R.I. to tzw. współczynnik losowej zgodności o wartości zależnej od liczby porównywanych elementów. Współczynnik można odczytać z odpowiednich tablic.

Wynik obliczany jest przy pomocy sumy iloczynów wagi wariantu dla określonych kryteriów i wag poszczególnych kryteriów. Następnie tworzony jest ranking wariantów. Wagi i oceny uszeregowane według oceny zostały przedstawione w tabeli 3.

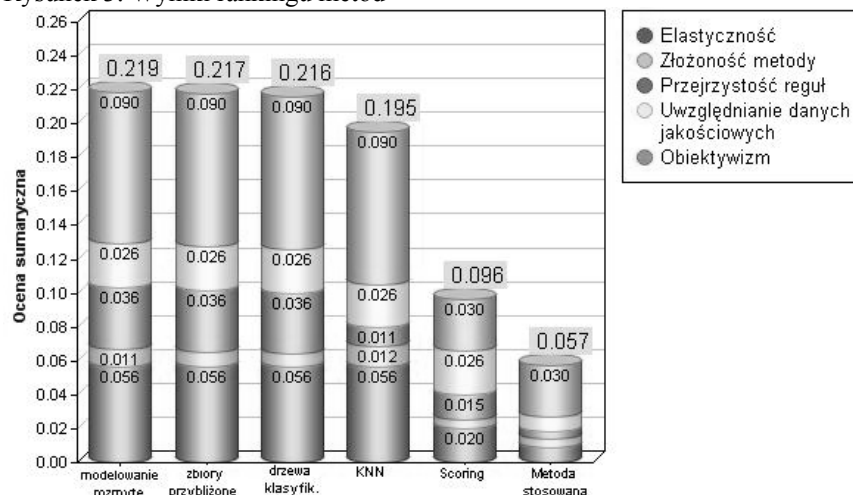
Tabela 3. Wagi wariantów dla kryteriów i uszeregowane oceny wariantów

	O.	E.	U.C.J.	P.R.	Z.M.	Ocena
Mod rozmyte	0,21	0,22	0,19	0,26	0,23	0,219
Zb. Przybliżone	0,21	0,22	0,19	0,26	0,18	0,217
Drzew Klasyf.	0,21	0,22	0,19	0,26	0,16	0,216
KNN	0,21	0,22	0,19	0,08	0,24	0,195
Scoring	0,07	0,08	0,19	0,11	0,10	0,096
M. Stosowana	0,07	0,03	0,06	0,03	0,09	0,057
waga kryterium	0,42	0,25	0,14	0,14	0,05	

Źródło Opracowanie własne

W badanym przykładzie obydwa współczynniki nie przekraczają wartości 0,01 na każdym poziomie, co w praktyce oznacza prawie pełną spójność ocen i preferencji (dopuszczalny stosunek zgodności C.R. jest równy 0,1). Ranking końcowy przedstawia się w sposób przedstawiony na rysunku 3.2.

Rysunek 3. Wyniki rankingu metod



Źródło: Opracowanie własne na podstawie AHP Projekt

Wyniki wskazują przewagę trzech metod: modelowania rozmytego, zbiorów przybliżonych i drzew klasyfikacyjnych. Metody te uzyskały bardzo podobne oceny, co wskazuje na brak dużych różnic w preferencjach między nimi. Tym samym każda z nich wymaga przeanalizowania pod kątem stosowalności w systemach decyzyjnych oceny kredytobiorców. Metody te zwyciężyły w rankingu ze względu na niezależność od subiektywnych ocen ekspertów i możliwości szybkiej aktualizacji reguł, a były to czynniki które oceniono jako najbardziej istotne. Nie bez wpływu pozostała możliwość ujmowania danych jakościowych oraz przejrzystość ocen (reguły w postaci If ... Then ...). Najniżej oceniono metodę aktualnie stosowaną przez badany bank do oceny kredytobiorców. Nie uwzględniała w ocenie wielu danych jakościowych oraz nie generowała reguł, dlatego jej przydatność w zastosowaniu w systemie ekspertowym jest najniższa. Wybrane metody oceniano następnie pod kątem skuteczności na przykładzie zastosowania do oceny zdolności kredytowej klientów banków udzielających kredyty dla rolnictwa. Metody te mogłyby być stosowane również dla innych klas kredytów. System ekspertowy oparty na regułach pochodzących ze zbiorów przybliżonych lub model rozmyty mogą z powodzeniem służyć jako system decyzyjny wspomagający inspektora kredytowego dla dowolnej klasy kredytów. Warunkiem koniecznym jest jednak posiadanie odpowiedniej ilości próbek uczących tak, aby reguły charakteryzowały się wysokim poziomem wsparcia.

LITERATURA

- Bodin L. , Gass S.I. (2003) On teaching the analytic hierarchy process *Computers & Operations Research*. Volume 30. 10.
- Migut G. (2003) Modelowanie ryzyka kredytowego, Materiały Konferencyjne „Zastosowanie Statystyki i Data Mining w Finansach”, Warszawa 2003r.
- Niderliński A. (2006) Modelowo-regułowe systemy ekspertowe rmse”, Wyd. Pracowni Komputerowej Jacka Skalmierskiego, Gliwice.
- Omkarprasad S. Vaidya, Sushil Kumar (2006) Analytic hierarchy process: An overview of applications *European Journal of Operational Research*. Volume 169. 1 2006r.
- Staniec I. (2007) Wybrane aspekty działalności polskich banków”, Materiały z internetu: <http://www.oizet.p.lodz.pl/istan/> data pobrania: 08-03-2007r.

Applying AHP to valuating methods fitting to credit borrower classification expert system

Abstract: The fundamental activity of most commercial banks is crediting. The risk of such activity is quite high, so banks are obliged to check borrowers possibilities to pay debt. But banks may decide about method of that measure. Using expert system in such cases seems to be good idea. The problem with that is to find method which will generate logic and precise rules. In this paper author uses AHP method to choose methods which are best fitted to use as rule generators for expert system.

Keywords: AHP, credit ability

THE USEFULNESS OF MULTINOMIAL LOGIT MODELS IN EXAMINING THE RELATIONSHIP BETWEEN THE OWNERSHIP FORM OF AGRICULTURAL ENTERPRISES AND THEIR FINANCIAL-ECONOMIC RESULTS

Stanisław Jabłonowski, Andrzej Kluza

Katedra Informatyki, SGGW

e-mail: staja@mail.sggw.pl

Summary: From the year 1994 onwards, the magazine "Nowe Życie Gospodarcze" has been publishing financial-economical data of the best three hundred agricultural enterprises selected from groups of such enterprises that had chosen to participate. Makers of these lists have taken into account enterprises which have come into being from the property of state treasury in the nineteen nineties. The present paper takes up the problem of utilizing (to a rather limited extent) of multinomial logit models for studying the relation between the ownership form of agricultural enterprises and some financial-economic characteristics, describing economic effectiveness.

Keywords: agricultural enterprises, ownership forms, economical effectiveness indicators, multinomial logit models

STUDY OBJECTIVE

Our investigation is restricted to the so-called "top 300 lists", presented by "Nowe Życie Gospodarcze". In this data, we find a description of economic entities that have arisen after the year 1989 from the estate that had remained in the hands of A.W.R.S.P.'s (then A.N.R.¹). In the paper, the question is raised whether it is possible to build econometric models where the qualitative explained variable corresponds to the form of property, while explanatory variables are constituted by financial-economic characteristics of agricultural industry enterprises. We ask also whether building such models makes sense, since it is not possible to use these models to observe causal links between the explanatory variables and the explained variable (which here is not an effect but cause of differences in the financial-

¹ A.W.R.S.P. = Agency for Agricultural Property of State Treasury; A.N.R. = Agricultural Property Agency

economic results). So, we want to focus on examination of existence of statistical links between features, ignoring indices of cause and effect. The causal link between explanatory variables and the explained one is important in qualitative variables' models [Cramer, 2001]. An approach presented below will be useful, in spite of that we obviously won't profit fully of such models. In the case of our data there exists a number of obstacles in building an econometric model in which a form of ownership would be an explanatory or one of explanatory variables (it will be explained later). So, analyzing the populations' data, it would be necessary to use common descriptive statistics only, and to use also some measurements of dependences between features.

The bodies from "top 300 lists" have various ownership and legal forms. Models are built separately for each year of the eleven year-time period under analysis. The models consider the set of three hundred entities featured by "Nowe Życie Gospodarcze" on the "top 300 list" for a given year.

The authors attempt to study the following idea: does a statistical link between economic performance and the ownership form of companies in these sets exist or not? We want to know which financial-economic features are significant in the constructed models. Does the utilization of models allow us to obtain greater benefits regarding the analysis of the sets under investigation or do we arrive at better results from the usual descriptive statistics?

An econometric model with a qualitative explained variable with many variants may take the form of a multinomial logit model. This form has been used in the paper.

DATA DESCRIPTION

Lists of three hundred best companies and other agricultural entities that had come into being from the property of state treasury have been scientifically processed by the I.E.R.i G.Ż (Institute of Agricultural Economics and Food Economy), while the initiators of their preparation have been the A.W.R.S.P. (Agency for Agricultural Property of State Treasury) [now A.N.R. - Agricultural Property Agency] and the magazine "Nowe Życie Gospodarcze".

The first list concerned the year 1994 and included only two hundred units. However, in the following years, the lists described three hundred best companies of the agricultural sector (from among those that would respond to the survey). The entities taken into account have been enterprises that had emerged from the property of state treasury and that remained in the hands of the A.W.R.S.P (restructured into the A.N.R. in 2003). They received varying ownership and legal status.

Rankings of companies were drawn up on the basis of responses to questionnaires sent out to a large number of units (e.g., to 2110 units in 1998). Usually, only less than 20% addressees would dispatch their answers (for example, in 1998, 386 entities completed the survey). The basic rating was originally created

on the basis of the added value indicator, then on the basis of the profitability index and, since 2000, according to a special synthetic measurement tool (then slightly modified).

Data concerning the best companies have been published in „Nowe Życie Gospodarcze” and supplemented by comments in the form of articles written by eminent specialists in the field of agricultural economics [e.g. Leopold, Ziętara, 1999]. The sole standard descriptive statistics have been used in the comments. The data was used in many analyses presented in other titles [e.g. Franc-Dąbrowska, 2008; Grontkowska, 2007]. This paper utilizes eleven lists from the years 1995 - 2005.

In individual lists, the following information is given for particular enterprises: current position in the rating and the position from the previous year (if the object was featured on the list the previous year), name of the company and its voivodship of origin (unless there was no consent), the form of property management (ownership), the EKD (later PKD) code which denotes the dominant socio-economic activity of an economic entity. Further, the features' definitions and their designations have been given. The designations come from the authors' research of this paper and are used in the further description. The absence of certain digits in the numbering is due to the fact that some features published in the ratings have been neglected, since they do not contribute significant to any of the econometric models constructed. Detailed information on the measurement tools and indicators included on the lists can be found e.g. in [Guzewicz et al., 2006]. The successive features are as following:

indicator of profitability of economic activity (operating activity in the 1995) (W6), return on total assets (W7), return on equity (W8), the rate of value added (W9), current ratio at year-end (W10) and quick ratio at year-end (W11), equity to assets ratio (W12), financial results to total debt ratio (W13), labour efficiency (W14), total emoluments (W15), sales income (W16), net financial balance (W19), average employment in year (W20), arable land area (W22) and soil quality class (W23). In addition, two variables have been attached: employment per 100 hectare of arable land (W24) and land productivity (emoluments per 1 hectare of arable land) (W25), calculated on the basis of data from the lists.

MODELLING METHOD

The model type that frequently becomes used in case the feature to be explained has a small number of variants, but still more than two, is the multinomial model of qualitative variables, otherwise known as the model of multiple choice. However, in case of disordered categories, where determination of the order of the nominal feature variations proves impossible, researchers usually choose the logit multinomial model. Such models are used in different fields. For example Mikhail, Walther and Willis [1998] use such a model to research the phenomenon if a change of brokers' office by an financial analyzer has had place

on effect of the quality of prognosis made by him. Cramer [2003] analyzes different stages of private cars' ownership according to different features of households. Seo and Mendelsohn [2008] describe how South American farmers adapt to climate changes by changing crops by developing a multinomial logit model of farmer's choice of crops. In the present analysis, we are interested in the form of property management that a companies utilize. This feature accepts 9 variants for the data in question. These are forms of ownership (abbreviations used in the "top 300 lists" are given in brackets): the enterprises purchased from the Agency (PZA), mixed units, i.e. these that have purchased some part of the land, while the rest is leased (PM), entities in that a company established by employees holds the lease of the utilized plot (DSP), leases concluded by natural persons (DOF), other leases (DP), companies wholly owned by the Agency (JSA) and administered enterprises (A). Number of administered enterprises (A) covered by the ratings gradually decreased (from 102 in 1995 to 0 in 2004), but since 2001, agricultural production cooperatives (RSP) and the other units (PJ) are also included in the ratings. There are only between 2 and 10 cases of the latter entity (PJ) and this code name usually denotes units that belong to scientific institutions. In this paper, entities with the legal form A (since 2001) and PJ have been disregarded. Overall, it can be said that the entities with the legal form A and JSA have conducted business on behalf of their mother Agency, while the rest - on their own account.

For particular "top 300 lists" from consecutive years, multinomial logit models have been built. The explained variable is a form of ownership, but its new categories are included, having applied four methods of classifying ownership forms – as shown in the tab. 1 (the symbols SP1, SP2, SP3, SP4, that denote the explained variable, correspond to the way of aggregating):

Table 1. Methods of assigning the form of property to the created variants

Property's form → ----- Way of clustering	PZa	PM	DSP, DOF, DP	JSA	A (until 2000)	RSP (since 2001)
Sp1 (4 variants)	1	1	2	3	4	4
Sp2 (5 variants)	1	2	3	4	5	5
Sp3 (3 variants)	1	1	1	2	3	3
Sp4 (3 variants)	1	1	2	3	3	3

Source: authors' decisions

Potential explanatory variables have been mentioned on the "top 300 lists", but in the course of building the models in question some of them have been eliminated. Finally, for each of the eleven years, one model has been left, featuring the explained variable (out of SP1, SP2, SP3 and SP4) that occurred in the model with the best characteristics.

In the section “Study Objective” the following is stated: the resulting models are to describe the sets of 300 companies from the “top 300 lists”. Unfortunately, the models cannot be generalized to cover the entire set of enterprises that had arisen from state treasury property managed by the A.W.R.S.P. (and later the A.N.R.) for particular years. The “top 300 lists” cannot be treated as representative samples selected from the entire set because it is not the case, as described above. Therefore, conclusions derived from constructed models refer only to entities listed on the “top 300 list” for the year in question.

MATHEMATICAL APPROACH TO THE MODEL

The applied econometric model type derives from Theil [Theil, 1969]. In this model, the aim is to estimate the values connected with probabilities of obtaining different categories by the explained variable. If Y is a qualitative variable that adopts $J+1$ variants and if P_{ij} denotes the probability of obtaining the j -th variant by this variable for the i -th element of the sample (population, once generalized), then the multinomial logit model explains the quotients of probabilities $P_{ij}/(P_{i0} + P_{ij})$, $j = 1, \dots, J$, $i = 1, \dots, n$. Index $j = 0$ corresponds to the chosen basic variant. This explanation is done by means of a function that represents cumulative distribution of logistic distribution. The argument of this function is a linear combination of explanatory variables and the vector of structural parameters of the model. However, there is a number of structural parameter vectors. The vectors are different for each j ($j = 1, \dots, J$).

Let us denote the vector of structural parameters for the category j of $k+1$ size by β_j , $j = 1, \dots, J$, if we assume that there are k explanatory variables (the free coefficient on the first position).

In the case of a model built on the basis of individual data it can be summarized as

follows: $\frac{P_{ij}}{P_{i0} + P_{ij}} = F(\mathbf{x}_i^T \beta_j)$, where \mathbf{x}_i is the vector (of dimension $k+1$) consisted

of the number 1 and of explanatory variables' values for the i -th element of the sample (population), $i = 1, \dots, n$, $j = 1, \dots, J$ ².

Since $F(\cdot)$ is the logistic function, finally: $\frac{P_{ij}}{P_{i0} + P_{ij}} = \frac{1}{1 + \exp(-\mathbf{x}_i^T \beta_j)}$,

$j = 1, \dots, J$ for i -th object, $i = 1, \dots, n$.

² As the ratios of certain probabilities are explained and not single probabilities, so it can be shown, that there is no contradiction with the condition $\sum_j P_{ij} = 1$.

Hence, it can be deduced that: $\text{Ln} \frac{P_{ij}}{P_{i0}} = \mathbf{x}_i^T \boldsymbol{\beta}_j$, for each category $j = 1, \dots, J$ [Gruszczyński, 2002]. Predicted category for the i -th element of the sample (after generalization, for any element of the population) is selected after evaluation of J vectors of structural parameters' estimates. The estimates of expressions $\text{Ln} \frac{P_{ij}}{P_{i0}}$, $j = 1, \dots, J$, are to be calculated. Then this category of $J+1$ of the Y variable's variants possible to be chosen is to be selected for which the estimate of the likelihood of its occurrence is the highest.

MODEL VERIFICATION

Evaluating parameters of a multinomial logit model is most often carried out by means of the principle of maximum likelihood [Cramer, 2001]. Most considerations may be regarded as a generalization of the binomial logit model, see e.g. [Dudek et al., 2006].

Long [Long, 1997] states that the maximum likelihood method should be applied for a large sample of at least 100 items, but preferably of about 500. The significance of parameters may be assessed in various ways, the most important being the likelihood quotient test. It is used for a larger set of parameters, while Wald statistics is used for individual structural parameters. All these statistical approaches have a chi-square distribution. Here a description of Wald statistics is skipped, indicating only that it has 1 degree of freedom.

The following provides a more detailed information on the likelihood quotient test. Logarithm of this ratio is calculated by the formula: $\text{LR}_j = -2(\ln \hat{L}_{R_j} - \ln \hat{L}_{UR})$, where $\ln \hat{L}_{R_j}$ means the logarithm of maximum of likelihood function for the model with the free coefficient, but which does not include the variable X_j from the k explanatory variables X_1, \dots, X_k (or without free coefficient and with all variables, when $j = 0$). The expression $\ln \hat{L}_{UR}$ is the logarithm of maximum of likelihood function for the full model, i.e. with the free coefficient and k variables.

Statistics LR_j for large samples follows the chi-square distribution with J degrees of freedom. The verified hypothesis is: $H_0: \beta_{1j} = \dots = \beta_{kj} = 0$, against the alternative H_1 : there exists r from the set $\{1, \dots, J\}$, that $\beta_{rj} \neq 0$. Statistics LR_j serves to its verification and the test is to be carried out separately for the successive variables and the free coefficient ($j = 0, 1, 2, \dots, k$).

The second hypothesis refers to the absence of statistical significance of all structural parameters of the explanatory variables. The verified hypothesis is: $H_0: \boldsymbol{\beta}_r = [\beta_{r0}, 0, \dots, 0]$ for all $r = 1, \dots, J$, against the alternative H_1 : there exists r from

the set $\{1, \dots, J\}$ and j from the set $\{1, 2, \dots, k\}$, that the parameter β_{ij} is different from zero. And here again the test of likelihood quotient is used to verify the hypothesis. Now it takes the form: $LR = -2(\ln \hat{L}_R - \ln \hat{L}_{UR})$, where $\ln \hat{L}_R$ denotes the logarithm of maximum of likelihood function for the model reduced to only a constant, while $\ln \hat{L}_{UR}$ is the logarithm of maximum of likelihood function for the full model, i.e. with the free coefficient and k variables.

For large samples the LR statistics follows the chi-square distribution with $k \cdot J$ degrees of freedom.

In assessing the compatibility of the model with empirical data a lot of different measurements are used. They are similar, in its conception, to the classical coefficient of determination R^2 .

One of frequently used measurements is so-called McFadden pseudo- R^2 [Manski et

al., 1977]: $R^2_{McFadden} = 1 - \frac{\ln \hat{L}_{UR}}{\ln \hat{L}_R}$, where $\ln \hat{L}_R$ stands for the logarithm of

maximum of likelihood function for the model reduced to only a constant, while

$\ln \hat{L}_{UR}$ is the logarithm of maximum of likelihood function for the full model, i.e.

with the free coefficient and k variables. $R^2_{McFadden}$, as well as classical R^2 , accepts values between 0 and 1, a better model fit is characterized by its higher value.

Similarly, count- R^2 [Gruszczynski, 2002] takes values between 0 and 1, a better model fit is characterized by its higher value. It is the quotient of the number of correct predictions to the number of all elements of the population, for which predictions have been estimated.

After having verified the model, an interesting analysis can be made. E.g.

from the formula: $\frac{P_{ij}}{P_{i0}} = \exp(x_i^T \beta_j)$, in which the left side can be called the “odds

ratio of j category to the basic category”, the following conclusion can be easily

derived: $\exp(\beta_{j1})$ informs how many times this odds ratio increases, when X_1 increases by a unit, ceteris paribus. This ratio increases with the increase of X_1 , if only $\beta_{j1} > 0$.

Similarly, it can be shown that $\frac{P_{ij}}{P_{ik}} = \frac{\exp(x_i^T \beta_j)}{\exp(x_i^T \beta_k)}$, and therefore we

conclude that $\exp(\beta_{j1} - \beta_{k1})$ informs how many times the quotient $\frac{P_{ij}}{P_{ik}}$ increases

if the variable X_1 is increased by a unit, ceteris paribus. Hence, if $\beta_{j1} > \beta_{k1}$, then with the increase of X_1 , the quotient increases, or the probability of category j relative category k increases.

RESULTS OF MODEL EVALUATION

The authors have decided to build models for successive years, involving all potential explanatory variables. Then the a posteriori elimination method has been used to obtain end models that have positively verified hypotheses about the significance (on the level of $\alpha = 0.1$ at the most) of structural parameters (all three hypotheses discussed above). Count- R^2 had to be greater than 50%, while the McFadden pseudo- R^2 had to equal circa 0.20 at least. For each of the eleven years, one model remained – the one with the explained variable (from among SP1, SP2, SP3 and SP4) that has occurred in the model with the best characteristics.

As the sample size for each year was about three hundred (in case of 1995, the data of two companies was removed because of their deficiencies; for some years, the sample size decreased slightly due to the omission of entities of legal form PJ and legal form A since 2001), so it can be stated that the requirements pertaining to the minimum number of statistical units were met [Long, 1997].

Table 2. Test results to prove the lack of statistical significance of all structural parameters of the model

Model	$-2 \ln \hat{L}_R$ or $-2 \ln \hat{L}_{UR}$	LR	Degrees of freedom	P-value
Only constant	634.31			
The full model	437.17	197.15	12	0

Source: the SPSS software results

The SPSS software version 11.5.0 (16 Nov. 2002) has been used to complete necessary calculations. Results of the model estimated on the basis of data concerning 291 objects from the list of 2005 are presented below. SP4 has been chosen as the explained variable, while the selected explanatory variables are W7, W8, W12, W14, W20 and W24. Following the description of these results, the summary of findings for all models is also presented.

Table 3. Test results to prove the lack of statistical significance of successive structural parameters of the model

Parameter	$-2 \ln \hat{L}_{R_j}$	LR_j	Degrees of freedom	P-value
Constant	443.42	6.25	2	0.044
W7	461.3	24.13	2	0
W8	492.62	55.45	2	0
W12	466.67	29.51	2	0
W14	466.09	28.92	2	0
W20	455.01	17.85	2	0
W24	467.21	30.05	2	0

Source: the SPSS software results

Table 4. Evaluation of parameters (B) for successive categories (SP4 is the explained variable), together with test results to prove the lack of statistical significance for individual structural parameters (3 is the variant of reference)

variant	parameter	B	Wald	degr. of freed.	P-value	95% confidence interval for Exp(B)	
						lower limit	upper limit
1	Constant	1.69	3.11	1	0.08		
	W7	-0.057	4.28	1	0.04	0.944	0.894
	W8	0.124	18.69	1	0.00	1.133	1.070
	W12	-0.056	24.45	1	0.00	0.945	0.925
	W14	0.036	20.29	1	0.00	1.036	1.020
	W20	-0.013	8.11	1	0.00	0.987	0.979
	W24	0.05	9.14	1	0.00	1.051	1.018
2	Constant	2.453	5.87	1	0.02		
	W7	-0.143	18.58	1	0.00	0.867	0.812
	W8	0.178	36.63	1	0.00	1.195	1.128
	W12	-0.046	15.53	1	0.00	0.955	0.934
	W14	0.027	10.82	1	0.00	1.027	1.011
	W20	-0.013	7.39	1	0.01	0.988	0.979
	W24	0.055	10.92	1	0.00	1.056	1.023

Source: the SPSS software results

We shall describe the selected explanatory variables more precisely: W7 is return on total assets (ratio of value added to total assets used by the company in the end of 2005), W8 - return on equity (ratio of net financial result to level of equity in the end of 2005), W12 - equity to assets ratio (ratio of equity to assets in the end of 2005), W14 - labour efficiency (ratio of value added to average employment in 2005), W20 - average employment in the year (average number of people employed in 2005), W24 - employment per 100 hectare of arable land (the

variable $W20*100$ divided by arable land area). McFadden pseudo- R^2 for the model of 2005 is equal to 0.311.

Table 5. Prediction results according to the model for 2005 (the count- R^2 can be found in the lower right corner, expressed in per cents)

Categories				
Observed	Predicted			Percentage of correct predictions
	1	2	3	
1	41	22	22	48.24
2	22	56	13	61.54
3	11	5	99	86.09
Percentage totally	25.43	28.52	46.05	67.35

Source: the SPSS software results

Table 6. Summary results for all models

Year	Explained variable	McFadden - R^2	Count - R^2	Explanatory variables
1995	Sp4	0.736	93.6%	W12, W16, W20
1996	Sp4	0.655	90%	W6, W7, W9, W10, W12, W13, W20, W23, W24, W25
1997	Sp4	0.602	86.3%	W6, W11, W12, W13, W16, W20
1998	Sp3	0.477	81.7%	W12, W14, W20
1999	Sp3	0.329	79%	W13, W19, W20
2000	Sp4	0.317	68.7%	W6, W9, W12, W20
2001	Sp4	0.324	68.7%	W6, W12, W14, W20
2002	Sp3	0.319	74.7%	W10, W12, W14
2003	Sp3	0.3	74.3%	W7, W12, W13, W20
2004	Sp4	0.187	56.8%	W6, W12, W13, W14
2005	Sp4	0.311	67.4%	W7, W8, W12, W14, W20, W24

Source: The author's own calculations based on the SPSS software results

SUMMARY OF RESULTS AND CONCLUSIONS

As shown by Table 6, in all models the variables explained are SP3 or SP4, so the remaining variables are these concerning three variants of the ownership form (cf. Tab.1). The best model is the model for 1995 data. Models for subsequent years provide a worse match, but McFadden- R^2 is generally higher than 0.3, and count- R^2 usually reaches at least 70%.

Testing the significance of remaining variables has given very good results. It may be stated that verification of statistical models has been positive. Thus, a statistical link is visible between financial-economic performance and the form of ownership of the enterprises from the "top 300 lists". As Table 6 indicates, this relationship is more pronounced in the earlier years of the time period in question. Not in every

year the same economic characteristics become significant in the constructed models. The explanatory variables that occur in the greatest number of models are as follows: W12 (10 models), W20 (9), W6 (5), W13 (5), W14 (5), W7 (3). Other variables - W9, W16, W24 – have occurred in two models, and W8, W10, W11, W19, W23 and W25 – in one model. The variables that appears most commonly are: W12, which denotes the ratio of equity to assets at the end of the year - expressed as a percentage - and W20, meaning the average employment in the year (measured in the number of persons).

Interpretation of a model will be shown on the example of the 2005 model. SP4 has become the explained variable. It attains three variants: the first means entities wholly or partially purchased (PZA, PM) , the second - leases (DSP, DOF, DP) and the third - companies wholly owned by the Agency and agricultural production cooperatives (JSA, RSP) (in 2005, administered entities, A, were not present).

The direction of the impact of explanatory variables on the odds ratio of both variant 1 and variant 2 to variant 3 (category 3 is the basic one) remains the same, because the parameters corresponding to the variables retain the same signs for variant 1 and variant 2. And so: an increase in the variable W7 decreased these quotients, while an increase in the variable W8 increased them - and so on, as seen in Table 4. Unfortunately, because of the fact that the ownership form is the cause and not the effect of differences of values of variables W7 or W8 we only can deduce that W7 reaches greater average values for variant 3 and W8 – lower than for variants 1 and 2.

Due to similar reasons the following analysis for e.g. variable W12 cannot be utilised well. For this variable (the ratio of equity to assets expressed in per cents), the estimated parameters are: $\beta_{1W12}=-0.056$, $\beta_{2W12}=-0.046$ (both significantly different from zero).

Therefore, the conclusion can be drawn that as the value of variable W12 increases, the quotient of probability of variant 1 to the probability of variant 3 decreases (0.945 times per unit of W12) and the ratio of probability of variant 2 to the probability of variant 3 also decreases (0.955 times per unit W12), *ceteris paribus*.

One could also presume that with an increase in the value of the variable W12 the ratio of probability of variant 1 to the probability of variant 2 decreases (0.99 times per unit of W12), *ceteris paribus*. It does not have to be so, however, due to the lack of significant difference between the parameters β_{1W12} and β_{2W12} , as seen in Table 4. One should concentrate on the 291 companies involved in the calculations, out of the three hundred entities found on the list for 2005.

Let us clear up why a form of ownership is not the explanatory variable, as the probable cause of economic discrepancies i.e. why the models show only a quantitative relationship between variables, without specifying the cause and effect relation. In the case of “top 300 lists” the enterprises in the ratings are ordered by a special synthetic measurement tool. This is a kind of weighted average of features

W6, W9, W13 i W14, which, by ratings' authors, best express different aspects of economical effectiveness. Because of many reasons, the approach presented in the present paper is more convenient to use. For example, it would be difficult to determine the explained variable which would show economical effectiveness. If the synthetic measurement instrument tool a.n., used in the analyzed ratings, is supposed to be the explained variable then only one variable (W14) from 4 variables included in it has a significant link to the form of ownership (for the 2005 model).

It would be also difficult to establish optimal grouping of the ownership forms into lower number of variants. They seem to be too numerous in the source data. It is visible from the presented approach that the variable corresponding to the form of property takes into account the optimum allocation of primary forms of ownership into new variants. It becomes clear which variables are significantly linked to the ownership form.

Another important question is whether it would not be sufficient to use the standard approach i.e. to calculate usual descriptive statistics for the analysis of these sets, possibly with carrying out tests to check the discrepancies in averages.

Of course it would be of worth to put into use the known measures of relationship between features [Encyclopedia ..., 2008], like rank correlation coefficients (of Spearman, Kendall), or correlation coefficients based on chi-square test (Chuprov T-coefficient, Cramer's contingency coeff., Yule's coeff. of association ...). However, application of these measures (of Spearman, Kendall) requires the values of the variable Y to be ranged and each chosen order is very arguable. On the other hand, setting ranks to values of potential explanatory variables, which are continuous, would be meaningful when the space of their possible values would be constrained. It could be done by laying down variants' values corresponding to certain brackets of these variables' values. Arrangement of number and borders of these brackets is very arguable, too. Correlation coefficients based on chi-square test also require an arrangement of number and scopes of intervals of continuous variables' values. So, different results are obtained for different arrangements. Finally, calculations of these different coefficients may only provide an intuition about relationship between the variable Y and the others. It is impossible to fix objectively, which variables are stronger and which weaker correlated.

Without minimizing the rank of the standard approach, it appears that the development of presented models is at least competitive. For example, for the year 2005 the average values of the variable W14 may be calculated and the variant 1 of the variable SP4 becomes equal to 90.59, for the variant 2: 74.69 and for the variant 3: 51.5. From the analysis of variable W14 in the model of the year 2005 it can also be concluded that the average values of this variable for variants 1 and 2 are significantly higher than for variant 3.

First of all the approach selected by authors of the present paper shows how forms of ownership should be grouped into a smaller number of variants. In

addition, a comprehensive approach is imbedded in the model: the set of variables found in it has a significant link to the variable that corresponds to the form of ownership. Hence the conclusion which financial-economic characteristics mostly coincide with the variant of ownership form.

Recapitulating, despite of considerable limiting of profits flowed from qualitative variable's model's building, the approach seems to be advantageous anyway. First of all it allows to look at the examined populations comprehensively and it is competitive in relation to the standard approach.

LITERATURE

Cramer J.S. (2001), *An Introduction to the Logit Model for Economists*, TCL, London

Cramer J.S. (2003), *Logit Models from Economics and Other Fields*, Cambridge Univ. Press, Cambridge

Dudek H., Dybciak M. (2006), Application of the Logit Model to Analyze the Test Results, *Zesz. Nauk. Ek. i Org.Gosp. Żywn.*, No. 60, Warsaw, p. 81-92 (in Polish)

Encyclopedia of Statistical Sciences (2008), J.Wiley&Sons

Franc-Dąbrowska J. (2008), Separation of Property from Management and Efficiency of the Management of Selected Factors of Production, *Roczniki Nauk Rolniczych, Serie G, T.95, z.3 / 4*, p. 70 – 82 (in Polish)

Grontkowska A. (2007), Changes in Assets and Results of ANR's Partnership Enterprises in the Years 1996 - 2005, *R.N.R., S. G, T.94, z.1*, p. 66 – 73 (in Polish)

Gruszczyński M. (2002), *Models and Forecasts of Qualitative Variables in Finance and Banking*, OWSGH, Warsaw (in Polish)

Guzewicz W., Kulawik J., Osuch D. (2006), How the Ranking has Been Made. *Methodological Notes, N. Ż. Gosp.*, 22/430 (App. "Lista 300"), p. 24-25 (in Polish)

Leopold A., Ziętara W. (1999), The Fifth List, *Nowe Życie Gospodarcze*, 39/227, p. 19 – 25 (in Polish)

Long S. (1997), *Regression Models for Categorical and Limited Dependent Variables*, Sage Publ., Thousand Oaks, London

Manski C.F., McFadden D. (eds.) (1977), *Structural Analysis of Discrete Data with Econometric Applications*, MIT Press, Cambridge, Mass.

Mikhail M.B., Walther B.R., Willis R.H. (1998), Does forecast accuracy matter to analysts, *Social Science Research Network Electronic Library*, www.ssrn.com

Seo S.N., Mendelsohn R. (2008), An analysis of crop choice: Adapting to climate change in South American farms, *Ecological Economics*, 67, pp.109-116.

Theil H. (1969), A Multinomial Extension of the Linear Logit Model, *International Economic Review*, 10, pp. 251-259.

Przydatność wielomianowych modeli logitowych do badania związku formy własności przedsiębiorstw rolniczych z wynikami finansowo-ekonomicznymi

Streszczenie: W artykule podjęty został problem pewnego ograniczonego wykorzystania wielomianowych modeli logitowych do zbadania związku formy własności przedsiębiorstw rolniczych z pewnymi ich charakterystykami finansowo - ekonomicznymi, opisującymi efektywność gospodarowania. Do badań wykorzystano dane z tzw. „list 300” publikowanych przez "Nowe Życie Gospodarcze". W kolejnych, traktowanych odrębnie latach wybierano po trzysta najlepszych przedsiębiorstw rolnych spośród tych, które chciały wziąć udział w rankingu. Twórcy list brali pod uwagę przedsiębiorstwa, które powstały z majątku Skarbu Państwa w latach dziewięćdziesiątych XX w

Słowa kluczowe: przedsiębiorstwa rolnicze, formy własności, wskaźniki efektywności gospodarowania, wielomianowe modele logitowe

APPLICATION OF EBLUP ESTIMATION TO THE ANALYSIS OF SMALL AREAS ON THE BASIS OF POLISH HOUSEHOLD BUDGET SURVEY

Alina Jędrzejczak^{1,2}, Jan Kubacki²

¹Katedra Metod Statystycznych UŁ, ²USLORB

e-mail: j.kubacki@stat.gov.pl

Abstract: In the paper the results of small area estimation using empirical best linear unbiased predictor (EBLUP) for the data coming from Polish Household Budget Survey are presented. The results were obtained using small area models of household expenditures for regions. Estimation of sampling errors was conducted by means of the balanced repeated replication (BRR) technique. The estimation of EBLUPs and their corresponding mean square errors (MSE) was carried out using variance components technique. To calculate MSE of EBLUP the maximum likelihood method (ML) and restricted maximum likelihood method (REML) were used. The computation was made using SAE package designed for R-project.

Key words: small area estimation, empirical best linear unbiased predictor (EBLUP), household budget survey, variance estimation

INTRODUCTION

The main objective of many modern sample surveys is to provide estimates of totals, means and other parameters not only for the population but also for subpopulations (or domains) such as geographic areas and socio-economic groups. Direct estimates of domain parameters are based only on domain specific data. It is seldom possible to have overall sample size large enough to support reliable direct estimates for all the domains of interest. Therefore, it is often necessary to use indirect estimates that „borrow strength” by using values of the variables from related areas and thus increase the „effective” sample size.

The paper is focused on the following two areas:

- To present shortly the theoretical background of the EBLUP estimation for the basic area level models.

- To apply the EBLUP theory in order to improve the results of the estimation in the case of the household expenditures in Poland.

EBLUP ESTIMATION

Small area means or totals can be expressed as linear combination of fixed and random effects. Best linear unbiased prediction (BLUP) estimators of such parameters can be obtained in a classical way using BLUP estimation procedure. BLUP estimators minimize Mean Square Error (MSE) within the class of linear unbiased estimators and do not depend on the normality of random effects. Maximum likelihood (ML) or restricted maximum likelihood (REML) methods can be used to estimate the variance and covariance components, assuming normality.

The EBLUP procedure was applied in many important statistical surveys conducted all over the world. The pioneer work in this area was that of Fay and Herriot [Fay and Herriot 1979], where EBLUP technique was used for evaluating per capita income and some other statistics obtained for counties. Application of EBLUP estimators for the Survey on Life Conditions in Tuscany by Pratesi and Salvati [Pratesi and Salvati 2008] introduced the spatial EBLUP. EBLUP estimators were also discussed in detail in the EURAREA project [EURAREA consortium, 2004], where one of the discussed variables was household equivalent income. A report entitled “Social Exclusion and Integration in Poland: An Indicators-based Approach” [UNDP, 2006] prepared for UNDP (United Nations Development Programme) presented the benefits obtained from the EBLUP technique applied for average equivalent income.

The examples described below are based on a special kind of the general linear mixed model which is widely known as basic area level model [Rao, 2003]

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{v} + \mathbf{e} \quad (1)$$

where: \mathbf{y} - is $n \times 1$ vector of sample observations, \mathbf{X} - known matrix of explanatory variables, $\boldsymbol{\beta}$ - is vector of linear regression coefficients, \mathbf{v} - area-specific random effect vector, \mathbf{e} - sampling error vector.

It is usually assumed that \mathbf{v} and \mathbf{e} are independently distributed with mean 0 and covariance matrices \mathbf{G} and \mathbf{R} respectively.

EBLUP estimator for the small area model given by (1) has the following form :

$$\boldsymbol{\theta}_{EBLUP} = \mathbf{X}\boldsymbol{\beta} + \mathbf{M}\mathbf{G}\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) \quad (2)$$

where:

$$\boldsymbol{\beta} = (\mathbf{X}^T \mathbf{V}^{-1} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{V}^{-1} \mathbf{y}. \quad (3)$$

\mathbf{M} is the identity matrix, and \mathbf{G} is the matrix with non-zero diagonal and its values are equal to σ_v^2 which is the model variance. It is usually computed using special iterative procedure that applies the Fisher algorithm.

Fisher “scoring” algorithm is a form of Newton- Raphson method used to solve maximum likelihood equations numerically. In the case of basic area level models, the variance component σ_v^2 can be obtained by means of the following iterative formula:

$$\sigma_v^{2(a+1)} = \sigma_v^{2(a)} + [I(\sigma_v^{2(a)})]^{-1} s(\tilde{\boldsymbol{\beta}}^{(a)}, \sigma_v^{2(a)}) \quad (4)$$

where: $I(\sigma_v^2) = \frac{1}{2} \sum_{i=1}^m \frac{1}{(\sigma_v^2 + \psi_i)^2}$ denotes the Fisher information,

$s(\tilde{\boldsymbol{\beta}}, \sigma_v^2) = \frac{1}{2} \sum_{i=1}^m \frac{1}{\sigma_v^2 + \psi_i} + \frac{1}{2} \sum_{i=1}^m \frac{(\hat{\theta}_i - \mathbf{z}_i^T \tilde{\boldsymbol{\beta}})^2}{(\sigma_v^2 + \psi_i)^2}$ is a score function, ψ_i - sampling variance for the region “ i ”.

PRECISION OF THE EBLUP ESTIMATION

Mean square error estimate (MSE) of EBLUP can be obtained from the following formula:

$$MSE(\theta_{EBLUP}) = g_1(\hat{\delta}) - b_\delta^T(\hat{\delta}) \nabla g_1(\hat{\delta}) + g_2(\hat{\delta}) + 2g_3(\hat{\delta}) \quad (5)$$

where δ is a variance dependent parameter. Using this formula we usually assume that the mean square error of EBLUP is the sum of three main elements g_1 , g_2 and g_3 which are described by the following equations:

$$g_1(\hat{\delta}) = \text{diag}(\mathbf{G} - \mathbf{G}\mathbf{V}^{-1}\mathbf{G}) \quad (6)$$

$$g_{2i}(\hat{\delta}) = (\mathbf{X}_i - \mathbf{m}_i^T \mathbf{G}\mathbf{V}^{-1}\mathbf{X})(\mathbf{X}\mathbf{V}^{-1}\mathbf{X})^{-1}(\mathbf{X}_i - \mathbf{m}_i^T \mathbf{G}\mathbf{V}^{-1}\mathbf{X}) \quad (7)$$

$$g_{3i}(\hat{\delta}) = (\mathbf{m}_i^T (\mathbf{V}^{-1} - \mathbf{G}(\mathbf{V}^{-1}\mathbf{V}^{-1}))\mathbf{V}(\mathbf{m}_i^T (\mathbf{V}^{-1} - \mathbf{G}(\mathbf{V}^{-1}\mathbf{V}^{-1}))^T)\mathbf{I} \quad (8)$$

where \mathbf{m}_i is a vector with zeroes for all elements with exception for the element having an index i while \mathbf{I} is the inversed Fisher information matrix.

The additional bias factor $b_\delta^T(\hat{\delta}) \nabla g_1(\hat{\delta})$ emerges only in case of the maximum likelihood (ML) estimation and can be calculated using some special formulas (see: Rao, 2003, page 129). This drawback of the ML method, connected with not taking into account the loss in the degrees of freedom coming from the estimation of regression parameters $\boldsymbol{\beta}$ can be omitted using REML.

Restricted (or residual) maximum likelihood (REML) is the method for fitting linear mixed models that produces unbiased estimates for variance and covariance parameters. In the case of REML the computation algorithm is almost identical except that in the procedure that evaluates σ_v^2 a special \mathbf{P} matrix instead of \mathbf{V}^{-1} matrix is used as follows:

$$\mathbf{P} = \mathbf{V}^{-1} - (\mathbf{V}^{-1}\mathbf{X}(\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{V}^{-1}) \quad (9)$$

The MSE for REML variant is computed using a simpler formula. Here the bias factor is neglected:

$$MSE(\theta_{EBLUP-REML}) = g_1(\hat{\delta}) + g_2(\hat{\delta}) + 2g_3(\hat{\delta}) \quad (10)$$

APPLICATION ON THE BASIS OF THE POLISH HOUSEHOLD BUDGET SURVEY

Sampling frame for Polish HBS is a list of census enumeration areas (CEAs) prepared by means of the census data. It is updated annually according to the increase of dwellings due to the completion of new buildings, and the decrease of dwellings due to the demolition of buildings and changes in administration division of the country. In some cases the Census Enumeration Areas are linked together, to achieve the minimum primary sampling units' size, that is set as 250 dwellings for urban areas and 150 for rural areas. The two-stage sampling plan is used and primary sampling units (PSU) are selected with probabilities proportional to their size using Hartley-Rao method. In each PSU 24 dwellings are selected (with 2 dwellings for each survey month and this dwellings are interviewed also in the same month next year). Additionally, in each PSU 150 dwellings are selected as a reserve sample, which is used in the case of non-response. The stratification is done by 16 voivodships, and in each voivodship, according to the class of size of localities and urban-rural criteria. Large cities constitute separate data. The number of strata in each voivodship ranges from 3 to 12. Altogether there are 96 strata.

In order to reduce the effect of unequal selection probabilities for Primary Sampling Units (PSU) the survey results are specially weighted. The starting point of each dwelling weight is the inverse proportion to its inclusion probability. The non-response coefficients in Polish HBS are relatively high and considerably affect the socioeconomic structure of households in the sample. To reduce the non-response bias the special weights from internal, as well as from external sources (e.g. Labour Force Survey), are applied [Kordos et al., 2002]. Balanced half-samples technique (BRR) is used to determine the standard error.

In the paper the results for some models using data from Polish HBS are presented. These models were constructed using different sources of data that come from Polish Public Statistics and also from administrative data. Firstly we present the region (voivodship) model for per capita expenditures. The exploratory variable

is per capita value of GDP for regions. The ordinary linear regression model $y = a_0 + a_1x$ can be summarized as follows:

$$\hat{y} = 411,70 + 2,35x$$

$$t_{\alpha} \quad (12,01) \quad (6,46)$$

The value of determination coefficient (R^2) is 0,748, the value of corrected R^2 is equal to 0,730, the F statistics is 41,749 and standard estimation error is equal to 28,994.

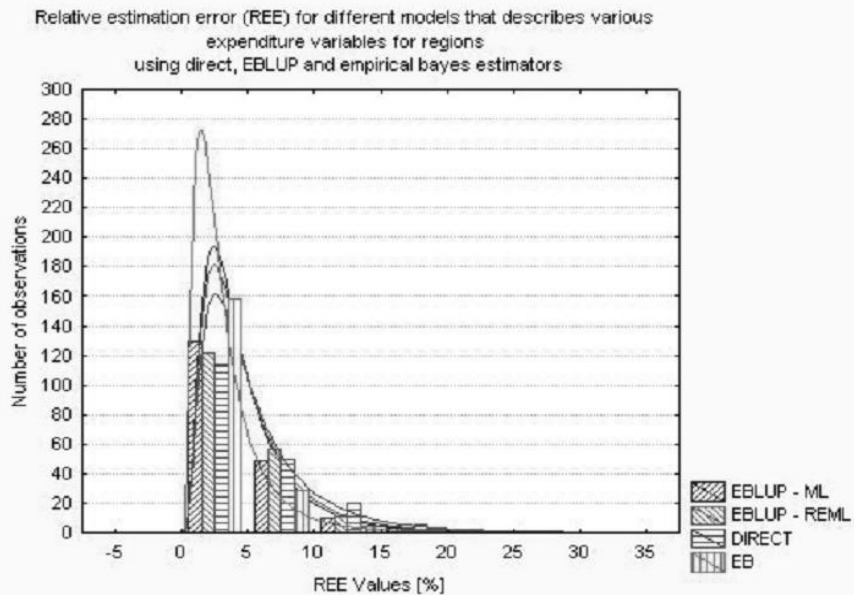
As it is showed in table 1, EBLUP estimator gives systematically lower MSE values and they are slightly greater for the REML variant. The reduction of relative estimation error (REE) for the EBLUP estimates is presented in figure 2. The figure presents the results for various models with different characteristics (including models for expenditures of consumer goods and services- food and non-alcoholic beverages, clothing and footwear, housing water, electricity, household equipment and routine maintenance of the house, health, transport, communication, recreation and culture etc).

Table 1. Per capita expenditures for regions- direct and EBLUP estimates

Region	Direct estimates	Variance for direct estimates	EBLUP - ML estimates	MSE EBLUP - ML	EBLUP-REML - estimates	MSE EBLUP-REML
Dolnośląskie	647,79	518,93	650,41	299,38	650,14	328,68
Kujawsko-Pomor.	590,69	401,36	605,40	258,07	603,75	279,49
Lubelskie	593,77	107,58	590,99	98,19	591,41	100,17
Lubuskie	630,53	724,69	621,54	345,22	622,25	386,97
Łódzkie	654,95	164,02	648,37	137,85	649,32	142,92
Małopolskie	615,67	79,74	615,46	73,92	615,48	75,14
Mazowieckie	792,14	203,72	787,08	200,98	787,81	203,33
Opolskie	655,04	314,92	632,88	224,96	635,53	239,90
Podkarpackie	566,18	170,67	569,54	145,62	569,03	150,60
Podlaskie	592,42	44,06	592,29	42,56	592,31	42,88
Pomorskie	661,88	86,81	659,07	79,86	659,52	81,31
Śląskie	655,84	180,23	659,15	150,91	658,68	156,59
Świętokrzyskie	569,74	598,05	584,30	323,31	582,91	357,60
Warmińsko-Mazur.	570,39	295,84	580,77	216,06	579,45	229,61
Wielkopolskie	597,73	164,07	613,17	139,03	610,92	143,93
Zachodniopomorskie	654,69	415,63	645,63	263,35	646,59	285,77

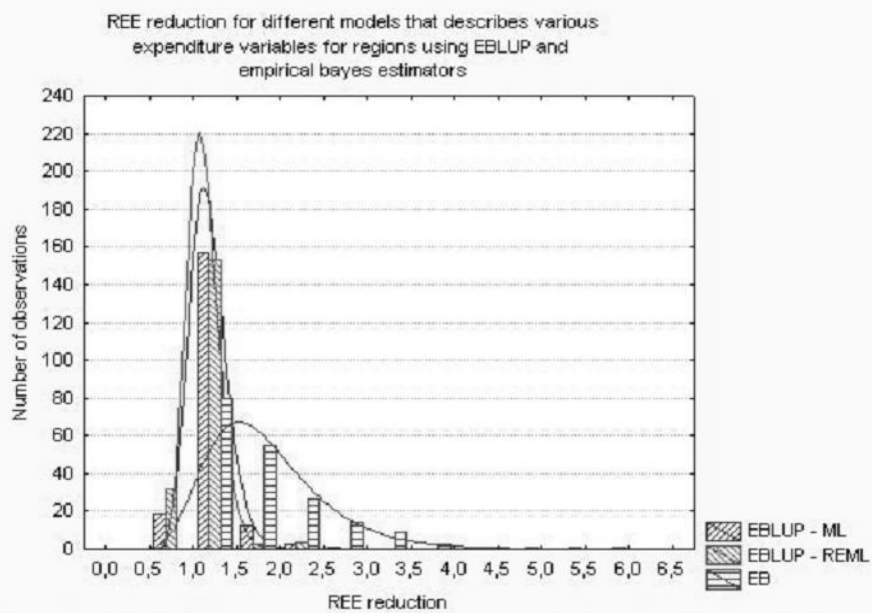
Source: own calculations using HBS data, WesVar software and SAE package for R-project

Figure 1. Relative estimation error values for direct, EBLUP and EB estimators



Source: own calculations using HBS data, WesVar and Statistica software and SAE package for R

Figure 2. Relative estimation error reduction for EBLUP and EB estimators



Source: own calculations using HBS data, WesVar and Statistica software and SAE package for R

In order to assess the variability of variance elements g_1 , g_2 and g_3 new measures of such variability are introduced, that allow to evaluate the share of each element in the overall MSE value. For g_1 element it can be expressed as follows:

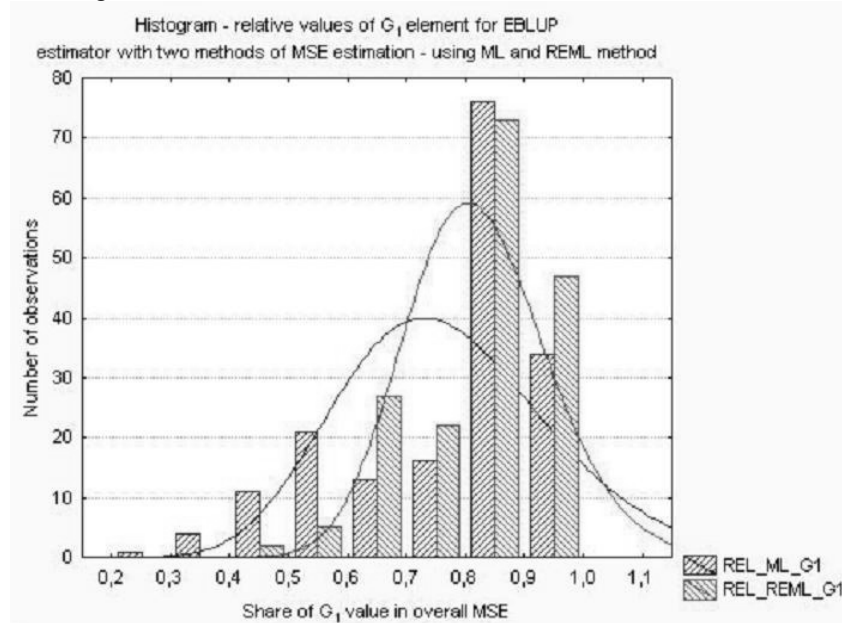
$$\text{Rel}g_1 = g_1 / \text{MSE}(\theta_{EBLUP}) \quad (11)$$

Analogical measures can be obtained for g_2 and g_3

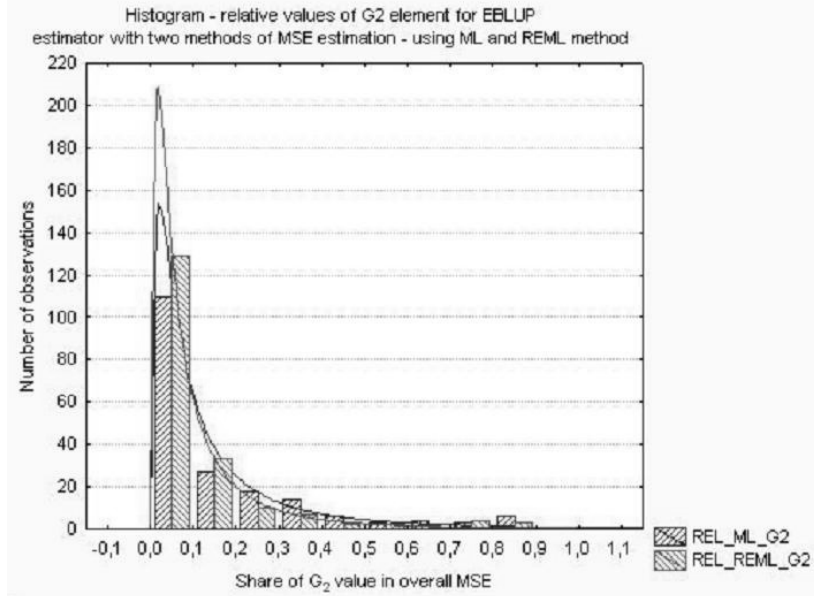
According to the results presented in this paper, using EBLUP technique improves the quality of the estimation. Observed reduction of MSE proves this. However, because of the method of MSE estimation, this results can be rather treated as good approximation.

Values of MSE obtained using EBLUP technique are more adequate than the values obtained using the naive EB procedure. This result is consistent with the theory. More details concerning the procedure for empirical Bayes (EB) estimation can be found in [Bracha et al, 2004]. It can be supposed that in the future some kind of compromise will be reached to become certain about the reliability of the methods discussed here. The comparison of different implementation methods of estimation may also be useful in this field.

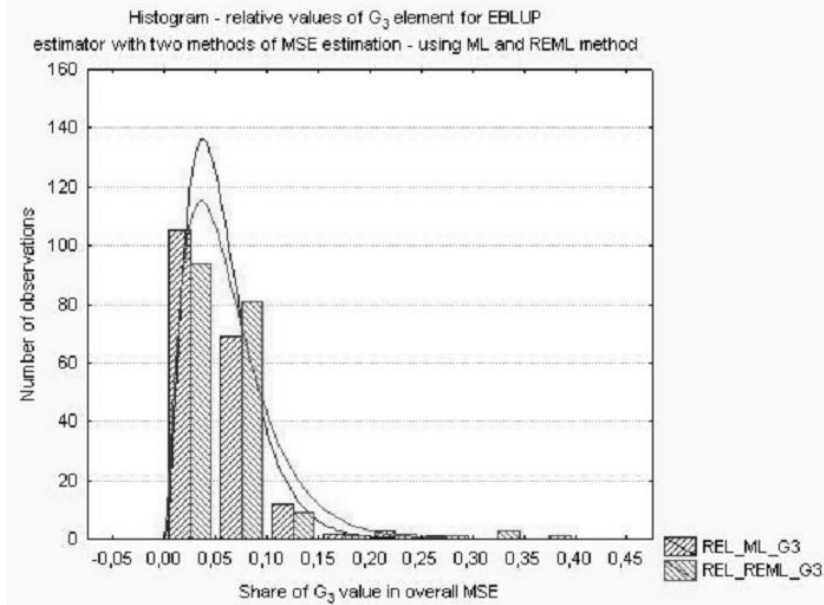
Figure 3. Histogram of relative G_1 element values for two variants of EBLUP estimation



Source: own calculations using HBS data, WesVar and Statistica software and SAE package for R

Figure 4. Histogram of relative G_2 element values for two variants of EBLUP estimation

Source: own calculations using HBS data, WesVar and Statistica software and SAE package for R

Figure 5. Histogram of relative G_3 element values for two variants of EBLUP estimation

Source: own calculations using HBS data, WesVar and Statistica software and SAE package for R

5. CONCLUSIONS

EBLUP and other indirect estimators are likely to be more efficient than the corresponding direct ones as a result of “borrowing strength” from other domains in time and in space.

The EBLUP estimation procedure, based on a general linear mixed model, has an additional advantage of taking into account the between-area variation beyond that explained by the auxiliary variables included in a regression model.

Analyzing the empirical results presented in the paper one can easily notice that the EBLUP technique improves significantly the estimation quality. The mean square error reduction, as compared to the direct estimates, proves this. However, because of the method of MSE estimation, these results can rather be treated as good approximations.

The problem of effective MSE estimation is of great importance and should be analyzed in detail. In the paper the empirical distributions of mean square error components for EBLUP have been presented and discussed.

The obtained dependencies reveal that g_1 component has the main share in the overall MSE value. In some cases this share is greater than 90%, what may indicate that the implementation of MIXED procedure in SAS system, where only the first two elements are taken into account - see [Rao 2003, Part 6.2.7, p 105] - may give also proper results.

In most cases the share of g_2 component is not greater than a few percent but it is worth mentioning that for some models its contribution to the overall MSE value is much higher- even 90%. This result suggests that g_2 can also play an important role in MSE estimation.

The g_3 element's share is usually smaller than 5% but the most frequent values are smaller than 1% what easily explains its supplementary role. However, also because of this – it shouldn't be neglected in MSE estimation.

REFERENCES

- Bracha, Cz., Lednicki, B., Wieczorkowski, R. (2004) Wykorzystanie złożonych metod estymacji do dezagregacji danych z Badania Aktywności Ekonomicznej Ludności w 2003 roku. „Z Prac Zakładu Badań Statystyczno-Ekonomicznych” z. 299
- The EURAREA Consortium (2004), Project Reference, Volume I and III
- Fay, R.E., Herriot, R.A. (1979), Estimation of Income from Small Places: Application of James Stein Procedure to Census Data. *Journal of the American Statistical Association*, 74, pp. 269-277
- Kordos, J. (2005), Household surveys in transition countries. Chap. XXV in *Household Sample Surveys in Developing and Transition Countries*, United Nations, New York

- Kordos, J., Lednicki, B., Żyra, M. (2002), The household sample surveys in Poland. *Statistics in Transition*, vol. 5, No. 4, pp. 555-589
- Pratesi, M., Salvati, N. (2008), Small area estimation: the EBLUP estimator based on spatially correlated random area effects, *Statistical Methods and Applications*, 17, No 1, pp. 113-141
- Rao, J.N.K. (2003), *Small Area Estimation*, Wiley, London
- Software for ecological inference (2007) <http://www.bias-project.org.uk/software/>
- UNDP Polska (2006), Report "Social Exclusion and Integration in Poland: An Indicators-based Approach", Warsaw

Zastosowanie estymatorów eblup do analizy małych obszarów na podstawie badania budżetów gospodarstw domowych

Streszczenie: W artykule przedstawiono wyniki estymacji dla małych obszarów w Polsce otrzymane z wykorzystaniem estymacji bezpośredniej oraz metody EBLUP (*empirical best linear unbiased prediction*). Analizy prowadzone były na podstawie próby pochodzącej z Badania Budżetów Gospodarstw Domowych oraz informacji dodatkowych pochodzących ze źródeł administracyjnych. W celu przeprowadzenia estymacji EBLUP oszacowane zostały modele wydatków gospodarstw domowych, uwzględniające również zmienność pomiędzy obszarami. Oszacowania wariancji dla estymatorów bezpośrednich otrzymano za pomocą metody replikacyjnej BRR (*balanced repeated replication*). Estymatory EBLUP wraz z ich błędami średniokwadratowymi oszacowane zostały za pomocą techniki komponentów wariancyjnych. W celu uzyskania ocen błędów średniokwadratowych dla EBLUP zastosowano metodę ML (największej wiarygodności) oraz metodę REML (największej wiarygodności z restrykcjami). Obliczenia prowadzone były w pakiecie SAE działającym w środowisku R.

Keywords: statystyka małych obszarów, estymacja EBLUP, estymacja wariancji, badanie budżetów gospodarstw domowych

PREDICTIVE MARKETING MIX MODELLING IN FMCG PACKAGED FOOD CATEGORY INCLUDING PRICE AND ADVERTISING IMPACT

Tomasz Kolanowski

Wydział Badań i Analiz, UMWW

e-mail: T.kolanowski@umww.com

Abstract: Case study describing estimation of Marketing Mix econometric sales modelling on packaged food fast moving consumer goods product. Model bases on two step modelling approach, using ordinary least squares method. Adstock data transformation has been used to evaluate advertising impact with time distributed lag. Paper presents non-linear relations between market distribution, consumer price levels and advertising as well as 52 weeks forward sales forecast accuracy evaluation.

Keywords: Marketing Mix Modelling, Advertising, Sales Forecast, Fast Moving Consumer Goods (FMCG).

INTRODUCTION AND THEORETICAL APPROACH

The Following paper illustrates the estimation and evaluation of Marketing Mix econometric sales modelling on FMCG product. This technique is recommended in brand managing (Rószkowska 2002) and particularly in advertising management (Beliczyński 2007) and become very popular in recent years. This kind of modelling evaluates impact of marketing activities (independent variables) on sales of the product (dependent variable). As in every modelling, the key dependent factors have to be included into model. There are many theoretical frameworks identifying the most impactful marketing factors, but in this paper the most popular will be used. The 4P's theory was first formulated in 1960's (McCarthy & Shapiro 1975) and since this time very has been very well described and evaluated (Kotler 1999). This theory identifies key factors as Product, Price, Place and Promotion (4P). Product represents all physical features like size, packaging or quality. Place describes all issues related to point of sales: size of stores, number of stores or supplies. Price includes regular product price but also all discounts and price

promotions. Promotion represents all kind of advertising, public relations and merchandising.

One of the most difficult marketing objectives for marketing mix modelling is evaluating the effect of advertising. This part of promotion is usually the highest position in brand marketing budget (Zyman 1999) therefore should be included in proper modelling.

The most common method for advertising modelling is called Adstock which is similar to time distributed lag. The transformation represents fact that effect of advertising impact can be observed long after end of advertising campaign.

$$A_t = T_t + \lambda A_{t-1}, \quad t = 1, \dots, n \quad (1)$$

Where A_t is the Adstock at time t , T_t is the value of the advertising variable at time t and λ is the ‘retention’ or lag weight parameter. Inclusion of the A_{t-1} term, imparts an infinite lag structure to this model, with the effect of the first Adstock term, approaching zero, as t tends to infinity.

Adstock parameter λ can be interpreted in the models as percentage of effectively remembered ad contact from previous week plus contacts from current one (Broadbent, 1979). In literature it is frequently called “Ad Retention”, there are also two other popular terms which are describing the same phenomenon: “Decay Factor” simply equals $1-\lambda$ and “Advertising Half-Life” which represents time needed to halve advertising effect. All three coefficients can be simply recalculated to each other.

There are no strict norms for ad retention and there are two popular theoretical approaches. In the first, the predefined lambda is determined for the model, (usually 80% or 90%) probably based on classical Emshoff and Mercer research (1970). The other recommended method is to set this parameter during modelling (Hanssens 2003). Some newer academic analyses suggest the most typical half-life range to be around 7-12 weeks (Leone 1995), approximately equivalent of $\lambda = 90\%-95\%$.

Adstock is only a mathematical transformation; its real-world nature is still a matter of discussion. Generally there are two opinions; one is that Adstock represents forgetting ad contacts with time, the other that it is caused by behavioural habit (Jones, 2002) created by advertising.

Advertising can be divided into two areas: creation and media planning (Czarnecki & Korsak 2001). Media planning focus on finding best campaign timing, most effective media and the optimal investment levels (Sisstors & Bumba, 1996). The model was designed to help the planning process, so it is focused on constant in time advertising performance, rather than on a particular creative’s effectiveness.

CASE DESCRIPTION

The analysed brand is owned by international corporation mainly focused on processed food market. The product belongs to broad salty snacks category. The brand was introduced to the UK market two years before analysis and still is in its early life stage.

THE DATA

Because the analysis was needed urgently the following data was chosen:

Sales in Quantity – delivered by AC Nielsen from their store panel. The data represents sales in Units. Units should be interpreted as number of packages or kg of products. This distinction doesn't change interpretation, but helps shield clients' sensitive data. Sales Data provided on a weekly basis and represents sales in multiple stores.

Sales In Value - delivered by AC Nielsen from their store panel. Defined as the total amount of money (GBP) spent on the product at the cash points. Data provided on a weekly basis, aggregated on a national level, and represents sales in multiple stores.

ACV Percentage Distribution – All commodity value weighted retail distribution from AC Nielsen. The measure estimates the percentage of stores where the product can be purchased. Each store is weighted by its sales potential to reflect the real power of distribution. This is illustrated by an example; suppose that town X has 10 stores. However, one of them is a megastore and represents 50% of all trade in the town. If our brand was present only in this store, its ACV Distribution will be 50%. (Curry 1993). Data provided on a weekly basis, aggregated on national level and represents sales in multiple stores.

Ad Spend - delivered by Nielsen Media Research through AdDynamix. Data provided from an independent monitoring system. Methodology based on monitoring of actually broadcasted/issued advertisements across different media. Each advertisement is attributed to a specific brand, and its cost is estimated. The costs are provided based on the rate-card, which means an official published cost of the placement. The real costs are significantly lower and depend on individually negotiated discounts for clients. For a mid-size client in UK, we estimate a discount of 35%-45% off the rate-card price. The discount is constant year to year. Data provided on a weekly basis and aggregated on a national level.

No other data had been used in modelling.

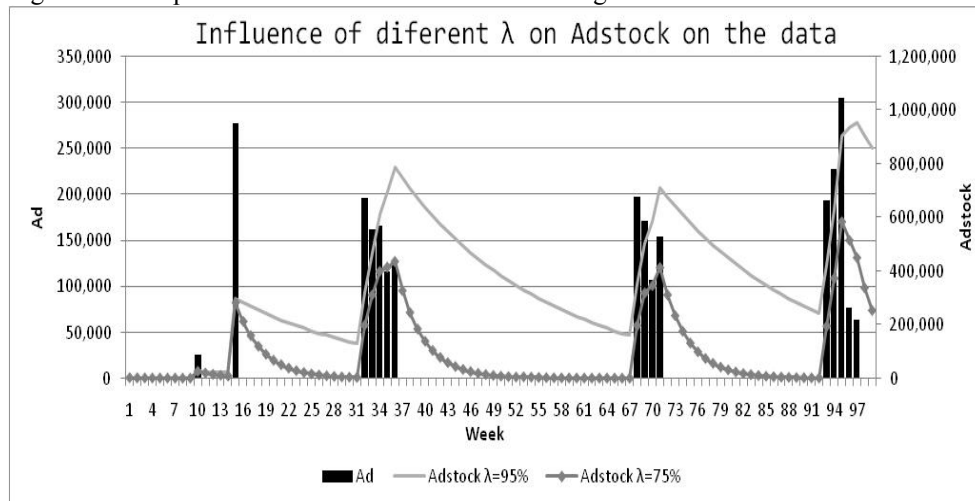
To find the representation of price, the sales (in value), were divided by sales (in volume) in each week. The result was an average purchase price for each week. This number helped to evaluate also effect of in-store promotions (like 2-for-1 offers), when display price is on the same level, but quantity of product increases.

This recalculated indicator also addressed multiple display prices in different stores within the same week. The indicators worked on total aggregated numbers

so for example, in a certain week 1000 units were sold for total amount of 800 GBP, the calculated average price was 80p. This indicator will be called “price” later on.

The advertising date were transformed using ad stock method. As previously mentioned, the most typical λ is in range 0.8 – 0.95. However, it is noted that λ varies strongly on different products and advertisements. In some product categories, it may be lower than 0.5, especially where direct response is a modelled variable. Therefore, in this case the best fitting λ was evaluated from whole theoretical range (0,1). There were five advertising bursts (campaigns) with different pressure during modelling period. According to different λ , their influence on the model varied significantly. (Fig 1)

Figure 1. Example of Adstock transformation according to λ .



Source: own calculations

MODELLING PROCESS

The first model tried to build simple single equation model, and failed. Error had a tendency to rise with time. It was for this reason, building a two-step model was considered.

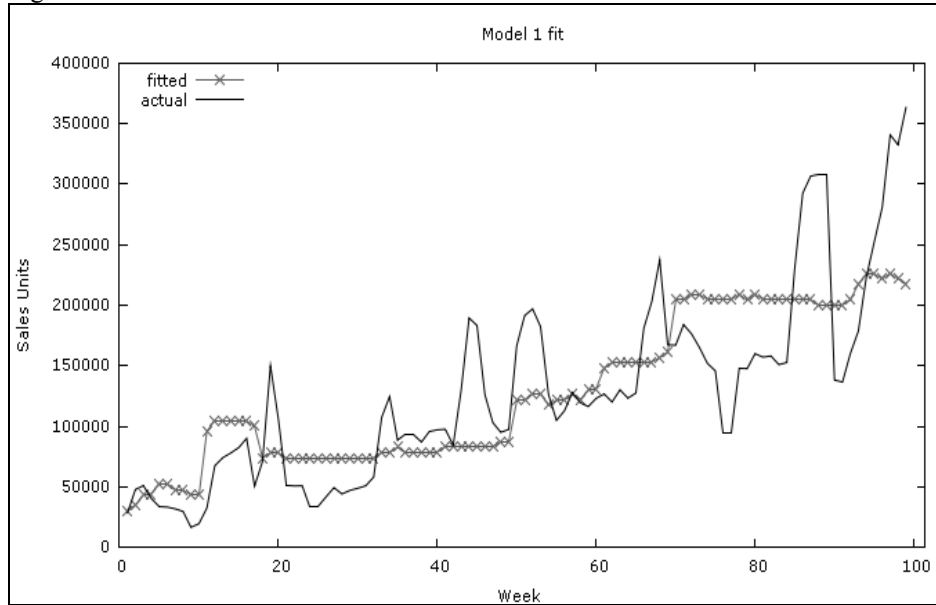
Sales were almost constantly increasing because brand was relatively new and still building its market distribution. ACV distribution was the highest correlated factor. The first model was built to evaluate regression only from ACV on sales.

$$S_d = \beta_1 * ACV \quad (2)$$

where: S_d = Estimated sales in units modelled by distribution, S_a = Actual Sales units

The ordinal least squares method was used. During modelling, the constant was removed because it had a low influence, and was statically not significant. It has also clear economic interpretation; when distribution (ACV) is equal zero, the product is not available in the market, therefore sales is null.

Figure 2: Model 1 fit



Source: own calculations

Residuals from the first stage model were transformed to relative fractional values with this simple formula.

$$S_r = \frac{S_a}{S_d} \tag{3}$$

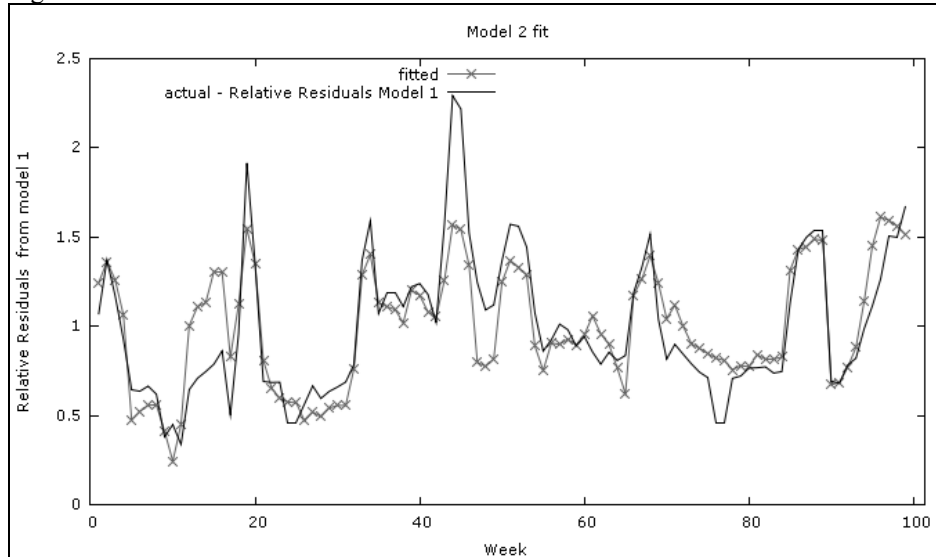
where: S_a = Actual Sales units, S_r = Relative residuals from model 1

The transformed residuals became a new time series for modelling in stage two. Again, the ordinal least squares method was used. The independent variables, Adstock and price, were used. Adstock λ was tested to find the best fit model – process was made by overall model R^2 , maximising by changing λ . (Fig. 2) Finally, the $\lambda=96.1\%$ was found as the best fit (equivalent of half-life 17.42 weeks). This level was also the most statistically significant (fig. 3).

$$S_r = \beta_2 \cdot AdSt + \beta_3 \cdot Price + \beta_4 \tag{4}$$

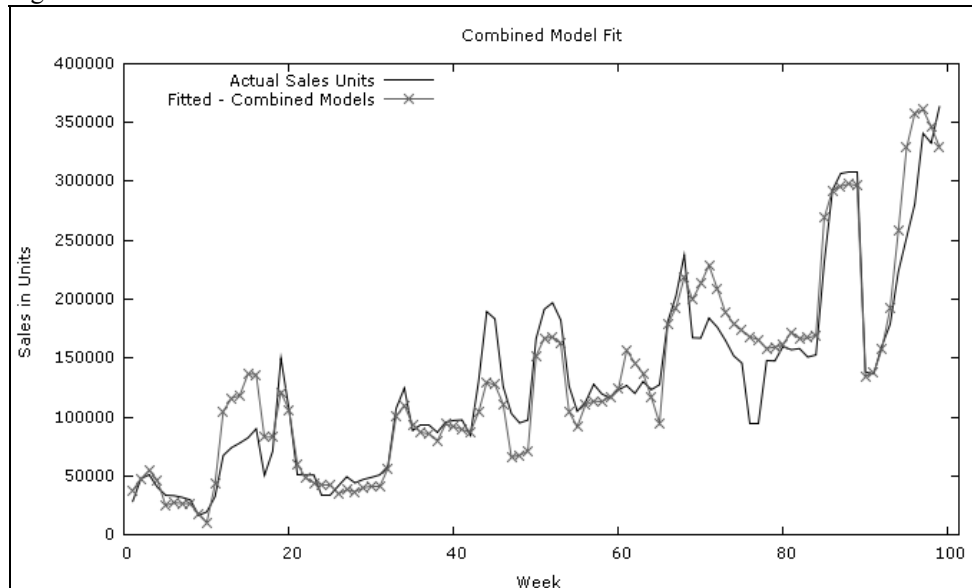
Where S_r = Relative residuals from model 1, S_a = Actual Sales units (measured), $AdSt$ = Adstock, $Price$ = average product price

Figure 3. Model 2 fit



Source: own calculations

Figure 4. Combined model fit



Source: own calculations

The models were combined to make calculation easier. The formula allows calculation of estimated sales in a single equation. The overall effect of applying two step models estimated sales on satisfying level (Fig 4). There were some prob-

lems with accuracy in three periods (around week 15, 50 and 75), but the general shape fitted to all local maximums and minimums.

$$S_Q = \beta_1 \cdot ACV(\beta_2 \cdot AdSt + \beta_3 \cdot Price + \beta_4) \tag{5}$$

where S_Q =Estimated sales in quantity (units)

It is worth noting that if number of units sold is estimated, and price for average unit is known, the next step would be to multiply equation by sides by price to gain value of sales. This transformation is acceptable (Hanssens 2003) in market mix modelling.

$$S_V = \beta_1 \cdot ACV(\beta_2 \cdot AdSt + \beta_3 \cdot Price + \beta_4) \cdot Price \tag{6}$$

where S_V = Estimated sales in value ($S_Q \cdot Price$)

To summarise the modelling process the two ordeal least squares regression were used. The first model estimates sales in units by β_1 parameter its R^2 equals 0.89. The second model evaluates effect of Adstocked ad activity β_2 and price β_3 on residuals from the first one. Its adjusted R^2 equals 0.72 and unadjusted 0.71. All coefficients including constant are highly statistically significant. (Tab 1)

Table 1. Equations parameters

Variable	Distribution	Adstock	Price	Constant
Coefficients	β_1	β_2	β_3	β_4
Value	4352.015595	4.60806E-07	-2.458696295	3.005917229
S(β)	155.1746507	7.98862E-08	0.179689217	0.169324788
stat t	28.04591844	5.768273588	13.68304863	17.75237558
p-value 2-sig	1.93928E-60	4.66054E-08	6.88719E-28	3.47167E-38
	***	***	***	***
p-value1-sig	9.6964E-61	2.33027E-08	3.44359E-28	1.73583E-38
	***	***	***	***

Source: own calculations)

MARKETING INTERPRETATION AND DISCUSSION

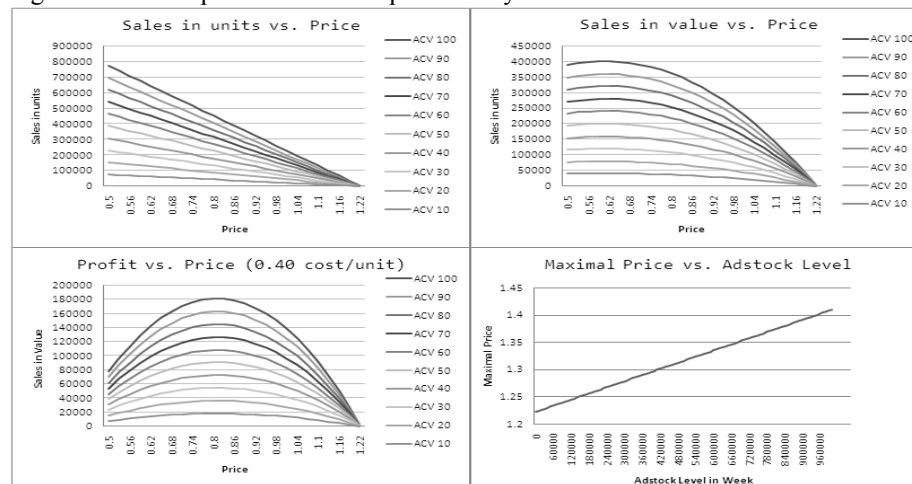
The model has quite an intuitive interpretation. For each percent point of weighted market distribution (ACV), the theoretical constant unit sales were generated. Then this theoretical level was modified by advertising and price factor. The advertising and price worked proportionally to build distribution. Simply speaking, the more people that are able to purchase the product, the more impactful price and advertising are. Price had a negative effect (higher price lower sales) and Adstocked advertising had a positive effect (more advertising, bigger sales). Advertising seemed to work very long-term, even after 17 weeks its effect was only halved. However, it is coherent with Leone (1995) works.

The Price, ACV and Adstock were highly connected and in the equation, and interact with themselves. Charts were quite difficult to read. The most important factor was ACV distribution, therefore wherever it was possible, different levels of ACV were charted.

The unit sales were related to price in a linear way. Lower price generated higher sales. When analysed with zero advertising the level 1.22 was the highest price that generated positive sales. However, for sales expressed in value, effects were non-linear. Maximal sales were generated when the price was set to 0.61 (no advertising).

Another interesting question is ‘What is the most profitable price level’? We did not have data regarding company cost levels per unit (manufacturing, packaging, transport etc) but it was assumed at 0.40. This was a little lower than the lowest observed price in analysed period. This assumption was done to show how the model can be used by clients. In this case, the optimal price level was 0.81 (without advertising) (Fig 5).

Figure 5. Price impact on sales and profitability



Source: own calculations

The effect of increasing price elasticity (because of advertising), was also observed. With null ad pressure the maximal price was 1.22, but with strong Adstock presence it rose above 1.4. The same shift occurred on optimal price for profitability, it rose above 0.91 from the effect of strong ad pressure. The effect of increasing customer's price elasticity as a result of advertising was identified as the most profitable on UK market (Binet & Field 2007) in recent IPA case study base.

Advertising works in long time, its effect despairs with ratio 96.1% per week. Product familiarity had a key effect on product sales and advertisements were well remembered. Each invested amount of money paid itself back 88% of its

Adstock in the first year, and the rest was present in the following years. The amount of generated sales was proportional to market distribution level ACV. There was also the effect of shifting consumers' price elasticity because of advertising. People seem to accept higher prices when the effect of advertising is present.

FORECASTING AND MODEL EVALUATION

Unfortunately because of external factors the model had never been used in planning. But when the following year's data became available, it was great opportunity to check model accuracy on real values. Again all data was taken from the same sources but further modelling was not made. The proposed approach belongs to ex-post methods and MAPE parameter has been used to evaluate model performance (Witkowska 2006).

This exercise checked how accurate the findings from the first two years modelling were, for product sales dynamics in the third year. In the 52 weeks, new promotions and price levels had been created, and also a new ad campaign was broadcast. There was also a significant increase in market penetration (ACV) from 50% on beginning to 62% by the end of forecast period. This meant that all significant factors were not stable (which makes a good testing period).

The model seemed to be useful in long-term predictions. The 52 forecasted weeks are the usual annual marketing planning period. The model predicted the whole period sales with 3.7% error and its mean absolute percentage weekly error (MAPE) was equal 12.4%. The model was also stable in time and had no tendency to accumulate error in time (fig. 6). Forecasts from the end period were not worse (and sometimes were better) than those from early data closer to the modelling period. (Table 2)

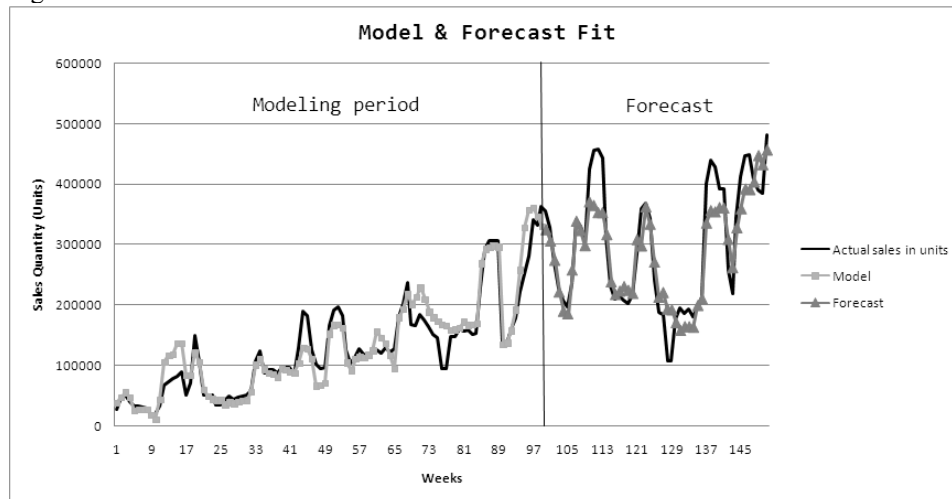
Table 2. Forecast Evaluation

13 week quarters	Q1	Q2	Q3	Q4
Absolute % Error - Total Period	8.1%	1.4%	1.3%	3.1%
MAPE weekly	8.2%	7.2%	22.2%	11.5%

Source: own calculations

It should be noted that all local minimums and maximums were properly indicated. Only the strength of sales response were underestimated in the lowest price (0.7) periods, around week 111 and 138. The prices at this level were almost not present in modelling period. This may indicate a nonlinear price response curve application for the following year's modelling.

Figure 6: Model evaluation



Source: own calculations

CONCLUSIONS

The case showed that market mix modelling can be successfully applied to brands, even when they are in early life stages or a dynamic growth period. The publicly available market monitoring data like AC NIELSEN can be a base for modelling and in situations where a company does not gather its own marketing data, this is sufficient enough to understand market dynamics. The most important conclusion is that properly built marketing mix models are able to forecast brand sales performance. This forecast can be accurate enough to set advertising strategies and marketing activities over a one-year period. This time scale allows brand managers to build their annual plans more precisely and improve their return on investment.

BIBLIOGRAPHY:

- Beliczyński, Jan. (2007) *Metody planowania i kontroli w zarządzaniu reklamą*, Wydawnictwo Akademii Ekonomicznej w Krakowie,, Wydawnictwo Akademii Ekonomicznej w Krakowie.
- Binet, L and Field, P (2007) *Marketing in the Era of Accountability: Identifying the Marketing Practices and Metrics That Truly Increase Profitability*, World Advertising Research Center 2007.
- Broadbent, S. (1979) *One Way TV Advertisements Work*, Journal of the Market Research Society Vol. 23 no.3.

- Curry, David J. (1993) *New Marketing Research Systems: How to Use Strategic Database Information For Better Marketing Decisions*, Published by Wiley, John & Sons, Incorporated.
- Czarnecki, A, Korsak, R (2001) *Planowanie mediów w kampaniach reklamowych.*, Published by PWE.
- Emshoff, J. R. and Mercer A. (1970) *Aggregated Models of Consumer Purchases*, *Journal of the Royal Statistical Society, Series A*, 133:1, 14-32.
- Hanssens, D. M. Parsons, L. J. Schultz R. L. (2003) *Market Response Models: Econometric and Time Series Analysis*, Published by Springer.
- Jones, John Philip (2001) *The Ultimate Secrets of Advertising*, Edition: 3, Published by Sage Publications.
- Kotler (1999) *Marketing. Analiza, planowanie, wdrażanie i kontrola*. Published by Warszawa.
- Leone, R.P. (1995) *Generalizing what is known about temporal aggregation and advertising carry-over*, *Marketing Science*, 14, G141-G150.
- McCarthy, E. J and Shapiro S. J. (1975) *Basic Marketing*, published by Irwin-Dorsey.
- Rószkiewicz, Małgorzata (2002): *Narzędzia statystyczne w analizach marketingowych*, published by C.H.Beck.
- Sissors, J. Z. Bumba, L. (1996) *Advertising Media Planning*, Published by McGraw-Hill.
- Witkowska, Dorota (2006): *Podstawy ekonometrii i teorii prognozowania*, Published by Oficyna Ekonomiczna.
- Zyman, Sergio (1999) *The End of Marketing as We Know it* Published by Harper-Business.

Predyktywne modelowanie marketing mix na przykładzie marki z kategorii dóbr szybkozbywalnych opakowanej żywności z uwzględnieniem wpływu reklamy i ceny

Streszczenie: Artykuł ilustrujący estymację modelu ekonometrycznego sprzedaży na przykładzie produktu z kategorii dóbr szybkozbywalnych opakowanej żywności. Model opiera się na podejściu dwukrokowym z zastosowaniem metody najmniejszych kwadratów. Do modelowania wpływu reklamy z rozłożonym opóźnieniem zastosowano transformację Adstock. Artykuł opisuje nieliniowe zależności pomiędzy dystrybucją produktu na rynku, ceną produktu dla końcowego konsumenta i nakładami reklamowymi jak również ocenę trafności prognozy modelu dokonaną na 52 tygodnie naprzód.

Słowa kluczowe: Modelowanie Marketing Mix, Reklama, Prognoza sprzedaży, Dobra szybkozbywalne (FMCG)

ANALYSIS OF BANKING PERFORMANCE IN POLISH COMMERCIAL BANKS. SYMPTOMS OF THREATS

Witold Luciński
Wydział Ekonomiczny, PŁ
email: wl53@wp.pl

Abstract: Worldwide tendency to decrease interest rates is a major threat to the Polish banks in the long term. In the future, the traditional way of conducting banking activities will have to be modified. The starting point for creating new strategies for Polish banks should be to depart from net interest income as the core model of generating banking income and embrace the prompt development of broadly understood non-interest activities such as private equity. Shift from typical of commercial banking sector loan-deposit to capital market activity may create many benefits for the Polish banking system.

Keywords: net interest margin, commercial banking system, private equity

INTRODUCTION

Since the early 1990s, Europe has seen the extension of the range of products in the banking market. The traditional financial services based on the transformational function of banks, such as providing credit and offering deposits, started to generate smaller profits and began to decline [Koeppen J. 1999 and Schildbach J. 2008]. In turn, the more profitable, individual consultancy services focused on solving the clients' problems became increasingly widespread in the financial market. Moreover, the practice of investing banks' own funds as well as those of their clients in the capital market has spread. [Dębski W. 2005]. This is the effect of the deregulation processes observed in the early 1990s. Monetary integration and the policy of a low, stable interest rate pursued by the European System of Central Banks provided new impulses for a further reduction of margin, which was also triggered by the increasing competition in the banking sector [Baka W. 2005]. Consequently, Europe has seen institutional transformations leading to the gradual disappearance of the division into commercial and investment banking [Solarz J.

1997 and Ostrowska E. 2005]. At the same time, the position of the securities market has grown while the banks' financing of businesses through credit flow has declined. This is a characteristic tendency called the "disintermediation" of banking services. [Korenik D. 2005]. It should be anticipated that this tendency is going to continue as according to a view widespread in the European Union, an efficient capital market is *raison d'état* for every modern country [Korenik D. 2005 and Czechowska D. 2004].

The author shares these views. However, as regards the Polish financial market, this kind of behaviour is not confirmed by business practice, which is shown in the analysis of data presented in the further part of the paper.

The condition of the financial sector is of particular importance in every country, as it is the venue where both economic entities and households keep or invest their savings. Moreover, without the flow of credit the development of modern countries is hardly conceivable. Thus, it is little wonder that in developed countries the value of assets amassed by this sector is many times their GDP [Schildbach J. 2008].

In many countries, Poland included, the functioning of the banking sector for the most part depends on the situation of the banks, which have been the most prominent actors in the financial market to date¹. Due to the huge gap between socialist and free market banking systems, the Polish banking sector has undergone dramatic transformations and has become a modern and profitable sector in terms of organizational solutions. Its stability has now been proven by withstanding the world banking crisis. In contrast to many more developed countries, the Polish financial sector, at least in the first several months of the crisis, did not witness collapsing banks or drastic actions by the government or the Central Bank aimed at preserving its stability.

However, even before the crisis, some symptoms of potential difficulties to the development of this sector had surfaced. Furthermore, these phenomena do not seem to disappear in the future; just on the contrary, they might intensify after the end of the worldwide events started by the subprime crisis in the USA. The author believes that the potential threats are rooted in the income structure of the sector and will increase in the future. The main goal of this study is to illustrate these problems. This article focuses on analysis of banks' net interest income and net interest margin.

ANALYSIS OF BANKING PERFORMANCE

Net interest income is the difference between the value of interest earned by a bank (mostly from its lending activities) and the value of interest paid on the capital it borrows in the financial market. In this article, this difference in absolute terms shall be called "net interest income", while in relative terms, whenever it

¹ Sytuacja finansowa banków w 2006 roku. p. 35.

refers to measures such as the volume of a bank's assets or earning assets, it shall be called a net interest margin (given in the form of a percentage).

Table 1. Breakdown of banking income by type of income in Polish commercial banks in the years 2001-2007, in %

	Specification	2001	2002	2003	2004	2005	2006	2007	Mean
1	Net interest income	51.6	53.3	55.0	55.5	57.6	58.2	57.8	55.6
2	Non-interest income	48.4	46.7	45.0	44.5	42.4	41.8	42.2	44.4
3	Fee and commission income	21.2	22.5	27.5	27.9	23.5	25.9	26.7	25.0
4	Foreign exchange income	18.8	14.9	12.6	10.8	13.2	9.2	8.9	12.6
5	Income from financial operations	6.3	8.0	3.2	4.5	3.0	3.3	4.1	4.6
6	Income from shares and securities	2.0	1.3	1.7	1.3	2.7	3.4	2.6	2.1
7	Income from banking activities	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

Source: own work on the basis of synthetic data from Sytuacja finansowa banków za lata 1998-2006 National Bank of Poland Warsaw 1999-2007 and Raport o sytuacji banków w 2007 roku Polish Financial Supervision Authority Warsaw 2008.

The structure of income of Polish commercial banks is shown in Table 1. Every year in the period under study, net interest income was the prevalent source of generating banking income in Poland. The table presents a gradual decline of its share in banking revenues in the years 1998-2001 (down from 69.7% to 51.6%)². Starting with 2002, this manner of generating banking income has been slightly but systematically growing. Its share in banking income at the end of the period under study was 57.8%.

Total non-interest income constitutes less than half of banking income in Poland, which is shown in Chart 1.

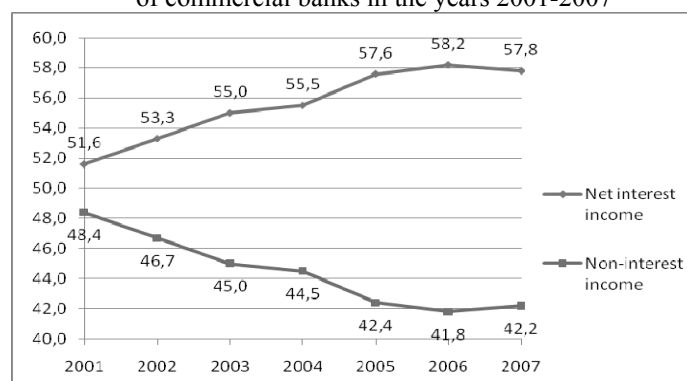
The amount of net interest income is subject to two factors. First of all, it depends on the large spread of interest rates on loans and deposits as well as on a large turnover, *i.e.* even at a low difference between the interest rates sizable net interest income may be generated by a substantial volume of transactions.

In terms of revenues, the core part of net interest income is constituted by receivables from the non-financial sector which form the greatest share of assets. They are the most profitable part of assets in Poland. Double-digit interest rates on loans are now widespread, while interest on interbank deposits barely exceeds 6%.

² Not shown in the table see Sytuacja finansowa banków za lata 1998-2007 National Bank of Poland Warsaw 1999-2008.

In the years 2004-2006, interest rates on one year deposits were approximately 4%. If one takes into consideration the so-called “Belka tax” (a capital gains tax), income from depositing money in banks by both individual and institutional clients has been largely marginalized in Poland.

Chart 1. Percentage share of interest and non-interest income from the banking activity of commercial banks in the years 2001-2007



Source: own work on the basis of synthetic data from Sytuacja finansowa banków za lata 1998-2006 National Bank of Poland Warsaw 1999-2007 and Raport o sytuacji banków w 2007 roku Polish Financial Supervision Authority Warsaw 2008.

Table 2 shows the relationship between net interest income and total banking income in reference countries³.

Table 2. Share of net interest income in total bank income in 2006, in %

Specification	Germany	Sweden	Portugal	Poland	Finland	Czech Republic	Ireland	Slovakia	Great Britain	Hungary	Greece	EU25
Indicator value	48.2	52.5	54.7	58.2	59.2	61.1	62.3	63.1	63.7	67.8	69.1	52.5
Position in the population under study	5	8	10	13	15	16	18	19	20	22	24	

Source: Own work on the basis of Schildbach J. European banks: The silent (r)evolution Deutsche Bank Research Frankfurt am Main Germany 2008, p. 30. Data on the basis of publications by the European Central Bank, International Monetary Fund, Swiss National Bank. The study covers EU25 countries + Switzerland. Maximum – Malta (72.9%), minimum – Switzerland 28.6%. The data calculated by the author are weighted means, where the weights are the relative sizes of the banking sectors

³ Data reported by the National Bank of Poland and the European Central Bank differ by 0.4 percentage points for unknown reasons, which translates into a relative error of 0.7%.

in the selected countries. The abbreviation EU25 means European Union states without Bulgaria and Romania.

The leaders in this classification are countries not shown in the table, *i.e.* Switzerland, Luxembourg, France and Belgium, whose shares are 28.6%, 33.3%, 36.8% and 47.5%, respectively [Schildbach, 2008]. Poland, with its share of slightly below 60% comes 13th in a sample of 26 elements, and 4th in terms of reference countries.

These data throw some light on the situation in Poland and Europe, but a clearer picture emerges upon analyzing the weighted mean of the share of net interest income in total bank income in selected groups of countries. The weights are the volumes of assets amassed in particular banking sectors in the population of countries under study. This is shown in Table 3.

Total population in the study includes EU25 and Switzerland. In this group of countries, the share of net interest income weighted by the size of banking sector assets was 51.3%. The European Union has a slightly worse result, which is due to the fact that Switzerland has an extremely low share of net interest income in generating bank income (28.6%) and the share of Switzerland's banking sector assets is relatively high in the population (over 5% of total assets), which puts this country in sixth place in Europe.

Table 3. Share of net interest income in total bank income weighted by the size of the banking sector in the European Union in 2006, in %

Specification	EU25+ Switzerland	EU25		6 European countries with the largest assets in the banking sector		5 countries with the lowest share of net interest income
		with GB	without GB	with GB	without GB	
Share of net interest income in %	51.3	52.5	48.6	50.6	44.8	42.1

Source: Own work on the basis of Schildbach J. *European banks: The silent (r)evolution* Deutsche Bank Research Frankfurt am Main Germany 2008, p. 30. Data on the basis of publications by the European Central Bank, International Monetary Fund, Swiss National Bank. The study concerns EU25 countries + Switzerland. The data calculated by the author are weighted means, where the weights are the relative sizes of the banking sectors in the selected countries. The abbreviation EU25 means European Union states without Bulgaria and Romania.

The other two indicators in the next two columns cover the twenty five countries of the European Union. The first indicator, marked with the note "with GB" (with Great Britain) covers EU25. In this population, the average share of net interest income in total income of the EU banking sector (EU25) was 52.5%.

However, as Great Britain has the largest banking sector in the European Union (24.8%), its high indicator strongly influences that for the whole population of EU25. Furthermore, Table 3 shows that Great Britain has one of the largest shares of net interest income in total bank income. If the influence of this country on the European indicator is omitted, its value will decrease to 48.6% (the column with the note “without GB”).

Following a similar train of thought, indicators were calculated for top 6 countries in terms of assets involved in the banking sector in the population under study. With Great Britain included, the indicator was 50.6%, while without it the indicator dropped to 44.8%. In the case of reference countries with and without Great Britain, the result was 57.4% and 51.7%, respectively.

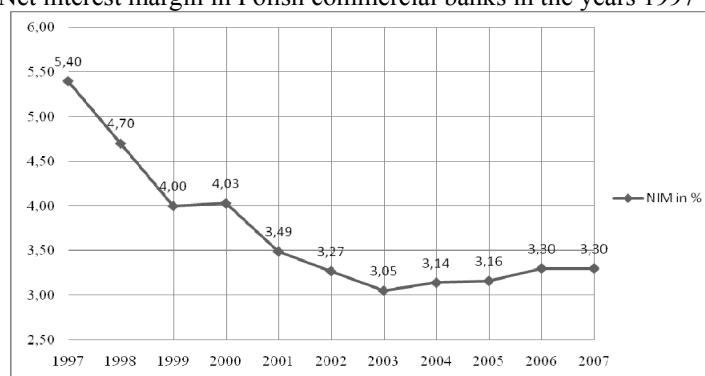
The last column presents the mean calculated for five countries with the lowest share of net interest income in total banking income in Europe, the indicator value being 42.1%.

The Polish banking sector as compared to indicators for the various groups of countries (58.2%, see Table 3) diverges significantly from the European figures, which are invariably lower than those in Poland. Therefore, in contrast to the leading banking sectors in Europe, Poland reveals a substantial, high prevalence of net interest income as the main source of banking income.

ANALYSIS OF NET INTEREST MARGIN

Net interest margin is defined as the relationship of net interest income to the average net assets (minus due interest on problem receivables) of the banking sector [Sytuacja 2003]. Its values over the period 1997-2007 are shown in Chart 2.

Chart 2. Net interest margin in Polish commercial banks in the years 1997-2007, in %

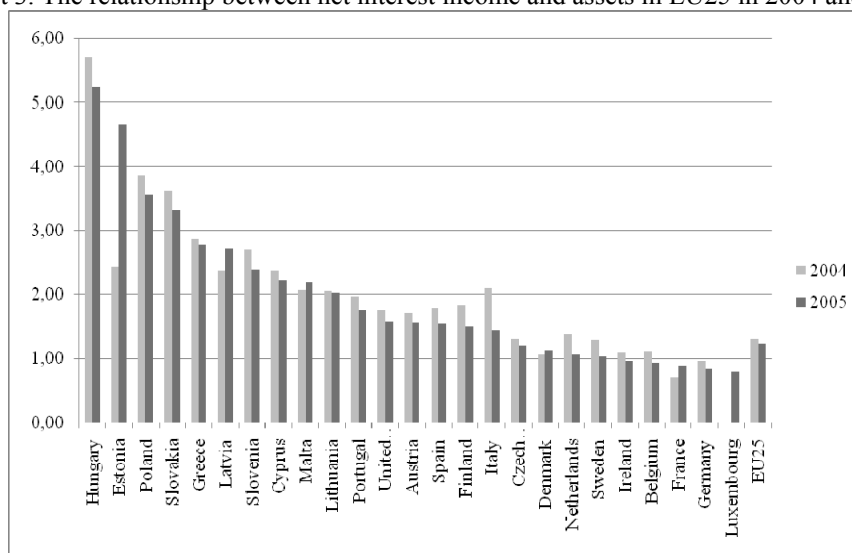


Source: own work on the basis of synthetic data from Sytuacja finansowa banków za lata 1998-2006 National Bank of Poland Warsaw 1999-2007 and Raport o sytuacji banków w 2007 roku Polish Financial Supervision Authority Warsaw 2008

In the period under study the net interest margin decreased substantially, although since 2003 it has been gradually gaining.

For the purpose of comparison of Polish margins with other European countries, data from the European Central Bank are cited. Unfortunately, the net interest margins quoted were calculated in a different way than in Poland. Thus, the indicators cited for Poland are not exactly the same as those given by the National Bank of Poland. However, the complete set of countries and their margins shown in Chart 3. enables a comprehensive overview of these phenomena in Poland as compared to other European countries⁴.

Chart 3. The relationship between net interest income and assets in EU25 in 2004 and 2005



Source: own work on the basis of EU Banking Sector Stability European Central Bank October 2005, pp.42-52 and EU Banking Sector Stability European Central Bank November 2006, pp. 52-60.

Poland is among countries with the highest net interest margins. Poland's net interest margins give it second place among reference countries and third in EU25. This situation shows that Poland is particularly dependent on lending, which is at the core of net interest income. At mostly double-digit interest rates on loans and at relatively low financing costs for banks, this situation creates a perfect handicap for the banking sector to achieve high profitability.

⁴ The Author hasn't gained access to the latest data in this area covering the whole EU. Certainly interest margins change from year to year in the individual countries but taking into consideration EU means they are rather stable in a longer period, keeping a slightly lowering trend, which may be observed in various ECB reports

In 2005, net interest margin in Poland was three times higher than the weighted average for the banking sector in the entire European Union. Among the reference countries, better conditions for banks in this respect are only found in Hungary. The other countries (including the Czech Republic) have lower or much lower margins.

In Europe, and especially in countries with the largest banking sectors, the erosion of net interest margins has long been observed, which has been reflected in successive European Central Bank reports [EU Banking 2005]. These reports mention the formidable challenges to the banking sector poised by the economic environment. As the Table shows, in many countries the existing circumstances do not even make it possible to maintain the difference between interests received and paid with respect to the volume of assets at a relatively stable level in subsequent years. In particular, this situation is true of loans, which are old, mature financial products. In the quoted report, analysts of the European Central Bank advance the thesis that this important source of banking income might become exhausted [EU Banking 2006].

Thus, the only way to preserve the margins is to increase the volume of credit, as due to fierce competition in the financial market raising the difference between the interest on receivables and liabilities of the bank would not be feasible. It seems that in order to sustain the profitability of the banks it is also necessary to increase their activity in generating profit other than net interest income: increasing income from fees and commissions, generating profit from selling various capital market products in order to take advantage of the development of the financial markets, and extending loans outside their domestic markets to the fast-growing economies, where margins are still high [EU Banking 2006]

CONCLUSION

In terms of the Polish banking sector, the question is to what degree the above-mentioned trends might apply to Poland. In the author's opinion, they might strongly influence the future shape of the Polish commercial banking sector.

The years 2003-2007 are characterized by stagnation or even a small increase in net interest margins (see Chart 2.), which conflicts with the European tendencies. In a market economy, competition should lead to a further decrease in margins. One could thus argue that the level of competition in the Polish financial market is not yet strong enough for banks to forfeit their comfortable stance. The situation is compounded by the substantial indebtedness of the State to the banking sector, which drains money out of the banking market and effectively increases loan prices.

In the time horizon of the next several years, net interest margins are not likely to decrease very fast and high interest rates will remain the central bastion of the banks' profitability. This thesis requires some substantiation and consideration of the problem in the short and long term. European Union regulations enforced a

free flow of capital, which theoretically should have triggered a rush of foreign lending institutions to the Polish financial market to seek high margins. However, this has not been the case. The western banks have penetrated the Polish market rather poorly to date. The number of branches of foreign lending institution has increased from 3 in 2004 to 14 in 2007, and the share of their gross assets in that period increased from 0.6% to 4.3% [Raport o sytuacji 2008].

Many banks that have opened their branches in Poland are companies not very well known to Polish customers [Sytuacja 2007]. The number of branches is certainly going to increase, but only in a longer time horizon. On the other hand, those banks which have been operating for a long time in the Polish financial market are certainly satisfied with their margins, which are substantially higher than in their countries of origin. Consequently, it seems that foreign banks are not going to substantially change net interest margins: they are realizing high profits and their number is still relatively small.

Furthermore, international banking is not very likely to decrease net interest margins any time soon, either. Even though nominal loans abroad have lower interest rates, their real cost in Poland may still be much higher due to higher premiums on risk demanded by foreign banks extending loans to Polish borrowers. Another factor increasing risk for both parties is exchange rate volatility, although Poland's accession to the Euro zone in several years will decrease these threats. Consequently, Polish citizens and companies do not use foreign lending services very often, they are not likely to start using them more in the near future.

A larger influx of foreign banks, forcing changes in the Polish market, will be very gradual and will take place over a longer time horizon. Therefore, in the long term the relatively high and stable margins acceptable to the Polish banks are eventually going to come to an end. The European economy will exert a strong competitive pressure on the market, which is going to necessitate changes in the prices of banking services in Poland. The openness of our economy will not make it possible for such an enclave to continue in the European or global financial market.

In summary, one may argue that the tendency to decrease interest rates is a major threat to the banks in the long term. In the future, the traditional way of conducting banking activities will have to be modified. The starting point for creating new strategies for Polish banks should be to depart from net interest income as the core model of generating banking income and embrace the prompt development of broadly understood non-interest activities potentially creating substantial incomes such as private equity. In the common opinion it can be a risky business. In fact such a consciousness may be the main reason of the phenomenon that this activity is almost given up in the Polish commercial banking sector. Private equity does not fit the banks' traditional rules because every case may be different, and this may be one of the most important reasons for its reluctance. On the other hand nobody expects a bank to provide 100 per cent of the funds in a risky project and there are wide opportunities for the diversification of risk.

Taking into consideration western experiences, process of expanding possibilities of earning money is obviously profitable for banks which are one of the main players in the European private equity market⁵. Following a similar train of thought there is no reason in rejecting these activities in Poland. Therefore shift from typical of commercial banking sector loan-deposit to capital market activity may create many benefits for the Polish banking system in the future.

BIBLIOGRAPHY:

- Baka W. (2005) Bankowość europejska Wydawnictwo Naukowe PWN Warszawa.
- Czechowska D. (2004) Procesy integracji sektora bankowego i ubezpieczeniowego Wydawnictwo Uniwersytetu Łódzkiego Łódź.
- Dębski W. (2005) Rynek finansowy i jego mechanizmy. Podstawy teorii i praktyki Wydawnictwo Naukowe PWN Warszawa.
- EU Banking Sector Stability European Central Bank November 2006.
- EU Banking Sector Stability European Central Bank October 2005
- Eurostat 2009.
- EVCA Yearbook 2008 Zaventem Belgium.
- Koepfen J. (1999): Strategie banków Europy CeDeWu Warszawa.
- Korenik D. (2005) Ważniejsze przemiany w środowisku działania banku w: Karpus P. [red.], Węclawski J. [red.] Przekształcenia rynku finansowego w Polsce Instytucje finansowe i mechanizmy ich funkcjonowania Wydawnictwo Uniwersytetu Marii Curie-Skłodowskiej Lublin.
- Ostrowska E. (2005) Banki na rynkach kapitałowych PWE Warszawa.
- Raport o sytuacji banków w 2007 roku, Polish Financial Supervision Authority Warsaw 2008.
- Schildbach J. (2008) European banks: The silent (r)evolution Deutsche Bank Research Frankfurt am Main Germany.
- Solarz J. (1997) Zarządzanie strategiczne w bankach Poltext Warszawa
- Sytuacja finansowa banków w 2002 roku. Synteza National Bank of Poland Warsaw May 2003.
- Sytuacja finansowa banków w 2006 roku. Synteza National Bank of Poland Warsaw May 2007.

⁵ Banking sector's fund raising amounted 17.7% of the total funds raised in the years 2000-2007 in the Europe (the second rank). Higher contribution came from pension funds - 23.1%, see EVCA Yearbook 2008 Zaventem Belgium

Analiza wyniku działalności bankowej w polskich bankach komercyjnych. Symptomy zagrożeń

Streszczenie: Ogólnoświatowa tendencja spadku stóp procentowych w długim okresie jest poważnym zagrożeniem dla polskich banków komercyjnych. W przyszłości tradycyjny sposób prowadzenia działalności bankowej w Polsce musi więc ulec zmianie. Punktem wyjścia powinno się stać odejście od dochodu odsetkowego netto, jako podstawowego źródła tworzenia dochodu, na rzecz szybkiego rozwoju szeroko rozumianej działalności pozaodsetkowej, w tym zwiększenia zaangażowania banków w operacje private equity. Przesunięcie akcentów z typowej jak do tej pory w polskich bankach działalności depozytowo-kredytowej w kierunku szerszego, niż dotąd, wejścia na rynek kapitałowy powinno w przyszłości przysporzyć polskiemu sektorowi bankowemu wielu profitów.

SIMULATION INVESTIGATION OF CONSEQUENCES FOLLOWING FROM TRANSPORT DECISIONS

Małgorzata Łatuszyńska

Instytut Informatyki w Zarządzaniu US
e-mail: mlat@wneiz.pl

Abstract: Different methods can be employed while investigating the consequences of transport decisions- for instance simulation methods. There are various techniques of formalizing the simulation model, from computer languages of general use and specialist simulation languages, to advanced functional simulation packages and systems. In order to illustrate the possibilities of using computer simulation model to investigate the consequences following from transport decisions, System Dynamics technique has been employed in the present article.

Keywords: computer simulation, transport decisions

INTRODUCTION

Two organizational planes can be distinguished in transport systems ([Bąkowski 1987, p. 136], [Szczepaniak 1996, p.p. 58, 173-174], [Rydzkowski & Wojewódzka-Król 1997, p. 38]), namely macroeconomic (international, national, regional) and microeconomic (transport companies, consumers of transport services). Transport decisions taken on a macroeconomic scale refer to the programming of transport system development and affect its effective functioning [Rydzkowski & Wojewódzka-Król 1997, p. 20].

Taken international plane into account, one strives after the creation of global transport market and, in order to do so, makes decisions about, for instance the harmonization of technical, environmental and socio-legal norms or the coordination of investments made in the scope of transport infrastructure of international importance [Szczepaniak 1996, pp. 187-194]. On the contrary, decisions taken on a national plane are aimed at shaping the state transport system, which consists, among other things, in the optimization of its growth as well as material develop-

ment, affecting the right course of transport processes and phenomena, and influencing the development of transport companies as well as proper transport standards ([Rydzkowski & Wojewódzka-Król 1997, pp. 38-40], [Grzywacz et al. 1994, p. 18]). As far as transport companies are concerned, decisions refer mainly to tactics and efficient actions concerning [Grzywacz et al 1994, p.137]: planning production actions, financial and material means needed; reorganization and development of company, employment; organization of production and support activity. By contrast, in the case of transport users the choice of means of transport is the main decision-making problem.

Every transport decision, regardless of plain or scale it refers to, leads to certain consequences. At the level of transport companies and transport users, it is desired to receive as great economic benefits as possible, whereas at macroeconomic level, other consequences are also crucial, e.g. consequences in the scope of natural environment or traffic safety¹.

Different methods can be employed while investigating the consequences of transport decisions. Sometimes it is enough to use a simple formula in spreadsheet, yet when decisions refer to issues important to a particular transport system (e.g. development of transport company, construction or modernization of national transport infrastructure) more advanced tool is required. Computer simulation model, understood (in accordance with Martin's definition) as a logical-mathematical presentation of a notion, system or actions programmed to be solved with the use of computer, can be such a tool [Martin 1976, p.13].

There are various techniques of formalizing the simulation model, from computer languages of general use² and specialist simulation languages³, to advanced functional simulation packages and systems. In order to illustrate the possibilities of using computer simulation model to investigate the consequences following from transport decisions, System Dynamics technique has been employed in the present article.

METHODOLOGICAL BACKGROUND TO SYSTEM DYNAMICS

System Dynamics (SD) is a simulation modeling method used mainly for the analysis of poorly structured problems with numerous interrelations among elements. It derives from the cybernetic approach to system analysis and allows to describe particularly complex systems in the form of interactive and combinatory

¹ Possible consequences following from transport decisions taken at a macroeconomic level have been thoroughly discussed in [Łatuszyńska 2004, pp. 45-59].

² The application of languages of general use in computer simulation as well as reasons behind devising the simulation languages and their qualities have been discussed, for example in [Naylor 1975, pp. 536-544].

³ Descriptions of different simulation languages have been provided, among other things, in [Naylor 1975, pp. 580-632]; [Tyszer 1990, pp. 101-144]; [Fishman 1981, pp. 113-152].

relations. One name which has been closely connected with SD since the 1950s is that of Forrester, who, together with his colleagues from the Massachusetts Institute of Technology (Cambridge, MA, USA), created its theoretical assumptions, based on traditional management theory, cybernetics and computer simulation ([Forrester 1971] and [Forrester 1961]).

The basis for building SD simulation models for quantitative analysis of system dynamic behaviors are stock and flow diagrams. Stock variables (also called state variables or levels) describe the states of the system, while flow variables (also called rate variables) depict the rates of change of stocks. Stocks are accumulations of their flows and are calculated as the integration of net inflows, i.e. [Sterman 2000, p. 194]:

$$Stock(t) = \int_{t_0}^t [Inflow(s) - Outflow(s)]ds + Stock(t_0) \quad (1)$$

with Inflow(s) and Outflow(s) denoting the values of the inflow and outflow at any time s between the initial time t₀ and the present time t. Conversely, the net flow determines the rate of change of any stock, i.e. its time derivative, by means of the following differential equation [Sterman 2000, p. 194]:

$$d(Stock)/dt = Inflow(t) - Outflow(t) \quad (2)$$

It can be noticed that the stock and flow notation provides a general way to characterize any process graphically ([Coyle 2000] and [Wolstenholme 1990]). Mathematical apparatus as well as principles of modeling that conform to the convention of system dynamics have been referred to in a great number of publications⁴. Course in system dynamics is available on the Internet⁵.

SD MODELS FOR INVESTIGATING TRANSPORT DECISIONS CONSEQUENCES

A wealth of evidence indicates that SD serves very well as a method for investigating consequences following from transport decisions. Examples of SD models built to support the process of transport decisions analysis have been shown in Table 1. The aforementioned models have been chosen in a more or less random way as it is difficult to refer to all information concerning the application of system dynamics in the aspect under discussion. Materials from international conference, that has been organized annually for 25 years by System Dynamics Society, are a valuable source of reference⁶.

⁴ For instance Kirkwood [1998], Biniak [2002], Tarajkowski [2008], Łatuszyńska [2008].

⁵ Road Maps: A Guide to Learning System Dynamics. Available on the Internet: <http://sysdyn.clexchange.org/road-maps/home.html>.

⁶ Detailed information in the form of conference materials is available on the following website: www.systemdynamics.org.

Table 1A. SD models for investigating transport decisions consequences: microeconomic scale

Model	Model purpose and consequences investigated
Microeconomic scale	
Model for forecasting the development of railway company [Gottschalk 1980]	<ul style="list-style-type: none"> • analysis of the influence of implementation of solutions improving the situation of company, including: <ul style="list-style-type: none"> - additional government loans, - freedom of shaping transport rates, - ceasing the exploitation of railway tracks in areas losing industrial base, - the influence that the implementation of solutions increasing the productivity has on the situation of company
Model of passengers transport system [Łatuszyńska 1992a]	<ul style="list-style-type: none"> • analysis of the situation of transport company in conditions of growing transport needs • investigating the influence of technological changes (new types of buses) and changes in values of different elements in transport services market (e.g. fuel prices, how frequent buses run etc.) on the situation of company
Model of transport-shipping company [Miłosz et al 1996]	<ul style="list-style-type: none"> • supporting the process of making strategic decisions • examining the effects of different strategic decisions taken in transport-shipping company

Source: own study

To explain the possibilities of using System Dynamics for investigating consequences following from transport decisions more precisely, the example of SD model has been presented below.

SHIPPING COMPANY MODEL

The model has been construed particularly for the sake of the present paper. The model is to investigate the situation of a hypothetical small shipping company possessing one merchant ship and offering transport services between European ports. The ship is used for carrying container cargo.

The aim of the model is to answer the question about the financial possibilities of the ship owner over several dozen years. Moreover, the model enables:

- to simulate possible development of the company's fleet and reveal measurable effects of this development,
- to show the influence of some changes in the system investigated on financial results of the shipping company,
- to gain an insight into solutions adopted in the company.

The idea of the model has been illustrated on the flow and stock diagram of the model identifying the major causal effects and feedback loops between variables playing an important role in the company under consideration (Fig. 1).

Table 1B. SD models for investigating transport decisions consequences: macroeconomic scale

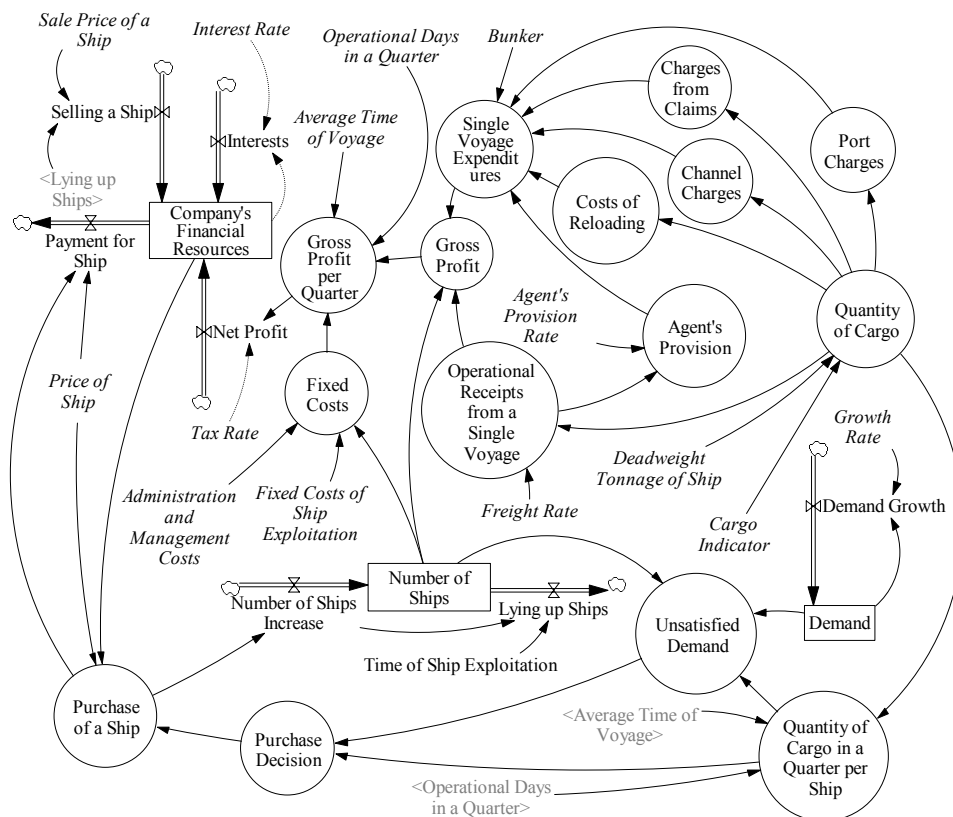
Model	Model purpose and consequences investigated
Macroeconomic scale	
Model for planning international carriage of goods [Łatuszyńska 1992b]	<ul style="list-style-type: none"> • testing transport policy scenarios • investigating the influence of different investment variants on changes in parameters of transport system under examination, e.g. speed, transport distance, transport routes etc.
ASTRA Model [ASTRA 2000]	<ul style="list-style-type: none"> • long-term strategic analysis at EU level • consequences investigated: modal shifts, generalized cost of transport, time savings, transport flows in network, network load, socio-economic indicators, impact on safety and the environment
Model for the 6th Pan-European Transport Corridor [Łatuszyńska 2005]	<ul style="list-style-type: none"> • research on effects of development of the specific transport corridor • effects estimated: transport costs per relation, transport time per relation.
Model for the market penetration of passenger cars with new drive train technologies [Bosshardt et al 2006]	<ul style="list-style-type: none"> • conceptualizing the main micro processes of technological change in the Swiss car fleet • investigating the influence of new drive train technologies, focusing on natural gas vehicles (NGVs), on energy consumption
Model for examining the effects of changes in land use and transportation planning [Stave & Dwyer 2006]	<ul style="list-style-type: none"> • analysis of the land use system • investigating the potential effects of land use and transportation systems on air quality, traffic congestion, and other quality of life factors
Model for testing of CO ₂ -emissions' reduction [Grit et al 2007].	<ul style="list-style-type: none"> • regulatory impact assessment for the transportation sector in the case of Germany • testing legal instruments for reduction of CO₂-emissions within the transportation sector
Model of Transportation Demand Management [Vakili et al 2008]	<ul style="list-style-type: none"> • comparing diverse policies of Transportation Demand Management (TDM) • investigating the impact of TDM strategies in the case of Tehran

Source: own study

Due to the fact that the model has a rather cognitive than application character, certain simplifying assumptions have been made. For instance, it has been as-

sumed that demand for services provided by amateur company is increasing, which leads to incompatibility between actual and potential transport. The incompatibility indicates that a decision about increasing the tonnage, which ship owner has at his/her disposal, can be made. If the company has accumulated essential financial means in a bank, it can buy a new means of transport – in this sense the number of ships increases and hence the fleet can carry bigger cargo. The greater the number of ships, the greater the global income from services and financial means (due to which it is more likely that the fleet will continue to develop).

Fig. 1. The flow and stock diagram of the shipping company model

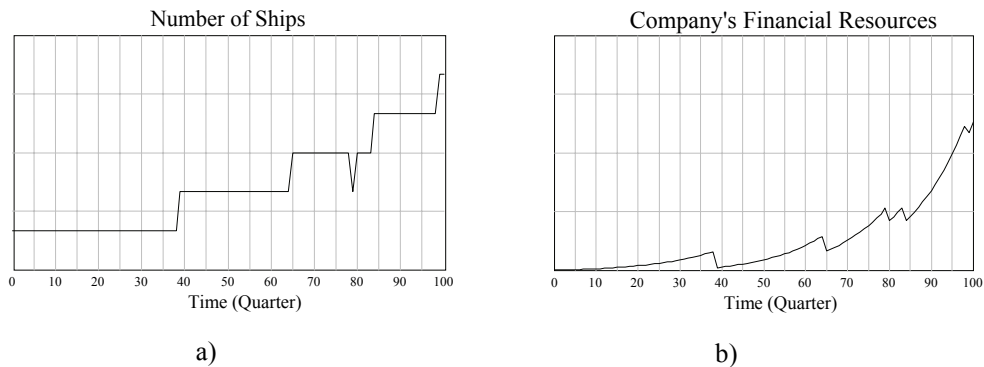


Source: own elaboration

The results of simulation experiments carried out on the model can be a source of information about the functioning of the system over time and about the influence that potential development of the fleet has on financial situation of company. It can also be used for investigating the impact that changes in parameters have on potential development and general situation of a firm.

Figure 2 shows the results of demonstration experiment that illustrate the influence of decision about increasing the tonnage (that company has at its disposal – Number of Ships) on the financial situation of company (Company's Financial Resources).

Fig. 2. Results of exemplary simulation experiment

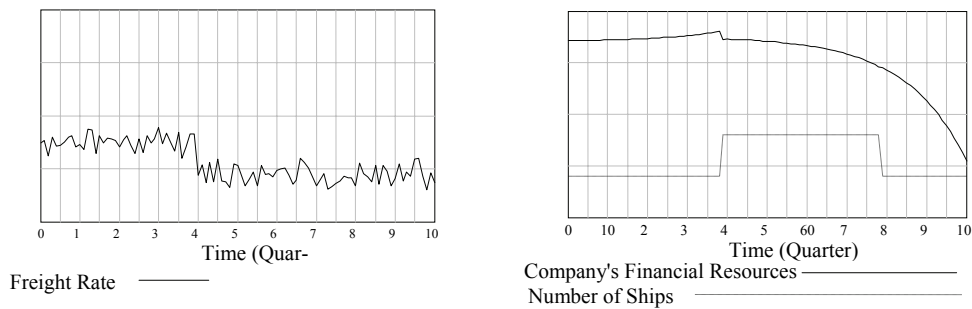


Source: own elaboration

The experiment was done for 100 quarters. It has been assumed that the tonnage will be increased when there is proper demand for services provided by the ship owner and he/she has his/her own financial means for financing the investment. For the sake of the experiment, it has been assumed that the ship owner gets rid of 20 ships after a 20-year operation, which has been shown in Figure 2a (decrease in the number of ships in the 80th quarter). Increasing the number of ships is accompanied by increasing financial means of company (Fig. 2b), which results from higher global income, on the assumption that freight market is stable. However, another scenario can be considered, namely that freight rate will fluctuate around a certain value expected and, what is more, will drop by several points in a given moment. Figure 3 shows results obtained for this scenario. As it can be noticed, tonnage increase is not plausible and financial situation of company is worsening considerably in such a scenario.

Summing up the discussion on the model presented, it should be emphasized that the experiments done are just a simple illustration of the possibilities of examining various decision-making scenarios in a certain transport system.

Fig. 3. Results of demonstration simulation experiment



Source: own elaboration

CONCLUSION

The analysis of the hitherto existing applications of computer simulation models, i.e. for anticipating consequences following from transport decisions taken at different organizational planes, proves that these models perform well in the aspect discussed. Employing such models, one can carry out practically unlimited number of simulation experiments that allow to find out consequences following from potential transport decisions without interfering in a living organism of the system. Furthermore, a detailed analysis of simulation results enables one to understand a certain transport system in a better way and facilitates the formulation of suggestions to its improvement, which would be difficult if not impossible to reach with the use of other methods.

BIBLIOGRAPHY

- ASTRA (2000) Final Report. Institut für Wirtschaftspolitik und Wirtschaftsforschung, Universität Karlsruhe.
- Bąkowski W. (1987) Podstawy organizacji i zarządzania w transporcie, Wydawnictwa Komunikacji i Łączności, Warsaw.
- Biniak Z. (2002) Elementy teorii systemów, modelowania i symulacji, INFOPLAN, www.finus.com.pl.
- Bosshardt M., Ulli-Beer S. A., Gassmann F., Wokaun A. (2006): Conceptualising the main micro processes of technological change in the Swiss car fleet, [in] A. Größler, E. A. J. A. Rouwette, R. S. Langer, J. I. Rowe, J. M. Yanni: Proceedings of the 24th International Conference of the System Dynamics Society, Nijmegen.
- Coyle, G. (2000) Qualitative and Quantitative Modeling in System Dynamics: Some Research Questions. *System Dynamics Review* 16(3)/2000.
- Fishman G.S. (1981) Symulacja komputerowa – pojęcia i metody. PWE, Warsaw.

- Forrester, J.W. (1971) Counterintuitive behavior of social systems, *Technology Review*, Alumni Association of the MIT, Cambridge-Massachusetts.
- Forrester, J.W. (1961) *Industrial Dynamics*. The MIT Press and Wiley. New York.
- Gottschalk P. (1980) *The Dynamics of Railroad Decline and Revitalization*, Norges Teknisk – Naturvitenskapelige Forskningsråd, Oslo.
- Grit W., Meyer G., Spengler T., Wansart J. (2007) *Regulatory Impact Assessment for the Transportation Sector – Case Study Germany*, [in] J. Sterman, R. Oliva, R. S. Langer, J. I. Rowe, J. M. Yanni (eds). *Proceedings of the 25th International Conference of the System Dynamics Society*. Boston, Massachusetts.
- Grzywacz W., Wojewódzka-Król K., Rydzkowski W. (1994) *Polityka transportowa*, Wydawnictwo Uniwersytetu Gdańskiego, Gdansk.
- Kirkwood C. W. (1998) *System Dynamics Methods: A Quick Introduction*. Arizona State University, www.public.asu.edu/~kirkwood/sysdyn/.
- Łatuszyńska M. (1992a) *Model systemu przewozów pasażerskich*, *Studia Informatica* Nr 5, Uniwersytet Szczeciński, Szczecin.
- Łatuszyńska M. (1992b) *Symulacyjne badanie rynku towarowych usług przewozowych w międzynarodowych korytarzach transportowych*, Uniwersytet Szczeciński, Szczecin.
- Łatuszyńska M. (2005) *Modelling of Consequences of Transport Corridors Development*. LIBERTAS Paper 62.
- Łatuszyńska M. (2004) *Modelowanie efektów rozwoju międzynarodowych korytarzy transportowych*. Wydawnictwo Naukowe Uniwersytetu Szczecińskiego, Szczecin.
- Łatuszyńska M. (2008) *Symulacja komputerowa dynamiki systemów*. Wydawnictwo PWSZ, Gorzów Wielkopolski.
- Martin F.F. (1976) *Wstęp do modelowania cyfrowego*, PWN, Warszawa.
- Miłosz E., Miłosz M., Mitraszewska I. (1996) *Modelowanie działalności przedsiębiorstwa transportowo-spedycyjnego*, Zeszyt 81 ITS Warszawa.
- Naylor T.H. (1975) *Modelowanie cyfrowe systemów ekonomicznych*, PWN, Warsaw.
- Road Maps: A Guide to Learning System Dynamics. Available on Internet: <http://sysdyn.clexchange.org/road-maps/home.html>.
- Rydzkowski W., Wojewódzka-Król K. (eds.) (1997) *Transport*, Wydawnictwo Naukowe PWN, Warsaw.
- Stave K. A., Dwyer M. F. (2006) *Lessons from LUTAQ: Building systems thinking capacity into land use, transportation, and air quality planning in Las Vegas* [in] A. Größler, E. A. J. A. Rouwette, R. S. Langer, J. I. Rowe, J. M. Yanni: *Proceedings of the 24th International Conference of the System Dynamics Society*, Nijmegen.
- Sterman, J.D. (2000) *Business Dynamics: Systems Thinking And Modeling For A Complex World*. Irwin/McGraw-Hill, Boston.
- Szczepaniak T. (ed.) (1996) *Transport międzynarodowy*, Państwowe Wydawnictwo Ekonomiczne, Warsaw.

- Tarajkowski J. (ed.) (2008) Elementy dynamiki systemów. Wydawnictwo Akademii Ekonomicznej w Poznaniu, Poznań.
- Tyszer J. (1990) Symulacja cyfrowa. WNT, Warsaw.
- Vakili K., Isaai M.T., Barsari A. J. (2008) Strategic Assessment of Transportation Demand Management Policies: Tehran Case Study. [in] Brian C. Dangerfield (ed.), Proceedings of 26th International Conference of the System Dynamics Society, Athens.
- Wolstenholme, E.F. (1990) System Enquiry: A System Dynamics Approach. Wiley, Chichester.

Symulacyjne badanie skutków decyzji transportowych

Streszczenie: W badaniu skutków decyzji transportowych można posłużyć się różnymi metodami, w tym symulacyjnymi. Istnieją różne techniki formalizacji modelu symulacyjnego, począwszy od komputerowych języków ogólnego zastosowania, poprzez specjalistyczne języki symulacyjne po zaawansowane funkcjonalnie pakiety i systemy symulacyjne. W niniejszym artykule, w celu zilustrowania możliwości zastosowania komputerowego modelu symulacyjnego do badania skutków decyzji transportowych, posłużono się techniką dynamiki systemowej.

Słowa kluczowe: symulacja komputerowa, decyzje transportowe

RESULTS OF MISTAKEN TIME PERIOD IN ANALYSIS IN THE CASE OF FRAMING EFFECT FOR SOME CAPITAL MARKETS' MODELS

Agnieszka Majewska, Sebastian Majewski

Katedra Ubezpieczeń i Rynków Kapitałowych US
e-mail: magnes@wneiz.pl; fisherman@autograf.pl

Abstract Economic activity around the world should be supported by analysis determined by the different level of difficulty of applied quantitative methods. The article presents the importance selecting a proper time window for analysis for correct diagnosing of situations based on information taken from TNI, BSOPM, Markowitz's, and Sharpe's models. TNI shows the role of expansiveness of media relations regarding the SE. The correct specification of the time window causes the proper valuation of price and risk. Behavioural finance flaming effect is a field of discussion in his research.

Keywords: behavioural finance, Black-Scholes Option Pricing Model, Tragic News Indicator, Sharpe's model, framing effect

INTRODUCTION

Economic activity around the world should be supported by analysis determined by the different level of difficulty of applied quantitative methods. Supporters of Becker's theory would claim that such a situation can be achieved by trying to calculate in the decision making process. Their opponents – behavioural economists – will explain such events as willingness to justify decisions or to blur responsibility for investment decisions. In both cases, the same problem is exact specification of time period for analysis.

The article presents the importance selecting a proper time window for analysis for correct diagnosing of situations based on information taken from TNI (tragic news indicator), BSOPM, Markowitz's, and Sharpe's models.

TNI shows the role of expansiveness of media relations regarding the Stock Exchange. The correct specification of the time window for TNI allows determination of how strong an influence negative information has on changes of the Stock Exchange indexes and how strong speculative information is (supported by BNI – bad news indicator and GNI – good news indicator). In the case of a capital market

model, correct specification of the time window causes the proper pricing of a financial instrument, and proper valuation of risk.

Statistical data used in this article are taken from the Warsaw Stock Exchange. Quotations of instruments of capital market in Poland from September 2008 to February 2009 and press information from Polish Press Agency are used.

The influence of wrong time period specification for three groups of financial market analyses were analysed in this paper:

- Technical analysis (Tragic News Indicator)
- Portfolio analysis (Markowitz and Sharpe models)
- Derivatives market analysis (BSOPM)

THEORETICAL ASPECTS

This article is closely tied with the prospect theory, because of its attempt to explain some motives to taking investment decisions.

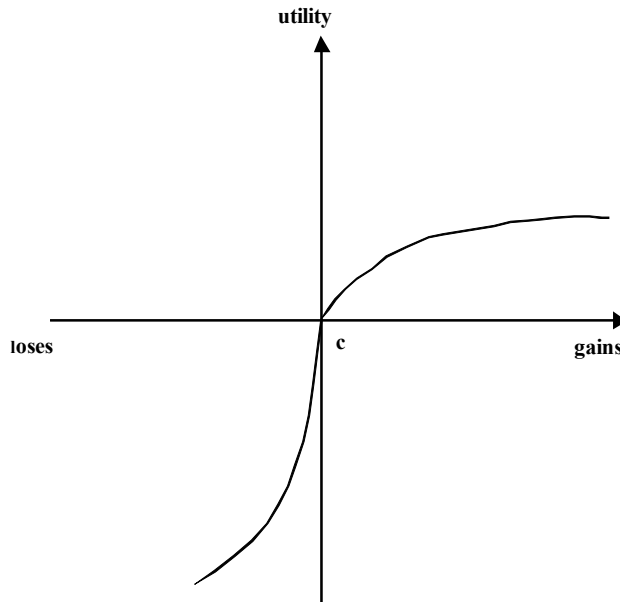
Human behaviour in the last twenty years has played an important role in building value of stock exchanges and creating interactions between market participants, etc. D. Kahneman and A. Tversky (1979) created a new theory, which has had to explain why people are not rational in the sense of G. Becker's theory (1990). In this work, the authors showed that there is significant difference between behaviour of investor incurring losses and gains. Figure 1 shows difference in level of emotions during a process of receiving information about results of economic activity.

The meaning of this chart is that a human is more sensible when incurring losses than earnings. This difference is about 2.5 times more for losses than earnings. It means that when an investor loses \$100 he has to earn about \$250 because of compensation losses (figure 1 shows such type of differences in utility of gains and losses).

The framing effect could be explained best by the two following citations:

1. "The term frame dependence means that the way people behave depends on the way that their decision problems are framed" (Shefrin 2000).
2. "The rational theory of choice assumes description invariance: equivalent formulations of a choice problem should give rise to the same preference order (Arrow, 1982). Contrary to this assumption, there is much evidence that variations in the framing of options (e.g., in terms of gains or losses) yield systematically different preferences (Tversky and Kahneman, 1986)."

Figure 1. Utility function for earning and losses with respect to point of reference c .



Source: [Rabin M., Psychology and economics, Department of Economics University of California – Berkeley, 1996], page 8.

MODELS USED IN THE RESEARCH

TRAGIC NEWS INDICATOR (TNI)

A point of view in most cases is formed by media. Mainly it concerns a perception of risk. People tend to assign greater risk to objects and events, on which the attention of media is more focused. This is due to form of presentation. For example, all plane crashes are presented in the form of pictures, stimulating imagination, yet car crashes are presented as statistical data. It is because the probability of plane crashes is smaller those of cars. However, if an ordinary consumer has been asked what kind of accident he fears more, most would choose plane crashes. Analogous with all information about crises, crashes and bankruptcy are overestimated.

There were many experiments, which aim to explain the influence of information on valuation of stocks in future. Andreassen (1987), presented fictitious news and quotations of stocks (positive and negative) to a selected group of investors. After that, investors had to estimate the new value of these stocks. An alternative group of investors in predictions of a stock's future value most often use regression as the mean process of a selected group. A selected group of investors predicted stable growth of stocks. This means that when an investor is under systematic press noise, he misuses knowledge and intuition in the valuation process.

Shiller (2001), described another such phenomenon – speculative bubble, which is closely tied with media influence on financial markets. A speculative bubble is an unsustainable increase in prices brought on by investors' buying behaviour, rather than by fundamental information. There are twelve factors that may explain existing or creation of speculative bubbles in the financial markets. One of them is an expansion in media reporting of business news.

Press economic information, taken from Polish Press Agency (PAP), could be treated as a source of media noise, which has an influence on the value of stocks quoted on the Warsaw Stock Exchange. All economic news should be divided into three groups: neutral information, positive information and negative information. For each type of negative information, it is possible to mark the most tragic news with respect to the economy. Negative information is a cause of an extreme behaviour of investors on the market and this a reason of focusing this research only on tragic information. As the most tragic news author recognized such news, which in its title have had such words as crisis, fall down, or panic. On the basis of this data, an indicator of media expansiveness was constructed. This indicator is a quotient of the number of articles in a chosen group of information, and the number of all PAP information in particular day.

$$TNI = \frac{NTN}{TNN} \cdot 100\% \quad (1)$$

where *TNI* – tragic news indicator; *NTN* – number of tragic headlines; *TNN* – total number of news items.

The authors used methods of correlation and regression analysis to identify the behavioural character of dependency between analysed variables. Because of a short time period in this research, statistical verification of significance of correlations is very important. The level of confidence in this case is set at 0.05.

In this research, correlation coefficients between rates of return of indexes of chosen Stock Exchanges were presented first. Next, the research will show how strong is the influence of media relations on the changes in prices of stocks quoted on the Warsaw Stock Exchange.

BLACK-SCHOLES PRICING MODEL

This part of the efficiency tests is based on the Black-Scholes pricing formula, with six kinds of volatility used in the valuation process. Classical BSOPM formula could be presented using equations:

For a call option in *t*-th moment:

$$c = S \cdot N(d_1) - X \cdot e^{-r(T-t)} N(d_2) \quad (2)$$

where

$$d_1 = \frac{\ln(S/X) + (r + \sigma^2/2) \cdot (T - t)}{\sigma \sqrt{(T - t)}} \quad (3)$$

$$d_2 = \frac{\ln(S/X) + (r - \sigma^2/2) \cdot (T - t)}{\sigma \sqrt{(T - t)}} = d_1 - \sigma \sqrt{(T - t)}. \quad (4)$$

where S – The value of the underlying asset in moment t , X – Strike price of option, r – Domestic risk-free interest rate, T – Time to option expiry, σ – Volatility of the underlying asset, $N(d)$ – normal cumulative distribution function of d_1 or d_2 , respectively.

With using put-call parity¹, we get an equation for European put option value:

$$p = X \cdot e^{-r(T-t)} N(-d_2) - S \cdot N(-d_1) \quad (5)$$

MV MODEL

Harry Markowitz created theory of effective selection of assets to portfolio (*MV – Mean-Variance Model*). This theory is a basis of most modern models in portfolio analysis. In this work, we use two classical formulas:

The first goal function:

$$S^2 = X' \cdot D \cdot X \rightarrow \min \quad (6)$$

where X – The vector of shares in portfolio, X' – the transposed vector of X , D – The matrix of variance and covariance of rates of return

With conditions:

$$1) X_i \geq 0 \text{ for } i = 1, 2, \dots, n$$

$$2) \sum_{i=1}^n X_i = 1$$

The second goal function:

$$R_p = \sum_{i=1}^n X_i \cdot R_i \rightarrow \max \quad (7)$$

With conditions:

1), 2) and

3) $S_p = a$, and a is a level of accepted risk

SHARPE'S MODEL

¹ Stoll H.R. (1969)

Sharpe's model is a simple econometric model, which aims to explain how sensible the effects of are rates of return are on changes in markets rate of return. The market rate of return is represented by that of the stock exchange index (in the Polish conditions WIG20).

We could show this model as an equation:

$$R_{it} = \alpha_i + \beta_i \cdot R_{mt} + u_t \quad (8)$$

where R_{it} – rate of return, R_{mt} – market rate of return, α_i , β_i – structural parameters of an econometric model estimated by OLS procedure.

In addition, total risk is described by the formula:

$$S_i^2 = \beta_i^2 \cdot S_m^2 + S_e^2 \quad (9)$$

Where S_i^2 – total risk parameter, S_m^2 – variance of market rates of return

EMPIRICAL RESULTS

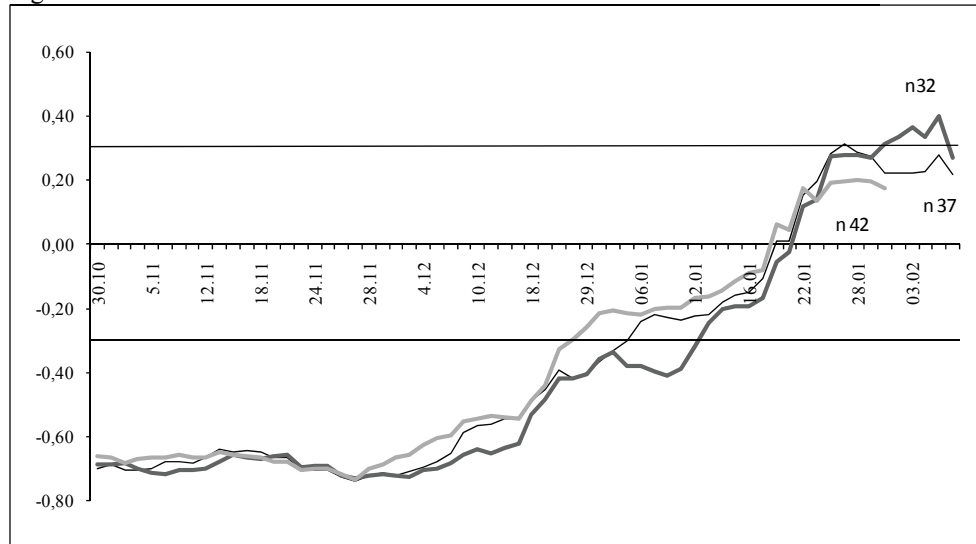
The correlation between TNI and changes in WIG20 was analysed in the beginning of the research. Authors observed changes in correlation coefficients, which were results of changing a time period (minimum 32 observations, and maximum 47 observations for the estimation of correlation coefficient). The results are presented in figure 2.

As shown in figure 2, generally the correlation coefficient implies that the importance of press information slowly loses power. In the beginning, the level of dependence was about -0.7, and then entered in the area of insignificance (two black lines on the figure). The information taken from n32-line is untrue because of significant, positive value of the correlation coefficient. Moreover, because TNI comes from technical analysis, the reaction of n42-line is too slow. So frame coefficient (FC) was constructed for identifying the proper time window. The FC is presented as an equation:

$$FC = \frac{RR}{VaR} \quad (7)$$

where RR – rate of return (logarithm), VaR – value of risk

Figure 2. Correlations coefficients between TNI and WIG20



Source: own research

Table 1. Changes in frame coefficient for WIG 20 with the hypothetical investment value 1000 PLN

	RR36	RR37	RR38	RR39	RR40	RR41	RR35	RR34	RR33	RR32	RR31
max	73,75	58,85	69,42	80,26	90,75	114,03	79,89	80,40	80,37	107,47	107,47
VaR	23,58	23,45	23,45	23,35	23,32	23,22	23,69	23,84	23,96	23,86	24,00
FC	3,12	2,51	2,96	3,44	3,89	4,91	3,37	3,37	3,35	4,50	4,48

Source: own research

The Authors made the assumption that if an investor loses \$100 he has to earn about \$250 because of compensation losses, so FC should be close to 2.5. From this table it is possible to identify the best time window from the perspective of frame effect. In this case, RR37 (rates of return for 37 observations) has had the best estimations.

One of the most important variables in the BSOPM model is the volatility of underlying asset. Accordingly, the authors tested the volatility for defining the proper time window for the analysed model. The results are presented in Table 2:

It was necessary to construct a new one frame coefficient for this model. In addition, if the relation of earnings to possibility of gains is a point of discussion in this article, so a value of FC should be estimated using an equation:

$$FC = \frac{RR}{\sigma} \quad (8)$$

where RR – maximum of logarithmic rates of return for analysed time period (8,15%), σ – risk parameter defined by standard deviation (sigma).

Table 2. The frame coefficient for WIG 20 volatility

No. of days	VaR	ln RR	FC	Sigma
70	6,771	-23,64%	2,92	2,80%
75	6,900	-8,95%	2,81	2,90%
80	7,030	-21,43%	2,71	3,01%
85	7,168	-33,92%	2,60	3,13%
90	7,260	-43,51%	2,54	3,21%
95	7,225	5,33%	2,56	3,18%
100	7,173	7,82%	2,60	3,14%
105	1,792	-9,84%	2,59	3,15%
110	7,129	-1,64%	2,63	3,10%
Range	0,488	51,33 p.p.		0,42 p.p.

Source: own researches

The proper time window for this situation is 90 days, because the FC indicator is closer to 2.5. Interestingly, this sigma parameter is highest in the analysed group. Also, a one day-risked value for the amount of 1000 PLN for this time period is highest, and the gain is the biggest one (real logarithmic rate of return amounted to -43,51%).

Next, research was conducted to verify the time period for the portfolio selection model MV in two cases: minimization of risk, and maximization of rate of return. The analysed portfolio was constructed on a base of the 20 shares, which have the greatest capitalizations on the Warsaw Stock Exchange. Results are presented in Table 3:

In this table, there are results of two diversification attempts. If the FC coefficient is constructed in the same way as in the case of BSOPM testing, it is possible to underline an 80-day period as the best estimation, but only for maximization procedure. It is a normal situation, but only for models in which a goal is to increase hypothetical earnings. Interestingly, in this case the best potential rate of return is not for the best frame coefficient.

Table 3. Empirical results for Markowitz model

No. of days	min			max		
	Sigma	RR	FC	Sigma	RR	FC
70	17,57%	20,48%	1,17	27,94%	62,43%	2,23
75	17,82%	19,97%	1,12	31,52%	76,42%	2,42
80	18,98%	23,39%	1,23	27,12%	66,82%	2,46
85	19,84%	15,02%	0,76	32,06%	69,58%	2,17
90	19,54%	13,61%	0,70	97,08%	36,80%	2,63
95	19,36%	8,87%	0,46	85,07%	37,33%	2,28
100	19,21%	7,16%	0,37	37,02%	70,71%	1,91
105	19,05%	6,00%	0,32	40,38%	76,01%	1,88
110	18,72%	4,09%	0,22	42,01%	74,82%	1,78
Range	2,27 p.p.	19,30 p.p.		69,96 p.p.	39,62 p.p.	

Source: own researches

Next, a test of proper time period was made for Sharpe's model for portfolios constructed using Markowitz's procedure. Results of this calculation are presented in Table 4:

Table 4. Empirical results for Sharpe's model

No. of days	beta	alfa	Rp
70	0,773	0,005	-17,81%
75	0,887	0,005	-7,41%
80	0,757	0,004	-15,81%
85	0,767	0,006	-25,44%
90	0,848	0,008	-36,14%
95	0,872	0,007	5,32%
100	0,887	0,005	7,44%
105	0,929	0,006	-8,55%
110	0,969	0,006	-0,97%
Range	0,212		43,58 p.p.

Source: own researches

Table 4 presents econometric model parameters (β , α), and hypothetic rate of return for the portfolio. The highest level of earnings is reached with a 100-day time period, and the worst for a 90-day period. A difference (spread) between these two estimations of R_p is 43.58 percentage points.

For these critical cases, the authors present risk calculations in groups of specific, market, and total in Table 5:

The strongest characteristic in risk estimation is the fact that differences between 90-days time period and 100-days time period are very small in terms of total risk. The structure of total risk determines the final result in the level of rate of return. The smaller level of specific risk is positively correlated with potential earnings.

Table 5. Risk calculations for Sharpe's model

No. of days	specific	market	total
90	1,67%	2,26%	2,81%
100	1,48%	2,37%	2,80%

Source: own researches

The final test was focused on the results of options pricing with using Black Scholes Option Pricing Model. There were two kinds of results, presented separately: for call options and put options. Final solutions of BSOPM are shown in Table 6:

Table 6. Differences in option prices and theoretical results of BSOPM

No. of days	range call	range put
70	155,53%	53,19%
75	133,85%	54,52%
80	121,63%	55,27%
85	102,10%	56,57%
90	101,35%	57,38%
95	100,32%	57,94%
100	85,34%	61,58%
105	98,54%	58,60%
110	96,66%	59,11%

Source: own researches

In the analysed case, there are significant differences between results for put and call options, but the 90-days period proper for standard deviation is not one of them. It is important because in BSOPM the authors used standard deviation as a parameter of volatility. The best result was chosen based on the range (call or put). This range was a difference between maximum and minimum percentage deviation of a theoretical price of the option from a market price. For call options, the proper time period is 100 observations, and for puts 70 in the analysed case.

CONCLUSIONS

This work shows the complications in making analyses of capital markets. Each investor waits for a method, which will give him enough time for making a decision. The most important factor is the interpretation of results. There are six separate conclusions to be drawn from this research:

1. There is no universal time period for financial models.
2. In case of technical analysis, even 6 days would change the meaning of an indicator.
3. For Markowitz's model, the best time was 80 and 90 days in the FC criterion.
4. We could use a FC goal function as maximization utility criterion.
5. For BSOPM the best results were for 70 days (put option), and 100 days (call option).
6. BSOPM is more effective for the level of exercise price, than for changing in the duration of volatility.

REFERENCES

- Barbies N., Thaler R. (2002), A Survey of Behavioral Finance, Handbook of the Economics of Finance, 2002
- Barber B. M., Odean T. (2001), Boys will be boys: Gender, overconfidence and common stock investment, Quarterly Journal of Economics, 2001
- Barber B. M., Odean T. (2001), Trading is hazardous to your wealth: The common stock investment performance of individual investors, Journal of Finance, LV, 2001.
- De Bont W., Thaler R., (1985) – Does the Stock Market Overreact?, Journal of Finance, Vol. 40, No 3, pp. 793-805
- French K., Poterba J. M. (1991) – Investor Diversification and International Equity Markets, NBER Working Papers Series No 3609
- International Corporate Finance (2005), Are Investor Home Biased? – Evidence from Germany, November 2005
- Kahnemann D., Tversky A. (1992), Advances in Prospect Theory: Cumulative Representation of Uncertainty, Journal of Risk and Uncertainty, 5, 1992, pp. 297-323
- Kahneman D., Tversky A., (1979), Judgment under Uncertainty: Heuristics and Biases, Science, New Series, Vol. 185, No. 4157. (September 27, 1974), pp. 1124-1131
- Kahnemann D., Tversky A. (1979), The Prospect Theory: An Analysis of Decisions Under Risk, Econometrica vol. 47, No. 2, March 1979, pp. 263-291
- Kilka M., Weber M. (2000) – Home Bias in International Stock Returns Expectations, Journal of Psychology and Financial Markets, Vol. I, No 3&4, pp. 176-192

- Koons L., McAnally M., Mercer M., (2001), How do investor judge the risk of derivative and non derivative financial items? The Univeristy of Texas, unpublicized report from researches, 2001
- Lewis K.K., (1999) – Trying to Explain Home Bias in Equities and Consumption, *Journal of Economic Literature*, Vol. 37, pp. 571-608, 1999
- Majewski S. (2007), Czy w budownictwie występuje efekt banki spekulacyjnej?, *Rynki finansowe*, UMCS Lublin 2007
- Majewski S. (2007), Efekt „banki spekulacyjnej” na Giełdzie Papierów Wartościowych w Warszawie?, *ZN US Metody ilościowe w ekonomii* Nr 450, Szczecin 2007
- Majewski S., (2009), The media and the prices creation in Poland, *International Journal of Management Cases*, Volume 11, Issue 1, April 2009
- Majewski S. (2005), Piramidalna struktura portfeli funduszy inwestycyjnych, *ZN US Finanse, rynki finansowe, ubezpieczenia*, Nr 413, Szczecin 2005
- Majewski S. (2006), Społeczna odpowiedzialność spółek publicznych czy celowy wybór Towarzystw Funduszy Inwestycyjnych, *ZN US Metody ilościowe w ekonomii* Nr 415, Szczecin 2006
- Rabin M., (1996), *Psychology and economics*, Department of Economics University of California – Berkeley, 1996
- Shefrin H. (2002), *Beyond Greed and Fear*, Oxford University Press, New York 2002
- Shefrin H., Statman M. (2000), Behavioral Portfolio Theory, *Journal of Financial Quantitative Analysis*, Vol. 35, No. 2, 2000
- Shiller R.J., (2001), *Irrational Exuberance*, Broadway Books, New York 2001
- Tarczyński W., (1997), *Rynki kapitałowe. Metody ilościowe*, Placet, Warszawa 1997
- Thaler R.H. (1999), Mental Accounting Matters, *Journal of Behavioral Decision Making* 12, p. 183-206, 1999
- Thaler R., Shefrin H., (1981) – An Economic Theory of Self-Control, *Journal of Political Economy*, Vol. 89, No 2, pp. 392-406, 1981
- Zaleskiewicz T., (2003) *Psychologia inwestora giełdowego. Wprowadzenie do behawioralnych finansów*, GWP Gdansk 2003
- Zielonka P. (2004), *Finanse behawioralne, Psychologia ekonomiczna*, Tyska T. (ed.), Gdanskie Wydawnictwo Psychologiczne, Gdansk 2004, pp. 334-367

Skutki błędnej cezurę czasowej w świetle efektu framingu dla wybranych modeli rynku kapitałowego

Streszczenie: Prowadzenie każdej działalności gospodarczej na świecie wymaga wspierania jej przez prowadzone na różnym poziomie abstrakcji analizy. W niniejszym artykule zostanie zaprezentowane znaczenie wyboru wła-

ściwego okna czasowego do analiz dla prawidłowego diagnozowania sytuacji wynikającej z kształtowania się wskaźnika TNI, modelu Blacka-Scholesa, Markowitza oraz modelu Sharpe'a. Właściwe wyznaczenie przedziału czasowego skutkuje prawidłową wyceną instrumentu finansowego czy ryzyka. Behawioralny efekt *framingu* jest tłem prowadzonych w artykule rozważań.

Słowa kluczowe: finanse behawioralne, model Blacka-Scholesa, wskaźnik TNI, model Sharpe,a, efekt framingu

APPLICATION OF EPSILON METHOD TO MODELING EXPECTATIONS IN CONSTRUCTION

Natalia Nehrebecka^{1 2}, Sylwia Grodkowska²

¹Katedra Statystyki i Ekonometrii UW, ²Departament Statystyki NBP
e-mail: nnehrebecka@wne.uw.edu.pl, Sylwia.Grodkowska@nbp.pl

Abstract: The epsilon method has been applied to examine the strength of relations among selected objective and subjective factors connected with a manager's predictions of their companies' development. The aim of this research was to study which variable has the strongest impact on business expectations in construction. The results offer compelling evidence that respondents rely both on their current opinion on enterprise as well as on general economic situation. The survey was carried out based on Polish data from 2000:1 to 2008:10.

Keywords: Epsilon method, relative weights, construction

INTRODUCTION

The primary aims of the business tendency survey are forecasting the economic activity, identifying the current economic condition of the country and predicting the direction of changes in the business activity. Business survey indicators are the valuable tool for economic analysis for all types of users. Moreover, this type of survey provides information on assessments and expectations of the economic situation by actors on the market. Conventional methods often prove insufficient in collection of information regarding business tendency surveys. Therefore, it is an useful source of information about changes in aggregated economic activity.

The business tendency survey in construction usually includes: capacity utilization, plans and expectations for the immediate future, managers' views on the overall economic situation¹. The selective key survey variables are used to construct confidence indicator which reflects the general business tendency.

¹ Business Tendency Surveys: A Handbook

This article aims to explaining the mechanism of formulation the respondent's expectation about business tendency in construction. It pursues factors which are closely associated with subjective expectations about future company's development. It investigates whether indices of current company's sentiment have a strong impact on future business tendency indicators. On the other hand, this paper also includes the objective, macroeconomic data, like unemployment or gross wages. The purpose is to check, if the signal from real economy strongly influences managers' prediction of development of their company. The leading hypothesis of the article is that respondents rely on their own evaluation rather than on the general economic situation.

Comparison between these two groups of regressors will be based on a relative weight method. We apply this method to estimate the relative importance of the predictor variable. The method and its usefulness will be presented in the next section. Section 2 discusses the data used in the empirical analysis. Section 3 reports the results which we obtained for the business tendency indicators. Finally, section 4 provides a conclusion.

THE EPSILON METHOD

Econometric modeling is facing a serious problem because of correlation between explanatory variables. The reason why the correlation occurs is interdependence of economic phenomena. In case of strong relationship between independent variables it is hard to separate the particular variable's influence on dependent variable from other variables. This problem makes economic analysis really complicated. It was the reason to search for a measure of relative importance of variables in multiple regression. The history of this research was described by Johnson and LeBreton (2004). Each of methods mentioned (zero - order correlations, standardized regression coefficients, not standardized regression coefficients, semi-partial correlation, the product measure, average squared semi-partial correlation, average squared partial correlation) has its drawbacks which makes this technique useless and leads to wrong conclusions.

These studies resulted in two new important methodologies - dominance analysis and epsilon method. First of these was introduced by Budescu (2003). Dominance analysis allows ranking the predictor variables. It is said that the variable A is dominate to variable B if A it is more useful than B in all subset multiple regressions. Given that the above definition is really strict, it is difficult to implement it for data sets of significant size, because it requires estimating submodels, where is the number of variables in the original model. The advantage of this method is that relative weights obtained from dominance sum to the model. Because of that relative importance of each variable can be calculated as the proportion of predictable variance for which it accounts [Johnson, LeBreton 2004].

This article applies the second approach - the epsilon method. It is much easier to implement than dominance analysis and lead to similar results [Johnson 2000].

The epsilon procedure presented below derives from Johnson (2000). The first step involves the following decomposition of predictor scores matrix:

$$\mathbf{X} = \mathbf{P}\mathbf{\Delta}\mathbf{Q}^T \quad (1)$$

where \mathbf{X} is matrix $n \times p$, \mathbf{P} is the full matrix of eigenvectors of $\mathbf{X}\mathbf{X}^T$, \mathbf{Q} is the full matrix of eigenvectors of $\mathbf{X}^T\mathbf{X}$, T denotes transposition, $\mathbf{\Delta}$ is the a diagonal matrix which contain the square roots of the eigenvalues of $\mathbf{X}\mathbf{X}^T$ and $\mathbf{X}^T\mathbf{X}$.

If none of the vectors from the matrix \mathbf{X} is collinear with another one, each of element of $\mathbf{X}\mathbf{X}^T$ and $\mathbf{X}^T\mathbf{X}$ is different from zero. In these case the best - fitting approximation of \mathbf{X} is:

$$\mathbf{Z} = \mathbf{P}\mathbf{Q}^T \quad (2)$$

The above equation ensures that the residual sum of squares between the raw predictors and the orthogonal variables is minimized. The matrix \mathbf{Z} contains z_k variables which are orthogonal to the original predictors x_j from the matrix \mathbf{X} . The matrix \mathbf{Z} is then used to obtain the set of relative weights. The calculation is based on regression \mathbf{X} on \mathbf{Z} :

$$\mathbf{\Lambda}^* = (\mathbf{Z}^T \mathbf{Z})^{-1} \mathbf{Z}^T \mathbf{X} = (\mathbf{Q}\mathbf{P}^T \mathbf{P}\mathbf{Q}^T)^{-1} \mathbf{Q}\mathbf{P}^T \mathbf{P}\mathbf{\Delta}\mathbf{Q}^T = \mathbf{I}^{-1} \mathbf{Q}\mathbf{\Delta}\mathbf{Q}^T = \mathbf{Q}\mathbf{\Delta}\mathbf{Q}^T \quad (3)$$

As the Z_k are not correlated, the correlation coefficients λ_{jk}^* are equal to the correlation between x_j and z_k . The squared elements of $\mathbf{\Lambda}^*$ defined as λ_{jk}^* are equal to the proportion of predictable variance in x_j accounted for by z_k . The relative contribution x_j to y can be approximated by $\sum_k \lambda_{jk}^{*2} \beta_k^{*2}$, where β_k^* is regression coefficient taken from least square estimation of y to z_k . The coefficient vector β^* is calculated as:

$$\beta^* = (\mathbf{Z}^T \mathbf{Z})^{-1} \mathbf{Z}^T \mathbf{y} = (\mathbf{Q}\mathbf{P}^T \mathbf{P}\mathbf{Q}^T)^{-1} \mathbf{Q}\mathbf{P}^T \mathbf{y} = \mathbf{I}^{-1} \mathbf{Q}\mathbf{P}^T \mathbf{y} = \mathbf{Q}\mathbf{P}^T \mathbf{y} \quad (4)$$

Each z_k is a linear combination of the x_j . The variance in y accounted for by z_k is, therefore, partitioned among the according to the proportion of variance in z_k accounted for by each x_j .

In this algorithm one can obtain the set of relative weights which sum to the model R^2 . Each weight is attributed to the particular raw independent variable and reflects the effect of the variable by itself and in combination with the others. Each is a linear combination of the. The variance is accounted for by is, therefore, partitioned among the according to the proportion of variance in accounted for by each.

THE DATA

The sample period for all results is 2000:1- 2008:10. Data have been provided by the Polish Central Statistical Office. The dependent variables are equal to the following construction expectation: expected general economic situation of the enterprise, expected prices of construction/assembly works carried out by the enterprise, expected employment in construction/assembly operations of the enterprise. These indicators are equal to the difference between weighted percentages of positive (the first option) and negative answers (the third option). Weighting is necessary as the companies which are taking part in the survey vary in size. The weights are based on the value of sold construction assembly production in current process². The questions are presented in table 1. Each year, ca. 2200 construction enterprises of more than 9 employees takes part in the survey.

The set of independent variables which could influence expectations consist of two parts. In the first one there are business tendency indicators of current economic situation in construction. This group includes: construction/assembly works carried out by the enterprise at domestic market, the guaranteed operating period of the enterprise, capacity utilization of the enterprise³. The second group consist of macroeconomic indicators: index of construction and assembly production in constant prices (2000=100), average monthly gross wages excluding payments from profit in construction, dismissals declared by enterprises⁴, average employment in construction, housing loans of Monetary Financial Institution⁵ to households, unemployment rate - registered⁶, dwellings under construction, sale of construction and assembly production, dwellings completed.

These data derived from Polish Central Statistical Office and National Bank of Poland.

The time series used in our analysis were seasonally adjusted using TRAMO/SEATS procedure [Maravall 2005]. It was done due to the fact that in each series the seasonality was detected and in most cases the seasonal component

² http://www.stat.gov.pl/cps/rde/xbcr/gus/PUBL_business_tendency_survey_in_manufacturing_2.pdf

³ in percentages

⁴ Number of persons which will be dismissed declared by enterprises

⁵ Excluding National Bank of Poland

⁶ Calculation based on number of unemployed registered in employment offices

was strong. In this situation using raw data could result in spurious correlation between variables.

Table 1. Business Tendency questionnaire

Question	Possible answers
Expected financial situation of the enterprise	<ul style="list-style-type: none"> – will be better – will remain unchanged – will be worse
Expected prices of construction/assembly works carried out by the enterprise	<ul style="list-style-type: none"> – will be better – will remain unchanged – will be worse
Expected employment in construction/assembly operations of the enterprise	<ul style="list-style-type: none"> – will be better – will remain unchanged – will be worse

Source: Monthly questionnaire in construction, Central Statistical Office, www.stat.gov.pl

The dependent variables used in our study are strongly correlated (absolute value of correlations ranging from 0,06 to 0,99) with an average absolute value correlation of 0,59. Because of strong correlation between variables, regression coefficients couldn't be used to measure of relative weight. It is the reason why in this case we have to apply the epsilon method.

RESULTS

Table 3 presents relative weights calculated for our dependent variables. These weights were rescaled by dividing them by the model R^2 and multiplying by 100, so they are equal to percentage of the model R^2 connected with each predictor.

EXPECTATIONS REGARDING FINANCIAL SITUATION

Results presented in Table 2 indicate that construction/assembly works carried out by the enterprise was accounted for almost 18% of predictable variable of expected financial situation of the enterprise. It is because this indicator denotes the number of orders which will be carried out.

The guaranteed operating period of the enterprise has also quite a high share in R^2 (12,6%) which only supports our feeling that current orders has a dominant role in predicting the financial situation. The longer the guaranteed operating period of the enterprise is the more optimism about chances of financial success the respondents' display.

Table 3. Relative weights

Variable	Expected financial situation of the enterprise	Expected prices of construction/assembly works carried out by the enterprise	Expected employment in construction/assembly operations of the enterprise
Construction/assembly works carried out by the enterprise	17,60%	14,70%	15,70%
Capacity utilization of the enterprise	2,30%	2,50%	2,40%
The guaranteed operating period of the enterprise	12,60%	12,70%	11,40%
Housing loans of MFI's to households	7,00%	7,30%	8,20%
Dwellings completed	0,50%	0,30%	0,70%
Dwellings under construction	8,00%	4,70%	9,20%
Average gross wages excl. payments from profit in construction	7,20%	14,30%	6,90%
Unemployment rate - registered	6,30%	6,90%	7,70%
Dismissals declared	10,10%	10,40%	8,30%
Average paid employment in construction	7,30%	7,30%	9,40%
Sale of construction and assembly production	8,60%	9,60%	7,30%
Index of construction and assembly production	12,50%	9,40%	12,90%

Source: own calculation

The current situation in construction sector (index of construction and assembly production) as well as sold construction and assembly production is also a valuable predictor of the future financial situation.

It is noticeable that data connected with employment are moderately important for respondents. For example, average monthly gross wages excluding payments from profit in construction which accounts for more than 7% of R^2 . It is a direct result of rapid growth in construction during past few years in Poland. In this period the wages in construction went up because of labor shortage. The moderate relative weight of unemployment rate and dismissals declared indicates that respondents follow the situation on the labor market, particularly the future tendency (the relative weight of dismissals declared is higher than the relative weight of unemployment rate).

The number of buildings completed plays no role in explaining the variation of dependent variable, due to the fact that it takes quite a long time to build a new building. On the contrary, respondents take into account the number of dwellings under construction as it is closely related with future price of dwellings. Value of housing loans of MFI's to households plays a similar role. This variable ac-

counts for 7% of R^2 and indicates the demand for new housing (the majority of buyers purchase property on the primary market).

Surprisingly, the relative weight of utilization capacity was below 2%. However, it is possible that enterprise could easily increase or decrease their utilization capacity to adjust it to the current needs. If so, this factor is not essential in explaining the dependent variables' variance.

EXPECTATIONS REGARDING PRICES IN CONSTRUCTION

The result of our survey shows that the expectations regarding prices in construction are similar to the outcome of the previous model. It is not a surprise that the average monthly gross wages excluding payments from profit in construction has almost the highest relative weight in this model. As the level of average wages is rising the company's profit is reduced. The company is unable to avoid these expenses, because it needs to fulfill contractual obligations. The other labor market's variables - unemployment rate and dismissals declared - have even higher influence on expectations regarding prices of construction than expectation regarding financial situation of the enterprise. These variables have a strong influence on the level of wage. Obviously, the wage claims are directly reflected in the prices of construction. On the other hand the influence on financial situation of the enterprise is not so strong - there are a set of other factors which also affect it.

Construction/assembly works carried out by the enterprise have a very similar wage as average monthly gross wages. It is because they are the fixed term contracts so provisions of an agreement cover the prices. On this basis the enterprise can estimate their future revenues and costs.

EXPECTATIONS REGARDING EMPLOYMENT IN CONSTRUCTION

In comparison to models discussed above, the average employment in construction has quite a strong influence on expectations of employment. It is because the high wage level reduces the enterprise's capability to increase employment.

It is not strange that the number of dwellings under construction and the housing loans of MFI's to households occur to be a satisfactory predictor of dependent variable. If dwellings are under construction it is obvious that enterprises need their employees to finish it, so they would not reduce the labor force in near future. The value of housing loans indicates that households are interested in property purchase. This trend has been observed in Poland since 2002, so respondents anticipate that it will remain unchanged.

CONCLUSIONS

For the purpose of this article we have applied the epsilon method in order to describe the weighting strategies used by respondents when making an overall performance evaluation. Although this procedure is valid when the dependent variables are correlated, it is recommended to consider that none of two variables measure the same thing. In this case one of them should be removed as the relative weight will be spread out of them [Johnson, 2000].

Results of our survey show that both the enterprise's evaluation of the current economic activity and the general economic situation in Poland has a strong impact on respondent's expectation about business tendency in construction. The most valuable predictors are construction/assembly works carried out by the enterprise and the guaranteed operating period of the enterprise. The average relative weight were equal to 16,0% and 12,2% respectively. On the other hand, average monthly gross wages excluding payments from profit in construction and index of construction and assembly production, which are included in general economic situation's indicators, have influenced explaining the variance of dependent variables (13,7% and 11,4% respectively).

REFERENCES

- Budescu D. (1993) Dominance analysis: a new approach to the problem of relative importance of predictors in multiple regression, *Psychological Bulletin*, no. 114
- Business Tendency Surveys: A Handbook, ODCE.
- Budżety gospodarstw domowych w 2007 r., Główny Urząd Statystyczny, Warszawa 2008 r.
- Dominitz J., Manski C. (1994) Using Expectations Data to Study Subjective Income Expectations, National Bureau of Economic Research
- Forsells M., Kenny G. (2002) The rationality of consumers' inflation expectations: survey based evidence for euro area, *European Central Bank Working Papers*, n. 163, Frankfurt
- Johnson, J. (2004) Factors Affecting Relative Weights: The influence of Sampling and Measurement Error, *Organizational Research Methods*, Vol. 7, no. 3, 283-299
- Johnson, J., LeBreton J. (2004) History and Use of Relative Importance Indices in Organizational Research, *Organizational Research Methods*, 2004
- Johnson J. (2000), A Heuristic Method for Estimating the Relative Weight of Predictor Variables in Multiple Regression, *Multivariate Behavioral Research*, no. 35
- LeBreton J., Ployhart R., Ladd R. (2004) A Monte Carlo Comparison of Relative Importance Methodologies, *Organizational Research Methods*, Vol. 7, no. 3
- Maravall A. (2005), An Application of the automatic procedure of TRAMO and SEATS; Direct versus Indirect Adjustment, *Banco de España Working Papers* n. 0524

Wall M., Rechtsteiner A., Rocha L. (2003) Singular value decomposition and principal component analysis, in: Berrar D., Dubitzky W., Granzow M., A Practical Approach to Microarray Data Analysis, Springer

Westerhoff F. (2006) Nonlinear expectation formation, endogenous business cycles and stylized facts, Studies in Nonlinear Dynamics and Econometrics, Vol. 10

Zastosowanie metody epsilon do modelowania oczekiwań koniunktury gospodarczej w budownictwie

Streszczenie: Metoda Epsilon została zastosowana do badania siły powiązania między wybranymi obiektywnymi i subiektywnymi czynnikami związanymi ze spodziewaną kondycją dyrektorów przedsiębiorstw w przyszłości. Celem artykułu było zbadanie, która zmienna ma największy wpływ na oczekiwanie koniunktury gospodarczej w budownictwie. Wyniki przedstawiają przekonujące dowody, że respondenci polegają zarówno na własnej opinii sytuacji przedsiębiorstwa, jak również na ogólny stan gospodarczy. Badanie zostało przeprowadzone w oparciu o Polskie dane z okresu 2000:1 do 2008:10.

Słowa kluczowe: metoda Epsilon, wskaźnik relatywnej wagi, budownictwa

NORMOWANIE ZMIENNYCH OPISUJĄCYCH OBIEKTY NIETYPOWE

Kesra Nermend

Instytut Informatyki w Zarządzaniu, US
e-mail: kesra@wneiz.pl

Streszczenie: W wielu badaniach problem stanowią obiekty nietypowe, których cechy opisywane są bardzo dużymi wartościami. Mogą one wpływać w sposób znaczący na wyniki badań z powodu zmniejszania zakresu zmiennych dotyczących obiektów typowych podczas normowania. Niekorzystny wpływ wartości nietypowych można zminimalizować przez wykorzystanie pewnych metod normowania. W artykule zostały przedstawione dwie tego typu metody: standaryzacja z ważonym odchyleniem standardowym oraz unitaryzacja z wartościami progowymi.

Słowa kluczowe: obiekty nietypowe, metody normowania.

WSTĘP

Zmienne biorące udział w tworzeniu miar syntetycznych, czy też wykorzystywane w grupowaniu, często są wyrażone w różnych jednostkach miary, na przykład w osobach na kilometr kwadratowy, czy też złotych. Zmienne takie są nieporównywalne. Zmienne mogą mieć też jednakowe jednostki miary, ale być nieporównywalne ze względu na różny zakres wartości. Przykładem może być średni miesięczny dochód na osobę i średnia kwota wydawana miesięcznie na kulturę (kino, teatr itp.). Zmienne te są porównywalne, jednak ze względu na większe wartości średniego miesięcznego dochodu na osobę ta zmienna będzie dominować. Jej ważność będzie znacząco większa niż ważność średniej kwoty wydawanej na kulturę, w związku z tym druga zmienna nieznacznie tylko wpłynie na wartość miary syntetycznej, czy też wyniki grupowania [Kukuła 2000].

W ogólnej formie formułę normowania cech można przedstawić następująco [Grabiński i in. 1989, Kolenda 2006]:

$$x'_{ij} = \left(\frac{x_i - A_i}{B_i} \right)^p \quad (1)$$

gdzie: x_i – wartość i -tej zmiennej dla j -tego obiektu przed normalizacją,
 x'_{ij} – wartość i -tej zmiennej dla j -tego obiektu po normalizacji, B_i – podstawa
 normalizacyjna i -tej zmiennej ($B_i \neq 0$), A_i, p – parametry.

Pojawienie się w badanym zbiorze obiektów o nietypowo dużych wartościach zmiennych jest poważnym problemem. Obiekty te powodują podczas standaryzacji „zbiecie” pozostałych wartości w pewnym niewielkim zakresie wartości. W przypadku unitaryzacji „zbiecie” to jest jeszcze dużo silniejsze. Efekt ten powoduje, zarówno podczas grupowania jak i tworzenia miar syntetycznych, słabą różniczalność obiektów ze względu na zmienną, w ramach której występują takie obiekty nietypowe. W skrajnym przypadku może dojść do sytuacji, w której dla tej zmiennej możliwe będzie jedynie rozróżnienie obiektów nietypowych od pozostałych, bez możliwości różnicowania obiektów typowych.

STANDARYZACJA ZMIENNYCH DIAGNOSTYCZNYCH

Po standaryzacji obiekty nietypowe uzyskują bardzo dużą wartość tej zmiennej, dla której posiadają nietypową wartość. W konsekwencji, o wartości tej zmiennej decydować będzie głównie wartość miary syntetycznej dla danego obiektu.

Problem obiektów nietypowych w przypadku standaryzacji można nieco złagodzić stosując zamiast odchylenia standardowego ważone odchylenie standardowe [Kozak i in. 2007]:

$$B_i = \sqrt{\frac{\sum_{j=1}^N \left(x_i w_j - \frac{\sum_{k=1}^N x_i w_k}{\sum_{k=1}^N w_k} \right)^2}{\sum_{j=1}^N w_j - 1}} \quad (2)$$

gdzie: w_j, w_k – wartość j -tej (k -tej) wagi.

Odpowiednie zastosowanie wag umożliwia zmniejszenie wpływu obserwacji odstających na wartość odchylenia. Obserwacjom dalekim od średniej nadawane są

wagi mniejsze niż obserwacjom leżącym blisko wartości średniej. W najprostszym przypadku można zastosować system wag zero-jedynkowych:

$$w_j = \begin{cases} 1 & \text{dla } \left| x_j - \bar{x}_i \right| \leq p \\ 0 & \text{dla } \left| x_j - \bar{x}_i \right| > p \end{cases} \quad (3)$$

gdzie: \bar{x}_i – wartość średnia i -tej zmiennej, p – ustalona wartość progowa (próg).

Ten system wag nadaje wagom wartość jeden jeżeli bezwzględna różnica pomiędzy obserwacją a wartością średnią jest mniejsza lub równa od zadanego progu, a wartość zero jeżeli jest większa.

Obiekty nietypowe występują w populacji dość rzadko - dwa albo trzy razy w zbiorze stuelementowym zawierającym zarówno obiekty typowe, jak i nietypowe. Wartość progu p można zatem uzależnić od przewidywanego prawdopodobieństwa wystąpienia obiektów nietypowych. Można na przykład przyjąć, że prawdopodobieństwo wystąpienia obiektu nietypowego jest mniejsze niż 5%. Wówczas, jeżeli założy się, że rozkład wartości zmiennej jest rozkładem normalnym, to wartość progu p można przyjąć jako równą w przybliżeniu dwóm odchyleniom standardowym 2σ . W przypadku ogólnym, kiedy rozkład jest nieznan, ale istnieje dla niego odchylenie standardowe, aby uzyskać podobne prawdopodobieństwo należałoby przyjąć w przybliżeniu $p = 5\sigma$. Należy jednak pamiętać, że zadane prawdopodobieństwo wystąpienia wartości nietypowej jest prawdopodobieństwem największym z możliwych. Jeżeli nieznan rozkład okazałby się rozkładem normalnym, to faktyczne prawdopodobieństwo wyniosłoby tylko około 0,0000006%. Ze względu na to, że sumowanie rozkładów „przybliża” sumowane rozkłady do rozkładu normalnego, większość rozkładów wartości zmiennych jest zbliżonych do rozkładu normalnego, a więc wartość progu z niewielkim błędem można szacować na podstawie założenia o normalności rozkładu. Wartość progu p można, zatem przyjąć jako 2σ .

UNITARYZACJA ZMIENNYCH DIAGNOSTYCZNYCH

Istnieją również metody zmniejszania wpływu obiektów nietypowych na unitaryzację. Przykładem tego typu metody jest metoda korekcji jasności i kontrastu stosowana w niektórych urządzeniach do automatycznego wykonywania odbitek lub też w niektórych aparatach cyfrowych. Przy przetwarzaniu zdjęć występuje podobny problem jak przy normowaniu zmiennych. Przy zapisie zdjęć jest pewien zakres wartości jasności (lub składowych kolorów) najczęściej $\langle 0;255 \rangle$. Jednak zdjęcie zaraz po zarejestrowaniu ma dużo większy zakres wartości, który musi być

zmniejszony do docelowego. W najprostszym przypadku można zastosować unitaryzację zerowaną (w przetwarzaniu obrazów nazywaną normalizacją), której wartości przemnaża się przez docelową wartość maksymalną, czyli najczęściej 255. Jednak pojawia się tu problem wartości nietypowych, które ze względu na dużą liczbę pikseli obrazu pojawiają się zawsze. Powoduje to, że większość wartości skumulowana jest w pewnej niewielkiej części zakresu docelowego. W konsekwencji zdjęcie wydaje się mało kontrastowe. Aby zwiększyć kontrast zdjęcia, w normalizacji używa się nie wartości skrajnych, a dwie wartości progowe, lewą i prawą, ustalane na podstawie specjalnego algorytmu, najczęściej opartego o analizę histogramu.

Podobne rozwiązanie można przyjąć do normowania zmiennych. Jako podstawę normalizacyjną można przyjąć wartości progowe:

$$B_i = x_{L_i} - x_{P_i} \quad (4)$$

gdzie: x_{L_i} – lewa wartość progowa i -tej zmiennej, x_{P_i} – prawa wartość progowa i -tej zmiennej.

Jest to pewna odmiana podstawy normalizacyjnej [Kukuła 2000, Nowak 1990]:

$$B_i = \max_j \left(x_j \right) - \min_j \left(x_j \right) \quad (5)$$

Wartości progowe podobnie jak dla zdjęć można wyznaczyć na podstawie histogramu częstości. Przy czym, w przypadku zmiennych konieczne jest wykorzystanie histogramu względnych częstości liczonego dla przedziałów wartości [Amir 2000]. Histogram taki liczy się dla zadanej z góry liczby przedziałów, bądź zadanej z góry szerokości przedziałów. Przy czym ten pierwszy przypadek jest o tyle wygodniejszy, że można określić, jaka mniej więcej powinna być minimalna liczba przedziałów. Najlepiej gdyby miała ona taką wartość, aby na jeden przedział nie wypadło średnio mniej niż dziesięć obiektów:

$$N_p \leq \frac{N}{10} \quad (6)$$

gdzie: N – liczba obiektów, N_p – liczba przedziałów.

W pierwszym etapie liczenia histogramu częstości wylicza się zakres wartości odejmując od wartości maksymalnej wartość minimalną:

$$zakr = \max_j \left(x_j \right) - \min_j \left(x_j \right) \quad (7)$$

Zakres wartości jest podstawą do wyliczenia szerokości przedziałów:

$$szer = \frac{zakr}{N_p} \quad (8)$$

Szerokość przedziałów umożliwia określenie granic poszczególnych przedziałów:

$$\begin{cases} g_{Lk} = szer(k-1) \\ g_{Pk} = szer k \end{cases} \quad (9)$$

gdzie: g_{Lk} – lewa granica k -tego przedziału, g_{Pk} – prawa granica k -tego przedziału.

Wyliczone w ten sposób granice definiują przedziały, przy czym jeden przedział jest zawsze domknięty dwustronnie, a pozostałe mogą być domknięte lewostronnie:

$$\langle g_{L1}, g_{P1} \rangle, \langle g_{L2}, g_{P2} \rangle, \dots, \langle g_{LN_p}, g_{PN_p} \rangle \quad (10)$$

lub prawostronnie:

$$\langle g_{L1}, g_{P1} \rangle, (g_{L2}, g_{P2}), \dots, (g_{LN_p}, g_{PN_p}) \quad (11)$$

Wybór sposobu domykania przedziałów, gdy liczba obiektów jest duża, nie ma znaczącego wpływu na wynik normowania zmiennych. Dla każdego przedziału określa się liczbę wartości zmiennej należących do tego przedziału. Powstaje w ten sposób histogram częstości. Wartości histogramu częstości zależą od liczby wszystkich obiektów oraz liczby przedziałów. Im więcej jest obiektów, tym większe wartości przyjmuje histogram częstości. Natomiast im jest więcej przedziałów, tym histogram przyjmuje mniejsze wartości. Aby uniezależnić wartości histogramu od liczby obiektów i liczby przedziałów dokonuje się ich przeskalowania:

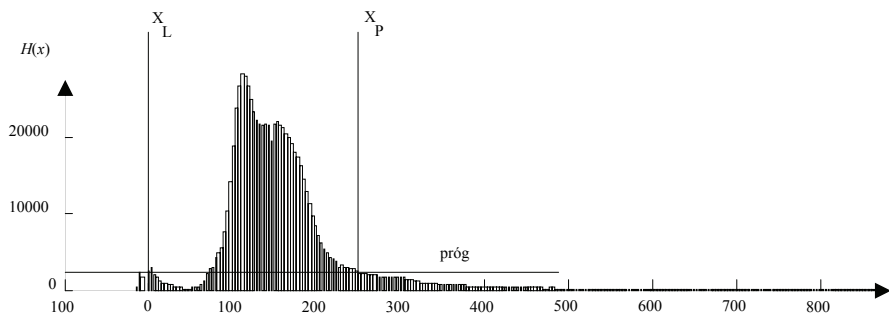
$$h_k = \frac{N_p h_{czk}}{N} \quad (12)$$

gdzie: h_{czk} – k -ty element histogramu częstości, h_k – k -ty element przeskalowanego histogramu częstości.

Przeskalowany histogram częstości jest podstawą do wyliczenia lewej i prawej wartości progowej podstawianej do wzoru (4). Rozłożenie wartości zmiennych charakteryzuje się występowaniem pewnego obszaru skumulowania większości wartości (rys. 1). Poza tym obszarem znajdują się nietypowe wartości zmiennej, występujące dość rzadko. Ze względu na losowy charakter występowania wartości nietypowych można je pominąć przy określaniu lewej i prawej wartości progowej. Lewą i prawą wartość progową można przyjąć za granice skumulowania wartości.

W celu określenia granic można przyjąć lewą i prawą minimalną liczbę elementów znajdujących się w przedziałach należących do skumulowania. Jako lewą granicę przyjmuje się pierwszy z lewej strony przedział, dla którego przekroczona została minimalna liczba elementów lewej strony. Podobnie, jako prawą granicę przyjmuje się pierwszy przedział z prawej strony, dla którego przekroczona została minimalna liczba elementów prawej strony. Ostatecznie wartościami x_{Li} i x_{Pi} będą środki wyznaczonych w ten sposób przedziałów.

Rysunek 1. Wyznaczanie wartości x_{Li} i x_{Pi}



Źródło: obliczenia własne

Wartości x_{Li} i x_{Pi} można także wyznaczać na podstawie odchylenia standardowego:

$$x_{Li} = \bar{x}_i - w_\sigma \sigma_i \quad (13)$$

oraz

$$x_{Pi} = \bar{x}_i + w_\sigma \sigma_i \quad (14)$$

gdzie: σ_i – odchylenie standardowe i -tej zmiennej, w_σ – współczynnik określający krotność odchylenia standardowego.

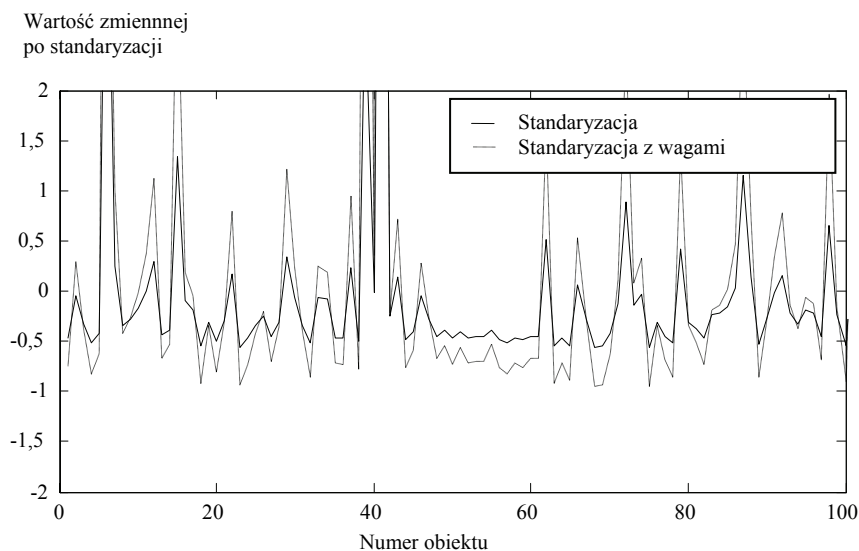
Na ogół wartość współczynnika w_σ przyjmuje się, jako jeden lub dwa. Ta metoda wyznaczania x_{Li} i x_{Pi} jest dużo prostsza od poprzedniej, jednak przy niesymetrycznych histogramach częstości nie gwarantuje, że obie wartości x_{Li} i x_{Pi} będą leżały w zakresie wartości zmiennych.

BADANIA EMPIRYCZNE

Normowaniu poddano wartości emisji zanieczyszczeń pyłowych powietrza z zakładów szczególnie uciążliwych. Badaniu poddano 379 powiatów. Dane zaczerpnięto z GUS-u, dotyczą one 2005 roku. Wskaźnik emisji zanieczyszczeń pyłowych wybrano ze względu na dużą jego rozbieżność w zależności od charakteru zakładów znajdujących się na terenie poszczególnych powiatów. Istnieje wiele powiatów o niewielkiej emisji zanieczyszczeń zbliżonej praktycznie do zera. Są to powiaty z obszarów nieuprzemysłowionych. Istnieje również wiele powiatów znajdujących się w rejonach mocno uprzemysłowionych gdzie emisja zanieczyszczeń jest znaczna. Ponadto istnieje niewielka liczba powiatów, na terenie których znajdują się bardzo duże zakłady przemysłowe, jak na przykład elektrownie. Zakłady te emitują kilkaset razy więcej zanieczyszczeń niż typowe zakłady. Stanowią one obiekty nietypowe zakłócające wartości miar syntetycznych.

Na podstawie wartości wskaźnika emisji zanieczyszczeń pyłowych wyznaczono wartości zmiennej przez podzielenie wartości wskaźnika przez liczbę zarejestrowanych w danym powiecie firm. Otrzymano w ten sposób zmienną: emisja zanieczyszczeń pyłowych na sto firm. Zmienną tą zestandaryzowano z wykorzystaniem zwykłego odchylenia standardowego oraz ważonego odchylenia standardowego. Zestandaryzowane wartości dla pierwszych stu wartości przedstawiono na rys. 2. Wykorzystanie standaryzacji z wagami spowodowało zwiększenie wahań zestandaryzowanej zmiennej.

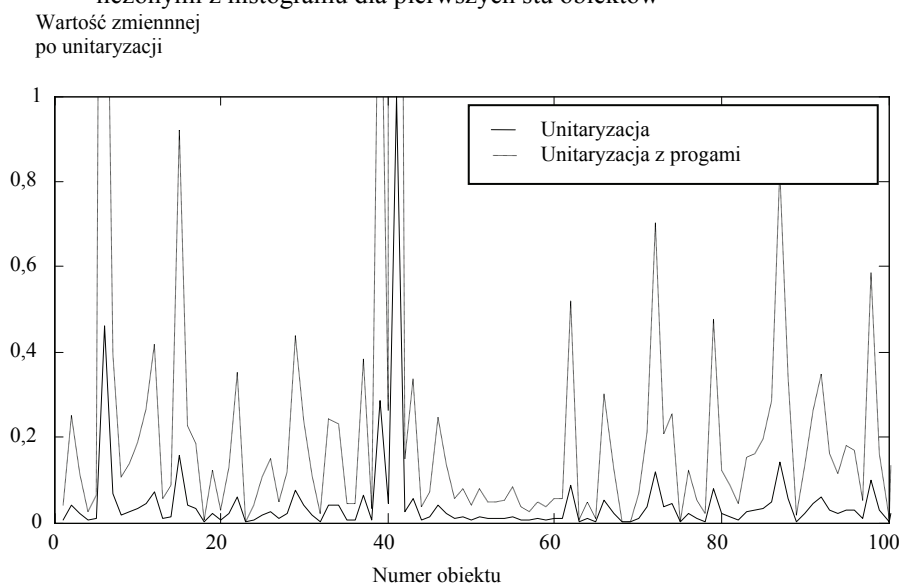
Rysunek 2. Porównanie standaryzacji i standaryzacji z wagami dla pierwszych stu obiektów



Źródło: obliczenia własne

Badaną zmienną podano unitaryzacji zerowanej oraz unitaryzacji z wartościami progowymi (rys. 3). Wartości progowe były wyliczone na podstawie odchyleń standardowych. Przyjęto jako w_σ wartość 1,5. Unitaryzacja z wartościami progowymi spowodowała zwiększenie oscylacji wartości, ale jednocześnie spowodowała znaczne przesunięcie wykresu w górę. To przesunięcie w górę jest spowodowane niesymetrią rozkładu wartości, przy czym w zależności od formy niesymetrii uzyskujemy przesunięcie w górę lub w dół. Te ostatnie jest szczególnie niekorzystne, gdyż pojawiają się wartości ujemne. W przypadku niektórych metod tworzenia miar syntetycznych jest to niedopuszczalne.

Rysunek 3. Porównanie unitaryzacji zerowanej i unitaryzacji z wartościami progowymi liczonymi z histogramu dla pierwszych stu obiektów



Źródło: obliczenia własne

Wady związanej z przesunięciem wykresu nie mają wartości progowe wyznaczone z histogramu. Histogram, który posłużył do wyliczenia wartości progowych miał sto przedziałów. Lewe i prawe progi wyznaczenia wartości progowych były równe i określone zostały na dwadzieścia elementów. W rezultacie uzyskano znaczne zwiększenie wartości wahań zmiennej przy bardzo małym przesunięciu wykresu. Niedogodnością tej metody wyznaczania wartości progowych jest fakt, że konieczna jest dość znaczna liczba obiektów do dokładnego określenia wartości progowych.

PODSUMOWANIE

Przetestowano różne metody eliminacji wpływu obiektów nietypowych na normowanie zmiennych. Wszystkie one powodują zwiększenie wahań wartości, co wpływa na zwiększenie rozróżnialności obiektów typowych ze względu na daną zmienną. W przypadku unitaryzacji, przy małej liczbie obiektów, wartości progowe można wyznaczyć przy pomocy odchylenia standardowego. W tym przypadku dobrze byłoby, gdyby rozkład wartości zmiennej był symetryczny. Przy dużej liczbie obiektów można wyznaczyć wartości progowe z histogramu. W tym przypadku warunek symetrii rozkładu wartości nie jest konieczny.

LITERATURA

- Amir D. A. (2000) Statystyka w zarządzaniu PWN, Warszawa
- Borys T. (1978) Metody normowania cech w statystycznych badaniach porównawczych, Przegląd Statystyczny, nr 2
- Grabiński T., Wydymus S., Zeliaś A. (1989) Metody taksonomii numerycznej w modelowaniu zjawisk społeczno-gospodarczych, PWN, Warszawa
- Kolenda M. (2006) Taksonomia numeryczna. Klasyfikacja, porządkowanie i analiza obiektów wielocechowych, Wydawnictwo Akademii Ekonomicznej im. Oskara Langego we Wrocławiu, Wrocław, ISBN 83-7011-805-4
- Kozak R., Staudhammer C., Watts S. (2007) Introductory Probability and Statistics: Applications for Forestry and Natural Sciences. CABI, ISBN 1845932757
- Kukuła K. (2000) Metoda unitaryzacji zerowanej, PWN, Warszawa, ISBN 83-01-13097-0
- Nermend K. (2008) Rachunek wektorowy w analizie rozwoju regionalnego, Wydawnictwo Naukowe Uniwersytetu Szczecińskiego, Szczecin, ISBN 978-83-7241-660-5
- Nowak E. (1990) Metody taksonomiczne w klasyfikacji obiektów społeczno-gospodarczych, PWE, Warszawa, ISBN 83-208-0689-5

Standardization of Variables Describing Untypical Objects

Abstract: In many investigations the problem of untypical objects, whose characteristics are described by very large values, appears. Such objects may affect significantly the investigations results due to the reduction of the scope of variables in the process of standardization. Negative impact of the untypical values can be minimized by the use of certain methods of standardization. The article presents two such methods: standardization with the weighted standard deviation and unitarization with threshold values.

Key words: untypical objects, standardization methods

IS MULTIPLE LINEAR REGRESSION THE PROPER TOOL OF MODELLING A BEHAVIOUR OF REAL SYSTEMS?

Jacek J. Nowak

Katedra Zarządzania SW im. Bogdana Jańskiego w Warszawie
e-mail: jacek.nowak@janski.edu.pl

Abstract: Methodological assumption that multiple linear regression is an adequate tool of modelling the behaviour of real systems is checked. To do this the experiment is organised on the basis of simple “real” system represented as finite discrete automaton. Main result is that in situation of “black box” modelling the approximation of output variables with multiple linear regressions (from several samples and under different conditions) may not fulfil any of criterions of feasible approximation of systems behaviour, also in situations where real relation between input and output variables is strictly linear and only one of variables is omitted.

Keywords: model of real system, multiple linear regression, real system structure, discrete automaton, „black box” modelling, quality of approximation

INTRODUCTION

After almost 40 years the main postulates of W. Leontief on modelling real economic systems are still actual¹. In present paper we take one of steps on this difficult road of verification of assumptions of modelling real systems.

Multiple linear regression is often used to describe then analyse, forecast or simulate the behaviour of such real systems as economic, environmental, technical and social ones. Most of these models are quantitative models based on statistical data. There still appear failures or unsatisfactory results when applying such quantitative models especially to forecast or simulate behaviour of economic systems. It

¹ We mean especially postulate: “*What is really needed, in most cases, is a very difficult and seldom very neat assessment and verification of these assumptions in terms of observed facts.*” [W. Leontief, 1971, p. 2].

is often pointed out by experienced practitioners and methodologists² that these failures are caused mainly by structural changes in modelled systems³ or by one or more principal factors from listed below:

- a) omitted or misspecified variables,
- b) multicollinearity (near multicollinearity) between explanatory variables,
- c) using bad data, (e.g. inadequate or incomplete or data with errors),
- d) stochastic nature of relations,
- e) series of data containing time trends,
- f) misspecification of the functional form of relations between variables,
- g) inadequate, incomplete (or lack of) theory of the modelled system.

After the period of significant creativity resulting in many ideas of models taking into account structural changes⁴ in recent years the greater attention in researches is paid to testing hypotheses of structural break⁵.

But still the more fundamental problem than hypotheses testing is how to construct the adequate model. The opinion is that if there exists no one of above mentioned reasons of failures of models' applications, excluding lack of or incomplete theory, then the regression model with good ratios of its stochastic structure ought to be proper for representing, forecasting or simulating given real system.

But models of real systems, especially econometric ones, are constructed often in situations of "black box" modelling. It means that the data, which modeller possess, are describing only input and output series but are not describing (internal) states of modelled system.

In [Nowak 2000] the following problem was stated: *Is the multiple regression an adequate tool of modelling and forecasting in "black box" situations under the assumption that there exists no one of above mentioned reasons of failures of models' applications (excluding lack of theory or incomplete theory)?*

To answer the above question there was conducted an experiment of modelling simple system under above assumption. We have generated several samples of observations, estimated parameters of linear regressions with two explanatory (input) variables and approximated endogenous (output) variables within and out of samples (using true, observed values of explanatory variables).

Instead of fulfilling the above assumption⁶ and very high values of determination coefficients, R^2 , the quality of approximation of dependent variables out of samples was very poor. The greater part of regressions produced mostly approxi-

² See e.g. Mayer (1993), Broemeling (1982), Hendry, Richard (1983), survey in Nowak (1981).

³ The main reason of these structural changes is the nature of economic (social) reality (cf Marschak, 1950 or Leontief, 1971).

⁴ See e.g. survey of such ideas in Nowak (2004).

⁵ See e.g. Bai (1999), Elliott and Müller (2006) or Juhl and Xiao (2009).

⁶ We could not avoid only some correlation between explanatory variables in some samples instead of generating their values on the base of random tables.

mation errors⁷ greater than 10% of really observed values of endogenous variables. The remaining regressions produced nearly half of such infeasible errors. There were several errors greater than 100% (!) of approximated values, some of them greater than 500%. The whole sample, consisted of all generated observations, numbered 34 elements. In addition, great variability of estimates of regression parameters was observed (from one sample to another sample).

There exist several practical criterion s of adequacy (feasibility) of approximations (or forecasts). We can take into account one or combination of at least three following criterions:

- 1) no error is greater than 10% of observed (approximated) value (very strong condition),
- 2) no more then 10% of errors is greater than 10% of observed (approximated) values,
- 3) mean of absolute values of relative errors is not greater than 10%.

Results of [Nowak 2000] showed that none of above criterions of feasibility of approximation was fulfilled by any regression from any sample.

The general conclusion of this investigations was that in situation of “black-box” and instead of lack of principal reasons of failures of modelling multiple linear regression is not the adequate tool of modelling behaviour of system.

The goal of present investigations is to check if the above described results were not casual and conclusion will be valid in situations of changed two types of conditions: A) with another initial states and B) quite different sequence of observations. To check this we must repeat the mentioned experiment with two parts according to A and B.

REPRESENTATION OF REAL SYSTEM

To make our experiment we must choose the point of view, and then the way of representation of real systems. This choice is very important from the methodological as well as from the quality of approximation point of view.

It is widely known that the reaction (output) of every real system does not depend only on impulse (input) acting on this system but also depends on a current state in which this system remains. Therefore we accept the way of representation of real systems as systems of inputs, states and outputs as it is accepted within automata theory in cybernetics or general systems theory.

In economic sciences possibly first who applied such representation of real economic systems were Kornai (1971) and Rey (1971) and (1979).

⁷ Approximation error we mean difference between observed and theoretical value of any output variable, where theoretical value is calculated from corresponding regression when “true” (observed) values of explanatory variables were used.

To say more exactly we accept as true the working hypothesis that every real system can be fully and explicitly described (on a proper level of generalisation) by the following quintuple

$$\langle X, Q, Y, \delta, \lambda \rangle, \quad (1)$$

where: X - the input set (containing inputs, x 's, i.e. actions of the environment on the system), Q - the set of states of the system, Y - the output set (enclosing outputs, y 's, i.e. actions of the system on its environment), $\delta: R \rightarrow Q: (x, q) \rightarrow q'$ - the state transition function, where $R \subseteq X \times Q$, $\lambda: R \rightarrow Y: (x, q) \rightarrow y$ - the output function, and X, Q, Y are finite. Formally, the quintuple $\langle X, Q, Y, \delta, \lambda \rangle$ is also a definition of so called discrete, finite Mealy automaton (or machine).⁸

If the "internal" structure (i. e. transition and output functions) of real system is not known the recognition of the way of performing inputs into outputs becomes the main problem of investigations and the "black box" approach can be applied here as an auxiliary, intermediate tool.

ORGANISATION AND PRINCIPLES OF THE EXPERIMENT

The experiment will consist of two phases as in [Nowak 2000].

In phase I we take some "real" system⁹ represented by quintuple (1), i. e. we know everything about its structure (transition and output functions). We fix its structure over the whole experiment – no structural changes. Then we randomly generate¹⁰ sequences of inputs in subsequent steps ("time units")¹¹ observing resulting sequences of states and of outputs of this system. Then we prepare data that are quantitative characteristics of these sequences of inputs and outputs (input and output variables).

In phase II we arrange the situation of the "black box" modelling, i. e. a modeller, as in many cases in reality, does not know (must forget) the structure of the modelled system which is this "real" system from phase I. Then to describe the behaviour of this system the modeller have only data which are quantitative characteristics of input and output elements in subsequent observations (time units)¹².

⁸ Theory of automata treats the automaton as only "an information processor" but the author following Kornai (1971) and Rey (1979) does not impose any constraints on the "nature" of elements of the system under consideration, described by the quintuple (1).

⁹ Despite of this artificially construction of the system used in experiment its structure has all basic features of real systems according to axioms listed in [Nowak 2003].

¹⁰ Random generation of inputs is applied to avoid time trends or periodic fluctuations in data, and, to minimalise probability of multicollinearity (or near multicollinearity) between explanatory variables.

¹¹ One can interpret these subsequent steps as time units but it is not necessary.

¹² We described in detail the "nature" of economic statistical data or the relation between statistical data and input and output elements of real economic systems in Nowak (1985).

Next we assume that the behaviour of this system (which is now limited to behaviour of input variables) can be adequately represented by linear regression between output and input variables as explanatory ones. Then we estimate parameters of such regression for samples of different sizes, approximate values of output (dependent) variables within and out of these sample (using always “true”, observed values of explanatory variables) and check adequacy of these approximations.

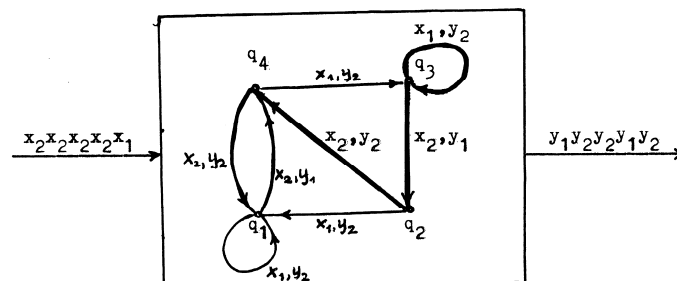
Particular principles of the experiment are as follows. In phase I we take into consideration the same simple system as in [Nowak 2000].

The structure of real system can be presented more conveniently, instead of the quintuple (1) with its algebraic description, as a graph¹³, and in general case, as the directed multigraph. So we present the structure of this system in more convenient form as the graph in fig. 1.

Elements of the graph representing the structure of system have following meaning:

- nodes represent states of system, q 's,
- arrows represent possible transitions of system from state q to state q' (at the arrow-head) - but not vice versa – when input x acts on state q ,
- labels of arrows, pairs (x, y) , denote what output is produced when input x acts on state q (the beginning of arrow);
- first element of label of every arrow shows input x which causes transition from state q to state q' represented by this arrow.

Figure 1. Representation of system G2 with its structure and behaviour in observation no. 2.



Source: Nowak (2000), p. 173.

Elements of a graph representing the structure of system have following meaning:

- nodes represent states of system, q 's,
- arrows represent possible transitions of system from state q to state q' (at the arrow-head) - but not vice versa – when input x acts on state q ,

¹³ By a graph we mean, in general case, a directed multigraph.

- labels of arrows, pairs (x, y) , denote what output is produced when input x acts on state q (the beginning of arrow);
- first element of label of every arrow shows input x which causes transition from state q to state q' represented by this arrow.

Fig. 1 represents an example of the behaviour of system G2, here during observation no. 2 where system's initial state was q_3 .

Principles of generating input sequences using random numbers table¹⁴:

1) watch succeeding digits of random numbers, row by row, from the left to the right, 2) if digit is odd (but other than 9), choose x_1 ; if digit is even, choose x_2 ; if digit is 0, omit it; if digit is 9, it means the end of given sequence of inputs.

It is assumed for simplicity that we use only one-element inputs and outputs. Then measuring quantities of input and output elements is simple and values of input and output variables are the frequencies of occurrence of input or output of given kind in given sequence of inputs or outputs.

If we denote any observation in our experiment as quadruple (y_1', y_2', x_1', x_2') , where symbols y_1', y_2', x_1', x_2' denote values of corresponding output and input variables, i. e. quantities (here: frequencies of occurrence) of, respectively, outputs y_1 and y_2 and inputs x_1 and x_2 in corresponding sequences then the quadruple representing observation no. 2 described in fig. 2 can be registered as $(y_1', y_2', x_1', x_2') = (2, 3, 1, 4)$.

To realise part A of our experiment (changed initial state) we generated (applying above principles) new 34 observations using the same sequences of inputs as generated in [Nowak 2000] but starting from other than q_1 initial states.

To realise part B we generated (applying above principles) quite new 34 observations using quite new sequences of inputs¹⁵ and starting from initial state q_1 .

In phase II we are in "black box" modelling situation. So we "forget" all information on transition and output functions. But we know all input variables the output variables are dependent of and we have all data concerning these input and output variables. Next we estimate (by least squares method) parameters of linear regressions of the form:

$$y_i' = \beta_0 + \beta_1 x_1' + \beta_2 x_2' + \varepsilon \quad (i=1,2), \quad (2)$$

where β_0 , β_1 and β_2 are structural parameters, ε is error term.

We do such estimations for samples of different sizes, which contain observations no. 1-6, 1-12, 13-24, 1-24, 25-29, 30-34 and 1-34. Then we check general goodness of fit, using R^2 , and significance of estimated parameters. The most important is checking the quality (feasibility) of approximation of values of output variables, y_i' s, the principal element of observed behaviour of modelled system.

¹⁴ We used tables published by Martin (1968). In experiment in Nowak (2000) we used random number rows from 1 to 4.

¹⁵ All randomly generated using tables by Martin (1968); rows from 10 to 16.

RESULTS OF NEW EXPERIMENT

Observations we have generated¹⁶ in phase I are contained in table 4 in Appendix. Results of estimation are presented in table 1.

Table 1. Estimates of full regression parameters and adjusted coefficient of determination. R^2 .

Full regression for y_1'					Full regression for y_2'			
Sample	Initial state: q_3 or q_4 ; old inputs				Initial state: q_3 or q_4 ; old inputs.			
	b_0	b_1	b_2	R^2	b_0	b_1	b_2	R^2
1-6	0,107*	0,071*	0,442	0,893	-0,107*	0,929	0,558	0,992
1-12	0,020*	0,072*	0,478	0,933	-0,020*	0,928	0,522	0,992
13-24	-0,052*	0,174*	0,476	0,948	0,052*	0,826	0,524	0,984
1-24	-0,013*	0,084	0,517	0,946	0,013*	1,017	0,572	0,987
25-29	-1,455*	0,214*	0,623*	0,629	1,455*	0,786*	0,377*	0,684
30-34	0,591*	-0,003*	0,429	0,982	-0,591*	1,003	0,571	0,982
1-34	-0,020*	0,077	0,506	0,939	0,020*	0,923	0,494	0,986
Full regression for y_1'					Full regression for y_2'			
Próba	Initial state: q_1 and new inputs				Initial state: q_1 and new inputs			
	b_0	b_1	b_2	R^2	b_0	b_1	b_2	R^2
1-6	-0,176*	0,054*	0,589	0,989	0,176*	0,946	0,411	0,999
1-12	0,024*	0,172*	0,408	0,900	-0,024*	0,828	0,592	0,985
13-24	-0,043*	0,127*	0,482	0,935	0,043*	0,873	0,518	0,989
1-24	0,000*	0,148	0,453	0,934	0,000*	0,852	0,547	0,989
25-29	0,179*	-0,011*	0,640	0,975	-0,179*	1,011	0,360	0,997
30-34	-0,339*	0,197*	0,551*	0,894	0,339*	0,803	0,449	0,979
1-34	-0,051*	0,131	0,485	0,923	0,051*	0,869	0,515	0,988

* means that the corresponding parameter was not statistically significant¹⁷ (level of significance = 0,10); R^2 - adjusted coefficient of determination.

Source: authors calculations.

Because all free terms were statistically insignificant we estimated also parameters of regressions without free term. Instead of insignificance of parameter β_1 in some regressions we did not withdraw variable x_1' from these regressions because they were basic factors the output variables depend of (according to true knowledge of modelled system). It was not surprising that estimates of parameters from different samples were also different. In two small samples we observed also change of sign of estimate of coefficient β_1 .

¹⁶ Observations generated from initial state q_2 for old input sequences were the same as for initial state q_1 .

¹⁷ Under assumptions of normality of distribution and independence of error term (checked by proper statistical tests) and of absence of multicollinearity.

Tables 2 and 3 present results of approximation of output values.

Table 2. Means* of absolute values of relative errors (percent errors).

	Sample							
	1-6	1-12	13-24	1-24	25-29	30-34	1-34	
	Initial state: q_1 ; old inputs							
Full ¹ y_1'	20	21	33	24	148	156	23	n=26
Full ¹ y_2'	14	13	13	11	152	96	12	n=34
Wft ² y_1'	22	21	31	23	96	80	23	n=26
Wft ² y_2'	11	12	14	12	12	13	11	n=34
	Initial state: q_3 or q_4 ; old inputs							
Full ¹ y_1'	16	16	25	19	42	14	18	n=27
Full ¹ y_2'	12	10,4	11	11	37	16	10,2	n=34
Wft ² y_1'	16	16	24	18	18	16	18	n=27
Wft ² y_2'	10,4	10,2	12	11	10,4	10,1	10,4	n=34
	Initial state: q_1 ; new inputs							
Full ¹ y_1'	25	23	20	20	29	25	22	n=30
Full ¹ y_2'	11	13	11	12	17	11	11	n=33
Wft ² y_1'	23	23	20	20	27	26	22	n=30
Wft ² y_2'	12	12	11	12	15	13	11	n=33

¹ Full – full regression ² Wft – regression without free term. * Calculated without errors when observed value was 0.

Source: authors calculations.

Synthetic evaluation of adequacy of approximation of output variables y' s by linear regression models from different samples is presented in table 3 in Appendix.

FINAL REMARKS

Observed variability of regression coefficient cannot be explained with structural changes, because the structure of modelled system was constant over the whole period of our experiment.

According multicollinearity; relatively moderate values coefficients of determination R^2 of regressions between explanatory variables we obtained, from 0,119 (for sample 1-6 and 1-34) through 0,55 to 0,94 (in two small samples, 25-29 and 30-34)¹⁸ for part A of experiment, and most of them from 0,524 to 0,685 (and 0,85 for samples 1-6 and 13-24) in samples for part B, give introductory information that there could be from two to four cases of near multicollinearity. It can be one of reasons why we could obtain some statistically insignificant parameters.

¹⁸ For which inputs were generated in special (different than for first 24 observations), not quite random way.

The first basic result of presented experiment is that instead of 1) very good values of adjusted coefficients of determination R^2 of almost all estimated regressions and 2) almost ideal conditions of modelling such as lack of structural changes, no time trends or errors in data, slight symptoms of multicollinearity (instead of few samples), variables was not misspecified etc. the quality of approximation of output variables with all estimated regressions was poor.

Notice that no one of estimated regressions did not fulfilled in any part of experiment no one of three mentioned in par. 1 criterions of adequacy, also not the most tolerant criterion, i. e. no mean of absolute values of relative errors was less (or equal) 10% (see table 3). Also increasing samples did not lead to substantial improvement of quality of approximations (see table 3).

The next result is that there were not essential differences or improvement in quality of approximation of behaviour of given system when initial states were changed or new input sequences was acting on the system in the same initial state as in first experiment¹⁹. So this new experiment in changed conditions confirms the results obtained in experiment from 2000yr.

It can be done more and more similar experiments with different sets of observations, new, more complex system, many-element inputs and outputs with "prices" of these elements, new modelling tools or more sophisticated use of regression. But if it is done in situation of "black box" modelling analogous to situation of inadequate, incomplete (or lack of) theory of the modelled system the process of modelling is subject to danger of misspecification of the functional form of relations between variables and, first of all, to omit²⁰ or to misspecify variables.

If estimates and approximations are not satisfactory there will be supposed that true relation between variables is nonlinear. But what functional form ought to have such nonlinear relation, if we do not know the structure of real system?

In situation in this experiment the true relation between input and output variables was linear (!) as below²¹:

$$y_1' = x_1' + x_2' - y_2' \quad (3)$$

The second basic result of presented experiment is that in situation of "black box" modelling the approximation of output variables with multiple linear regressions (from several samples and under different conditions) may not fulfil any of criterions of feasible approximation of systems behaviour, also in situations where real relation between input and output variables is strictly linear and only one of variables is omitted.

¹⁹ Only some improvement was in regressions from 5-element samples (25-29 and 30-34), but still with infeasible quality of approximation.

²⁰ We mean variables representing states of system.

²¹ Because of assuming very simple system and one-element inputs and outputs. But relation (3) in general may be no longer true under other assumptions than applied in our simple situation.

But what could be the reason to modeller to include into regression model the other output variable as explanatory variable if he had true information that outputs of type y_1 (therefore output variable y_1') are results caused only by factors named inputs x_1 and x_2 and internal (nonobserved) states?

And what would be the results of assuming and estimating any nonlinear relation between variables if true relation is linear?

REFERENCES

- Broemeling L., ed. (1982), *Structural Change in Econometrics*, "J. of Econometrics" vol.19, no.1, pp. 1-182; "Annals of Applied Econometrics" 1982-2.
- Elliot G., Müller U.K. (2006), *Efficient tests for general persistent variation in regression coefficients*, "Review of Economic Studies", Vol. 73, pp. 907-940.
- Hendry F., Richard J.-F.(1983) *The Econometric Analysis of Economic Time Series*, "Intern. Statist. Review" No. 2, pp. 111-148 (discussion pp. 148-163).
- Juhl T. and Z. Xiao (2009), *Tests for changing mean with monotonic power*, "Journal of Econometrics", Vol. 148, pp. 14-24.
- Kornai J. (1971), *Anti-Equilibrium*. North-Holland, Amsterdam.
- Leontief W. (1971), *Theoretical Assumptions and Nonobserved Facts*, "American Economic Review" Vol. LXI, no.1, pp. 1-7.
- Martin F.F. (1968), *Computer Modeling and Simulation*. Wiley, New York.
- Marschak j. (1950), *Statistical Inference in Economics: An Introduction*, in: "Statistical Inference In Dynamic Economic Models" (ed. T.c. Koopmans). Wiley.
- Mayer T. (1993) *Truth versus Precision on Economics*, Edward Elgar Publ.
- Nowak J.J. (2004), *On econometric approaches to building and verification of simulation model in situations of structural changes* (in Polish), in: "Symulacja systemów gospodarczych. Część II" (red. A. Balcerak i W. Kwaśnicki). Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław 2004 (Prace Naukowe Instytutu Organizacji i Zarządzania Politechniki Wrocławskiej nr 75), pp. 47-63.
- Nowak J.J. (2003), *Minimal structures of real systems and their models in discrete simulation* (in Polish), in: "Symulacja systemów gospodarczych" (red. A. Balcerak, E. Radośniński, B. Mielczarek), Prace Naukowe Instytutu Organizacji i Zarządzania Politechniki Wrocławskiej nr 74, Wyd. Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław 2003; pp. 143-161.
- Nowak J.J. (2000), *On essential limitation of multiple regression as simulation* (in Polish), in: "Symulacja Systemów Gospodarczych. Prace Szkoły Antałowska'2000". Wyd. Wyższej Szkoły Przedsiębiorczości i Zarządzania im. L. Koźmińskiego oraz Politechniki Wrocławskiej, Warszawa-Wrocław, pp. 169-189.
- Nowak J.J. (1985), *On Some Foundations of Economic Model Building*, in: "Summer MIZAR Workshop, 16.06-15-09, Louvain-la-Neuve 1985", "Cahiers des Mathematiques Appliquees aux Sciences Humaines", No. 1, Bruxelles.

Nowak J. (1981), *Evaluation of stability of econometric model and constancy of elements of its structure with application to forecasting* (unpublished doctoral thesis written in Academy of Economics in Cracow, Dept. of Econometrics and Economic Cybernetics – in Polish). Kraków-Warszawa.

Rey-Werszowec K. (1979), *Tentative de construction d'echelons de structures economiques* (paper at the "Colloque Structures Economiques et Econometrie", Lyon, France).

Rey K. (1971), *La planification sur le chemin de l'automatisation*, "La Pologne", No 7 (203) Juillet, pp. 7 and 21.

Czy liniowa regresja wieloraka jest odpowiednim narzędziem modelowania zachowania systemów rzeczywistych?

Streszczenie: W pracy sprawdzono metodologiczne założenie, że liniowa regresja wieloraka jest odpowiednim narzędziem modelowania zachowania systemów rzeczywistych. W tym celu przeprowadzono eksperyment, w którym wzięto pod uwagę prosty system "rzeczywisty" reprezentowany z pomocą dyskretnego automatu skończonego. Główny wynik tego eksperymentu mówi, że aproksymacja zmiennych wyjściowych z pomocą liniowej regresji wielorakiej (w sytuacji modelowania "czarnej skrzynki", na podstawie szeregu prób i w różnych warunkach) może nie spełniać żadnego z kryteriów dopuszczalnej aproksymacji zachowania modelowanego systemu, również w sytuacjach, gdy rzeczywista relacja pomiędzy zmiennymi wyjściowymi a zmiennymi wejściowymi jest dokładnie liniowa oraz gdy tylko jedna ze zmiennych jest pominięta.

Słowa kluczowe: automaty dyskretnie, model systemu rzeczywistego, liniowa regresja wieloraka

APPENDIX

TABLE 3. Quality of approximations of output variables with linear regression models.

Sample (Numbers of observations)	Regression for y_1 '		Regression for y_2 '	
	Full	Without free term	Full	Without free term
	Initial state q_3 or q_4 and old inputs			
1-6	- 14	- 20	+ 16	+ 12
1-12	- 15	-19	+ 9	+ 10
13-24	-17 ₁	-20 ₂	+ 14	- 23
1-24	- 20 ₂	- 16 ₁	+ 8	+ 13
25-29	- 20 ₃	- 26 ₁₀	- 21 ₄	+ 13
30-34	+ 13	- 26 ₂	+ 15 ₁	+ 15
1-34	- 17	- 17	+ 10	+ 12
	Initial state q_1 and new inputs			
1-6	- 18 ₁	- 17	+ 11	+ 10
1-12	- 23	- 23	+ 13	+ 13
13-24	- 17	- 17	+ 12	+ 12
1-24	- 17	- 17	+ 12	+ 12
25-29	- 21 ₃	- 21	- 18	+ 15
30-34	- 20 ₁	- 20 ₁	+10	+ 15
1-34	- 17	- 17	+ 12	+ 12

Meaning of signs and numbers in the table:

“- n“ means that majority of approximation errors were greater than 10% of really observed values of endogenous variables, and n is the number of such errors,

“+ n” means that most of approximation errors were not greater than 10%, and n is the number of errors greater than 10% of really observed values,

number appearing as subscript means number of errors greater than 100%,

“full” means linear regression with free term and two explanatory (input) variables.

Source: authors calculations.

TABLE 4. Observations for system G2 generated in new conditions of experiment.

No. of observ- ation	Initial state: q_3 or q_4				Initial state: q_1			
	„Old” input sequences				New input sequences			
	Quantities of inputs		Quantities of outputs		Quantities of inputs		Quantities of outputs	
	y_1'	y_2'	x_1'	x_2'	y_1'	y_2'	x_1'	x_2'
1	4	15	12	7	4	13	11	6
2	2	3	1	4	2	4	3	3
3	0	2	2	0	5	13	10	8
4	1	2	2	1	1	4	3	2
5	1	3	1	3	1	1	0	2
6	2	3	2	3	1	2	1	2
7	0	1	0	1	4	10	8	6
8	4	9	6	7	0	2	2	0
9	1	3	3	1	4	6	5	5
10	0	2	2	0	1	4	4	1
11	1	1	1	1	3	8	8	3
12	0	1	1	0	3	9	5	7
13	1	2	1	2	2	2	0	4
14	2	4	2	4	4	7	4	7
15	0	1	1	0	5	9	6	8
16	2	2	2	2	0	2	0	2
17	6	13	7	12	1	0	1	0
18	2	5	4	3	0	1	0	1
19	2	2	1	3	3	2	2	3
20	0	4	4	0	1	5	3	3
21	2	3	1	4	10	26	20	16
22	5	9	6	8	1	3	2	2
23	0	1	1	0	1	3	1	3
24	7	12	9	10	5	9	5	9
25	3	9	6	6	0	2	2	0
26	4	6	1	9	3	13	11	5
27	1	10	10	1	4	15	13	6
28	5	7	3	9	4	7	5	6
29	4	7	6	5	1	5	5	1
30	1	9	9	1	1	2	2	1
31	3	9	6	6	6	15	14	7
32	3	8	5	6	5	8	7	6
33	5	6	1	10	2	3	1	4
34	2	10	9	3	1	4	2	3

Source: authors calculations.

ANALYSIS OF ECONOMIC ACTIVITY MOVEMENTS IN THE CZECH REPUBLIC – FREQUENCY APPROACH

Poměnková Jitka
Wydział Finansów, MUAF
e-mail: pomenka@mendelu.cz

Abstract The primary purpose of this work is to identify the business cycle types in the Czech Republic. In the light of business cycles problematic, and viewed from the frequency analysis perspective, harmonic analysis will be used. From the range of de-trending techniques, the first order difference, linear filtering, unobserved component model and Hodrick-Prescott filter are used. In the case of Hodrick-Prescott filter, a cyclical fluctuation estimate with the derivation of smoothing parameter designed specially for the Czech Republic case is investigated. The aim is to distinguish types of cyclical fluctuations in the Czech Republic. A consequent analysis of potential sources of cyclical movement is done.

Keywords: harmonic analysis, deterministic trends, stochastic cycles, economic activity, business cycle

INTRODUCTION

In many studies, the economic activity of a country is discussed in respect to the business cycle theory. Recall for example Canova (1998, 1999) business cycle study from de-trending and dating turning-points points of view; also King and Rebelo (1993) using the Hodrick – Prescott (HP) filter in time and frequency domain or Harding and Pagan (2006) with the description of different measuring ways of business cycle. There is a variety of approaches to de-trending methods, approximations between filtering methods and effects of used techniques on obtained results. Nevertheless, the empirical analysis results are different from the data approach results. The correctness of the business cycle identification consists of stabilization function of monetary and fiscal policy.

Most of these studies work with advanced market economy, like the US economy having at disposal a large data sample size. The case of the Czech Republic is different. The type of economy in this country is denoted as transition because the economy was transformed from a central planned economy into a

market economy with shocks, structural breaks and with a small sample size of available data.

Notice that a series of works are motivated by the European Union enlargement when countries study their transition economies, separately or in comparison with Euroarea, comovement as well as aspects of evolution of domestic economy. Recall for example the work of Artis & all (2004) or Szapáry and Darvas (2008). The work of Wozniak and Paczinski (2007) using the finite order AR model for the calculation of time series spectra belongs to the topical (contemporary) pieces. Furthermore, Hallet and Richter (2004) examined the time-frequency framework in a way which allows to accommodate structural breaks and non-stationary variables.

Thus, this paper focuses on the business cycle in the Czech Republic from the frequency domain point of view using harmonic analysis for identifying possible types of cycles. Available data are the quarterly values of Gross Domestic Product (GDP) 1996/Q1 – 2008/Q3. Consequently, the analysis of potential sources of cyclical movement is done. The identification of sources of cyclical movements can significantly contribute to the effectiveness of economic policy during the economic activity stabilization.

EXPERIMENTAL PROCEDURES

The standard theory of Burns and Mitchel (1946) business cycle is used: “Business cycles are a type of fluctuations found in the aggregate economic activity of nations that organise their work mainly in business enterprises: a cycle consists of expansion occurring at the same time in many economic activities, followed by similarly general recessions, contractions, and revivals which merge into the expansion phase of the next cycle; in duration business cycles vary from more than one year to ten to twelve years; they are not divisible into shorter cycles of similar character with amplitudes approximating their own.” There exist two approaches to the business cycle. The first one is called the classical business cycle and is based on a fluctuation in the level. The second one is known as the growth cycle and is based on a fluctuation around a trend. The dating of cycles is usually done by identifying turning-points. The time period is usually too short; full of structural breaks and shocks. On the other hand, the growth cycle, where turning-points are characterized by changes relative to the trend, represents a more promising version of the business cycles (Artis & col, 2004). And, the dating procedure of the growth cycle is less sensitive to the growth trend than that of the classical cycle.

In the field of business cycle, cyclical movements of time series which exhibit trends and cycles are studied. If the models are designed to characterize the cyclical behavior, the trends are eliminated prior to the analysis. It is often useful to

build a model where both trend and cyclical behavior are in the model and eliminate the trend from the model and from the actual data in parallel fashion. The specification of the model is subject to the constraint that it must successfully characterize the trend behavior. Having satisfied the constraint, trends are eliminated appropriately and the analysis proceeds with an investigation of cyclical behavior. The isolation of cycles is closely related to the trends removal. Indeed, for time series exhibiting cyclical deviations around a trend, the identification of the trend automatically provides identifying the cyclical deviations as well. Notice that the removal of the trend will leave such fluctuation intact, and its presence can have a detrimental impact on inferences involving business cycle behavior.

In this paper, economic activity is measured by absolute values of GDP transformed to the natural logarithms and denoted as Y_t . For the analysis of business cycle, the additive decomposition method in the form

$$Y_t = g_t + c_t, \quad t=1, \dots, n, \quad (1)$$

is used. There are three main approaches to removing the time trend from macroeconomic time series. The first two approaches to the trend removal are de-trending and differencing. De-trending is usually accomplished by fitting the linear trend to log input values using the ordinary least square method (OLS). The third approach to de-trending involves the use of filters designed to separate the trend from the cycle, but given admission of slowly evolving trend. We use the Hodrick – Prescott filter, which has proven quite popular in business cycle applications. Thus, the set of chosen de-trending method is the following: the first order difference (FOD), autoregression with/without a constant (AR), regression analysis (linear and quadratic trend) and Hodrick – Prescott filter (HP filter).

In our paper, we are not discussing whether these methods are “the best” for the trend removing, or which one is more or less suitable and why. We just take some of standard methods, for example Canova (1998), Baxter and King (1999), Hodrick and Prescott (1980) or books Mills (2003) or Dejong, Chetan (2007), and search for statistically significant periods. The use of several methods for de-trending provides a robust periodicity analysis in the sense of cyclical movement source identification. As calculating the FOD as well as the estimating process of random walk with/without a drift and regression curve using the OLS are well known, we are not going to discuss it in detail. In the case of resultant deterministic models, a quadratic model and the AR (with or without a constant) were chosen according to the standard statistical techniques for the quality model evaluation

The procedure of HP filter was first introduced by Hodrick and Prescott in 1980 in the context of estimating business cycles. The HP filter decomposes Y_t (macroeconomic time series) into a nonstationary trend g_t (growth component) and a stationary residual component c_t (cyclical component)

$$Y_t = g_t + c_t, \quad t=1, \dots, n,$$

where g_t and c_t are unobservables. The measure of the g_t path smoothness is the sum of the squares of its second difference. The application of HP filter involves minimizing the variance of cyclical component subject to a penalty for the variance in the second difference of growth component

$$\min_{\{g_t\}_{t=1}^T} \sum_{t=1}^T (Y_t - g_t)^2 + \lambda \sum_{t=1}^T [(g_{t+1} - g_t) - (g_t - g_{t-1})]^2,$$

where $c_t = Y_t - g_t$. The parameter λ is a positive number which penalizes variability in the growth component series. As Harvey and Jager (1993) have shown, the infinite sample version of the HP filter can be rationalized as the optimal linear filter of the trend component. King and Rebelo (1993) find that the business cycle component and the second difference of the growth component must have the same moving average representation for the HP filter to be a linear filter which minimizes the mean squares error. That is, this filter minimizes the mean square error

$$MSE = (1/T) \sum_{t=1}^T (ets(c_t) - c_t)^2,$$

where c_t is the true cyclical component and $ets(c_t)$ is its estimate. Hodrick and Prescott found that if the cyclical component c_t and the second difference of growth component g_t ($\Delta^2 g_t$) are identically and independently distributed normal variables, $c_t \sim N(0, \sigma_c^2)$, $\Delta^2 g_t \sim N(0, \sigma_g^2)$, then the best choice in the sense of MSE for the smoothing parameter is $\lambda = \sigma_c^2 / \sigma_g^2$. In many papers, the recommended lambda value for the quarterly data is $\lambda=1600$, see Ahumada, Garegnani (1990), Guy, St-Amant (1997), Hodrick-Prescott (1980) and others. On the basis of this let us formulate the rule which helps us to find the optimal smoothing parameter.

Let us have an input time series of positive values Y_t , $t=1, \dots, n$, a set of lambdas $L = [100/n; 100 \cdot n]$ and an indexed set $I = [1, \dots, length(L)]$. Thus, for every $L_i \in L$, $i \in I$ we can calculate the Hodrick – Prescott estimate of growth and cyclical components of Y_t . Let the cyclical component c_t and the second difference of growth component g_t ($\Delta^2 g_t$) be identically and independently distributed normal variables, $c_t \sim N(0, \sigma_c^2)$, $\Delta^2 g_t \sim N(0, \sigma_g^2)$. Thereafter, the optimum value of smoothing parameter λ is such that

$$\lambda_{opt} = \sigma_c^2 / \sigma_g^2 \Leftrightarrow L_i \approx \sigma_c^2 / \sigma_g^2, \quad \text{for any } i \in I$$

where the difference $d = |L_i - \sigma_c^2 / \sigma_g^2| \leq 100/n$.

For the periodicity analysis in the time series, the following harmonic analysis was used. After removing the trend T_t , residuals were obtained, i.e.

$$e_t = Y_t - T_t, t = 1, \dots, n.$$

For the analysis of random sequence in the form

$$e_t = \mu + \sum_{j=1}^{n/2} (a_j \cos(\omega_j t) + b_j \sin(\omega_j t)) + \varepsilon_t, t = 1, \dots, n,$$

where μ , a_j , b_j and ω_j ($0 < \omega_j \leq \pi$) are unknown parameters, e_t is stationary process, a periodogram is usually used. In points $\omega_1, \dots, \omega_r$, the periodogram constructed for a given sequence of realization has relatively big values, and thus makes it possible to find estimates of the parameters $\omega_1, \dots, \omega_r$ including the value r (the corresponding statistical procedure is called the Fishers test of periodicity and it is discussed later in the text). The coefficients a_j , b_j are estimated using the ordinary least squares method (OLS) in the standard way; they represent a regression parameter for the j -th smoothing sinusoida, resp. cosinusoida, and the following formulas derived on the basis of OLS hold for their calculation

$$a_j = \frac{n}{2} \sum_{t=1}^n e_t \sin(\omega_j t), \quad b_j = \frac{n}{2} \sum_{t=1}^n e_t \cos(\omega_j t), \quad j = 1, \dots, n/2.$$

Next, the variability corresponding to the j -th smoothing sinusoida and cosinusoida is calculated according to the formula

$$\text{var}_j = 1/2(a_j^2 + b_j^2).$$

Testing the statistical significance of all possible periods and connected theoretical variance is done by the Fisher test (Anděl, 1976). As it was mentioned above, it is suitable to use some stationary test; therefore we propose the augmented Dickey-Fuller (ADF) test (Wooldridge, 2003).

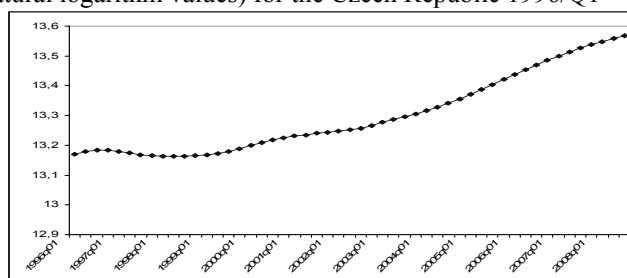
RESULTS

The input data set for the economic activity measurement in this paper are the values of Gross Domestic Product (GDP) in the Czech Republic. The available data are quarterly values 1996/Q1 – 2008/Q3. Notice that the GDP values in absolute form (y_t) had been transformed into natural logarithms and denoted as Y_t (figure 1) before the analysis.

In the first step of empirical analysis, the stationary test of input values was done. According to the ADF test, the logarithmic values of the Czech Republic

GDP are trend stationary on 5% significance level with lag $p=2$. We can permit that the effects of random disturbance disappear gradually over time because our variable grows along a trend. The chosen methods for de-trending were the first order difference (FOD, e_1), autoregression with/without a constant (AR, e_2 ; ARc, e_3), regression analysis (linear and quadratic time trend; e_4) and Hodrick – Prescott filter (HP filter, e_5 , e_6). Table I describes the OLS estimates of some models. In the case of deterministic model, the quadratic regression gives better results than the linear regression, thus this type of regression is used in the next analysis. The estimates of AR model without a constant as well as AR model with a constant have significant parameters and good statistics. Hence, both cases were involved in the analysis of the periods. The Hodrick – Prescott filter is calculated for two types of smoothing parameter λ . At first $\lambda = 1600$ (e_5) on a regular basis value for the quarterly data is taken. Secondly it is taken as a ratio of variances calculated from values for the Czech Republic $\lambda_{opt} = \sigma_c^2 / \sigma_g^2 = L = 115$ (e_6).

Fig. 1: GDP (natural logarithm values) for the Czech Republic 1996/Q1 – 2008/Q3



Source: Own investigation

Before applying harmonic analysis, the residuals obtained by the application of de-trending methods are tested for the zero mean stationarity (ADF1). Table II shows the results. Figure 2 shows the residuals in correspondence to the used de-trending method. The notation means the "name of detrending method, corresponding residuals", for example "FOD, e_1 ".

The basic assumption for the FOD procedure applications are that the trend component of the analysed time series is a random walk with no drift and that the cyclical component is stationary and both components are uncorrelated (Canova, 1998). Thus, $g_t = Y_{t-1}$, $t=1, \dots, T-1$, should be a random walk with no drift in the Czech Republic case. This condition can be taken as satisfied even if we know that a random walk with a drift gives statistically better results. But the stationarity test (ADF1) for the cyclical component (Table II, denotation FOD, e_1) showed a non-stationarity (even in the case of stationarity around a constant). For this reason the FOD cannot be taken as a de-trending method for the Czech Republic. From the

economic point of view, the values of ΔY_t corresponds to the classical business cycle concept more than the growth type of cycle. Even though the FOD does not give stationary residuals, we do the periodicity test to see the results for the classical business cycle.

Table I. Estimate of quadratic regression model, AR model with/without a constant for the Czech Republic

	Parameter estimate	SE	t-value	F-test	p-value	R_{adj}^2	n
Quadratic regression				4609,24	$1,4 \cdot 10^{-55}$	0,9946	51
const.	13,1757	0,0042	3127,1802		$4 \cdot 10^{-129}$		
t	-0,0023	0,0004	-6,0495		$2,1 \cdot 10^{-7}$		
t^2	0,0002	$7 \cdot 10^{-6}$	29,3705		$2,5 \cdot 10^{-32}$		
AR const				4280,1	$1,7 \cdot 10^{-72}$	0,9988	50
const.	-0,4527	0,0665	-6,8089		$1,4 \cdot 10^{-8}$		
Y_{t-1}	1,0346	0,0050	206,8868		$1,7 \cdot 10^{-72}$		
AR				$2,3 \cdot 10^8$	$4 \cdot 10^{-162}$	0,9796	50
Y_{t-1}	1,0006	$6,6 \cdot 10^{-5}$	15273,9762		$3 \cdot 10^{-165}$		

Statistical significance at the 1% (***), 5% (**), 10% (*)

Source: Own calculation

Table 2. ADF stationarity test results for residuals obtained using corresponding de-trending method

	FOD, e_1	quadratic, e_2	AR cons., e_3	AR, e_4	HP1600, e_5	HP115, e_6
ADF1	No	Yes, ***	Yes, ***	Yes, **	Yes, ***	Yes, ***
Lag		2	3	3	2	2

Statistical significance at the 1% (***), 5% (**), 10% (*) H_0 : non-stationarity of e_t , t-stat. > quantil, (No) H_1 : stationarity of e_t , t-stat. < quantil, (Yes)

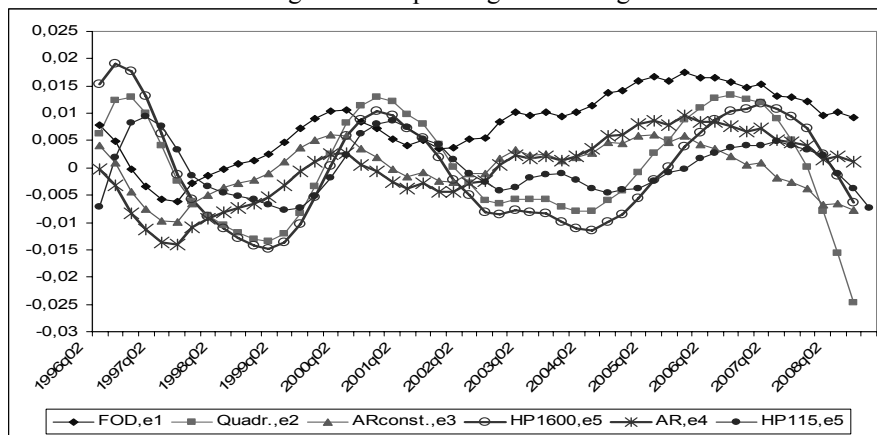
Source: Own calculation

The unobserved component (UC) model (Canova, 1998) for de-trending in the Czech Republic case is mentioned above as a random walk process with a drift and with a stationary cyclical component of finite order AR process type. Table I shows statistically significant values for the estimated model, Table II presents stationarity results for the corresponding cyclical components (AR cons., e_3). This method is, therefore, suitable. However, comparing the FOD and AR residuals shape and AR cons. residuals in Figure 2 we can see a similarity in the second half

of the period of the curve shape with the FOD residuals at first sight. Thus, even if the assumptions of UC method are satisfied, it was decided not to recommend this method as de-trending for the Czech Republic case.

In following step, the application of harmonic analysis for the identification of possible types of cycles on all given residuals is done. Notice that one value is lost if we use de-trending by the AR and FOD. The results of calculated periodicity and testing statistical significance are given in Table III. Two lengths of sample size and two corresponding periods for the used de-trending methods are caused by losing one value in the case of FOD and AR models. This difference is quite small, thus the periods in one column are considered as similar.

Fig. 2: Residuals obtained using the corresponding de-trending method



Source: Own investigation

Table 3. Results of statistical significance of periods for the residuals obtained using the corresponding de-trending method

periods	50	25	16,6667	12,5	10	8,33333	7,14286	6,375
FOD,e1		**						
ARc,e3	***	***	***	***	**			
AR,e4	***	***		**				
periods	51	25,5	17	12,75	10,2	8,5	7,28571	6,375
HP1600,e5	***	***	***	***	**	***	***	
HP115,e6		***	***	***		***	***	***
Quadrat.,e2		***	***	***				

Statistical significance at the 1% (***), 5% (**), 10% (*)

Source: Own calculation

As discussed above, the FOD technique and its residuals do not satisfy assumptions for the application and taking it as a de-trending method does not give a reasonable result. For economic reasons, these values correspond better to the classical type of business cycle theory, thus, we can use the results for considering the periodicity in the classical business cycle. If the analysis of Figure 2 is done, we can see some similarity in the FOD and AR residuals trending. In addition, the FOD residuals are not stationary, unlike the AR residuals. Thereby the periodicity analysis of AR residuals will be taken as confirmative for the FOD residuals.

DISCUSSION OF RESULTS

We can distinguish several types of cycles (Burda, Wyplosz; 2001) known as the Kitchin inventory cycles of 3-5 years, the Juglar fixed investment cycles of middle length 7-11 years and the Kuznets infrastructural investment cycles of 15-25 years. The longest type, the Kondratieff waves of long technological cycle of 45-60 years are not considered as the available sample size for the Czech Republic is small.

The success of dating growth cycles crucially depends on the quality of the approximation of trend component. In our case of the Hodrick-Prescott filter, a random walk with a drift and quadratic de-trending helped to identify the growth cycle, while the FOD and AR were used for the classical business cycle dating. All residuals time series denoted 6.25 years as a significant period. This cycle duration is of the Juglar type caused by fixed investment.

The periodicity of FOD residuals showed only 25 periods as significant, i.e. 6.25 years periodicity. More significant periods were found when we used the AR residuals. In the other cases the residuals describe the growth business cycle. Comparing all of them, the cycles of length 6.25, 4.25 and 3.19 years were detected. Additional cycles, shorter ones, with the cycle duration of 2,125-1,6 years, were detected when the Hodrick-Prescott filter was used. For the special λ , derived for the Czech Republic, the Hodrick-Prescott also identified the longest period of 12,75 years. For the ARc residuals, the results show a similar statistical significance period as the quadratic and Hodrick-Prescott filters. All these three cyclical components agreed on three kinds of periodicity, namely 6.375, 4.25 and 3.18 years. The comparison of classical and growth business cycle concepts validates the fact (Bonenkamp, 2001) that the classical cycle peaks come later in time than the growth cycle peaks also for the Czech Republic case.

Hence, the basic types of cycle can be found in the Czech economy evolution - the shortest Kitchin cycles caused by the inventories as well as the middle one Juglar cycles caused by the fixed investment. On the basis of harmonic analysis, the results completed by the economic research of cycle correlation

analysis were done. For the classical BC as well as for the growth BC, the Kitchin inventory and the Juglar fixed investment cycles were identified, and so were the Kuznets infrastructural cycles. In the case of the growth BC the short type of cycle was identified.

Table IV. Correlation coefficients between residuals and economic factors

Correlation	final consumption expenditure	final consumption expenditure of households	final consumption expenditure of general government	gross capital formation	gross fixed capital formation	changes in inventories	exports of goods and services	imports of goods and services	net exports of goods and services
Quadrat.,e2	0,1687	0,3039 **	-0,0973	0,4655 ***	0,4077 ***	0,1837	0,3444 ***	0,4638 ***	0,0590
ARc,e3	0,1278	0,2204	-0,0581	0,3645 **	0,2440 *	-0,2359	0,2643 *	0,2397	-0,1251
HP1600,e5	0,1945	0,3313 **	-0,0875	0,3981 ***	0,4713 ***	0,2445 *	0,3210 **	0,4283 ***	0,1913
HP115,e6	0,2036	0,2987 **	-0,0233	0,4618 ***	0,4372 **	0,2498 *	0,3351 **	0,5107 ***	0,1072

Statistical significance at the 1% (***), 5% (**), 10% (*)

Source: Own calculation

Hence, the economic factors possibly influencing cyclical fluctuation in the Czech Republic based on the data availability were chosen – the final consumption expenditure, final consumption expenditure of households, final consumption expenditure of general government, gross capital formation, gross fixed capital formation changes in inventories, exports of goods and services, imports of goods and services, net exports of goods and services in the period 1996/Q1 - 2008/Q3 as the percentage changed compared to the corresponding previous period. The aim of the study is to measure the tightness relation, i. e. the correlation coefficients between economic factors and de-trended time series (Table IV). Thereafter the source of cyclical movements in the Kitchin sense caused by the inventory was statistically significant only on 10% level for one de-trending method. This dependence is rather weak for the correlation coefficient result ($r = 0,24$). In the Czech Republic, the cyclical movements of short type are influenced more by the consumption of households; the correlation showed significant results for several de-trending time series on 5% significance level. Further, the source of cyclical movements in the Juglar sense was highly significant for the gross capital formation (the correlation from 0.36 to 0.46) and the gross fixed capital formation

(the correlation from 0.40 to 0.46) and it can be denoted as the middle one. The question is which type of cyclical movements is influenced by the export and import in the Czech Republic. The results of correlation showed a high statistical significance for the middle level of correlation.

CONCLUSION

This paper has focused on the business cycle in the Czech Republic from the frequency domain point of view using harmonic analysis for the identification of possible types of cycles. Deterministic as well as stochastic methods were applied for obtaining the cyclical fluctuation. According to the empirical results, two approaches to business cycle are discussed, the classical and growth types. Consequently, the analysis of potential sources of cyclical movement is done.

The chosen methods for de-trending were the first order difference, autoregression with/without a constant, quadratic time trend and Hodrick – Prescott filter. In the case of Hodrick – Prescott filter, two types of smoothing parameter were used ($\lambda=1600$ and $\lambda=115$). The success of dating growth cycles crucially depends on the quality of the approximation of trend component. In our case, the Hodrick-Prescott filter, random walk with a drift and quadratic de-trending helped to identify the growth cycle, while the first order difference and autoregression process without a constant were used for the classical business cycle dating. Next, harmonic analysis for the identification of possible types of cycles in all given residuals (including the first order difference and random walk process residuals) was done. All residuals time series denoted the period of 6.25 years as significant. This cycle duration is of the Juglar type caused by the fixed investment.

As mentioned in the introduction, the Czech Republic post-transforming economy disposes of a limited sample size of available data. Even if we study the classical BC concept, this length gives one possible period with respect to the shocks, structural breaks and development of the country. The growth business cycle reacts to this feature better and is less sensitive to the trend. Thereafter, the application of HP filter, random walk with a drift and quadratic regression detected a wide scale of significant periods. From the longest one (12.5 years), through the medium (around 6 years), to the short one (between 2-4 years). In both cycle concepts, the same effects of nested loops took place.

Thus, the basic type of cycles in the Czech economy evolution were identified, the shortest Kitchin cycle caused by the inventories as well as the middle Juglar cycle caused by the fixed investment. In case of growth BC the short type of cycle was also identified. Hence, the economic factors possibly influencing the cyclical fluctuation in the Czech Republic based on data availability were chosen and the correlation between the cyclical movements and economic factors was investigated.

Thereafter, the source of cyclical movements in Kitchin sense caused by inventory was statistically significant only on 10% level for one de-trending method and this dependence is rather weak. The cyclical movements of the short type cycles are influenced more by the consumption of households; the correlation in this case showed significant results for more de-trending time series. Further, the source of cyclical movements in the Juglar sense was highly significant for the gross capital formation and gross fixed capital formation and it can be denoted as middle one.

ACKNOWLEDGEMENTS

The results introduced in the paper are supported by the research intent n. MSM 6215648904 with the title “The Czech Economy in the Process of Integration and Globalisation, and the Development of Agricultural Sector and the Sector of Services under the New Conditions of the Integrated European Market”.

REFERENCES

- ANDĚL, J. (1976) *Statistická analýza časových řad*, SNTL, Praha, 271 pp.
- AHUMADA, H., GAREGNANI, M. L. (1999) Hodrick-Prescott Filter in practice, UNLP.
- ARTIS, M., MARCELLINO, M., PROIETTI, T. (2004) Characterizing the Business Cycles for Accession countries. CEPR-EABCN Conference of Business Cycle and Acceding Countries, Vienna.
- BAXTER, R., KING, R. G. (1999) Measuring Business Cycles: Approximate Band – Pass Filters for Economic Time Series. *Review of Economic and Statistics*, vol. 81, no. 4, pp. 575-593.
- BURDA, M., WYPLOZS, C. (2001) *Macroeconomics. A European text*. 3rd ed. Oxford: Oxford University Press, 572 pp., ISBN 0-19-877650-0.
- BURNS, A.F., MITCHELL, W.C. (1946) *Measuring Business Cycles*. New York, National Bureau of Economic Research, pp.590, ISBN: 0-870-14085-3 .
- BONENKAMP, J., JACOBS, J., KUPER, G.H., (2001) *Measuring Business Cycles in the Netherlands, 1815 – 1913: A comparison of Business Cycle Dating Methods*. SOM Research Report, No. 01C25 , Systems, Organisation and Management, Groningen. University of Groningen [online].
- CANOVA, F. (1998) De-trending and business cycle facts, *Journal of monetary Economic*, vol. 41, pp. 533-540.
- CANOVA, F., (1999) Does De-trending Matter for the Determination of the Reference Cycle nad Selection of Turniny Points?, *The Economic Journal*, Vol. 109, No. 452 (Jan., 1999), pp. 126-150.

- DEJONG, N. D., CHETAN, D. (2007) Structural Macroeconometrics, Princeton University Press, New Jersey, 338 pp., ISBN-10: 0-691-12648-8 .
- GUAY, A., ST-AMANT, P., (1997) Do the Hodrick-Prescott and Baxter-king Filters Provide a Good Approximation of Business Cycles? Université a Québec á Montréal, Working paper No. 53,.
- HALLETH, A. H., RICHTER, C. (2004) A Time-frequency Analysis of the Coherences of the US Business Cycle and the European Business Cycle, CEPR Discussion Paper no. 4751, London, Centre for Economic Policy Research.
- HALLETH, A. H., RICHTER, C. (2007) Time Varying Cyclical Analysis for Economies in Transition, CASE Studies & Analyses no. 334, Center for Social and Economic Research.
- HARDING, D., PAGAN, A. (2006) Measurement of Business Cycles, Research paper number 966, Melbourne, ISSN 0819-2642.
- HARVEY, A.C., JAEGER, A. (1993) De-trending, Stylized Facts and the Business Cycle. Journal of Applied Econometrics 8: pp. 231-47.
- HODRICK, R.J., PRESCOTT, E.C. (1980) Post-war U.S. Business Cycles: An Empirical Investigation, mimeo, Carnegie-Mellou University, Pittsburgh, PA. , 24 pp.,.
- KING, R. G., REBELO, S. T. (1993) Low frequency Filtering and Real Business Cycles. Journal of Economic Dynamics and Control. vol. 17, pp.207-231.
- MILL, T. C. (2003): Modelling Trends and Cycles in Economic Time Series, Palgrave Macmillan, 178 pp., ISBN 1-4039-0209-7.
- SZAPÁRY, G., DARVAS, Z. (2008) Business Cycle Synchronization in the Enlarged EU, Open Economies Review, Volume 19, Number 1 / February.
- WOOLDRIDGE, J. M. (2003): Introductory Econometrics: A modern approach. Ohio, 863 pp., ISBN 0-324-11364-1.
- WOZNIAK, P., PACZYNSKI (2007) Business Cycle Coherence between Euro Area and the EU New Member State: a Time-Frequency Analysis. Center for Economic Research and Graduate Education and the Economics Institute (CERGE-EI), Prague,.
- SCHUMPETER, J. A. (1939) Business Cycles. A theoretical, Historical and Statistical Analysis of the Capitalist Process, New York, 461 pp..

Analiza ruchów w zakresie aktywności gospodarczej w Czechach – podejście frekwencyjne

Streszczenie: Głównym celem pracy jest określenie rodzajów cykli koniunkturalnych występujących w Czechach. Mając na uwadze cechy cykli koniunkturalnych oraz częstotliwość ich występowania, w artykule zastosowano analizę harmoniczną. Spośród technik de-trending wybrano filtr „pierwszych różnic”, filtrowanie liniowe, model nieobserwowanych

komponentów oraz filtr Hodricka-Prescotta. Przy użyciu tej ostatniej metody zbadano oszacowanie zmienności cyklicznej z derywacją wygładzonego parametru, które zostało zaprojektowane specjalnie dla Czech. Celem badań jest wyodrębnienie typów cyklicznych fluktuacji w Czechach. W pracy przeprowadzono również analizę potencjalnych źródeł ruchów cyklicznych.

Słowa kluczowe: analiza harmoniczna, trendy deterministyczne, cykle stochastyczne, aktywność ekonomiczna, cykle koniunkturalne JEL.

Application of selected statistical methods to ICT used in Public Procurement System in Poland

Wacława Starzyńska, Justyna Wiktorowicz

Katedra Statystyki Ekonomicznej i Społecznej UŁ
e-mail: wacstar@uni.lodz.pl, justynawiktorowicz@uni.lodz.pl

Streszczenie: W niniejszej pracy podjęta zostanie próba wielowymiarowej analizy rozwoju elektronicznych zamówień publicznych oraz jego determinant w oparciu o dane pochodzące z wywiadów kwestionariuszowych, przeprowadzonych wśród podmiotów zamawiających, zobligowanych do stosowania Pzp. Będzie mieć miejsce testowanie wybranych metod statystycznych dla potrzeb pomiaru, klasyfikacji i analizy determinant rozwoju elektronicznych zamówień publicznych.

Słowa kluczowe: zamówienia publiczne, regresja logistyczna

INTRODUCTION

Public procurement is a necessary element of economies functioning in every part of the world. The public procurement market in the European Union is estimated at 16% of GDP [EIM 2004] and the market's value in Poland in 2008 stood at 109.5 billion zlotys, i.e. 8.6% of Polish GDP [*Sprawozdanie ...*, 2009], which suggests that the Polish public procurement market has only started to grow. In the next few years, it can be expected to expand faster, following a wider use of public aid co-financed by the EU funds. Therefore, instruments need to be implemented that could improve the effectiveness of public procurement systems and give a more important role to the SMEs. A broader use of ICT tools in public procurement, especially the computerization of its procedures, seems a good and relevant way forward.

The paper performs a multivariate statistical analysis of e-procurement development and of its determinants using data derived from surveys of the awarding entities covered by the Public Procurement Law. Statistical methods were especially applied in order to measure, classify and model the development of ICT-based public procurement.

COMPUTERIZATION OF PUBLIC PROCUREMENT PROCESSES IN EU – MAJOR TRENDS

Modernization of the public procurement market and its better availability may become a key factor strengthening European Union's competitiveness and providing enterprises with new opportunities, especially the SMEs. Improved competitiveness and efficiency of enterprises operating in the public procurement area that stem, among other things, from the development of e-procurement are likely to influence the whole economy and play an important role in achieving the goals of the Lisbon Strategy. New ICT techniques applied to purchase goods and services make competition stronger and public purchasing more efficient, especially in terms of time and financial savings. It is estimated that e-procurement may help the awarding entity and the economic operators to save as much as 5% of their expenses and up to 50-80% of transaction costs [Ramboll M. 2004]. Previous years' experiences show that an awarding entity using e-procurement procedures may save even 10-20% compared with one applying the traditional methods. It is so, because entrepreneurs operating IT procedures can offer better prices, while in a traditional system their offers are final [Wikariak 2009].

In 2006, 71% of the Member States already had functional ICT-based public procurement systems, while the remaining 29% were in the process of developing them. Poland belonged to the second group. Among the European forerunners of e-procurement systems there are Italy and UK. However, the implementation of e-procurement instruments does not necessarily mean that they are used more widely. A case in point is the electronic mode of tendering (Prince, Zeiden 2004). Studies conducted by GHK and Technopolis have revealed that e-procurement is not common in Europe: 58% of the surveyed companies have never submitted fully electronic tenders and 43% have not done that even partly. The proportion of companies submitting fully electronic tenders 'at least sometimes' is only 24%. However, there are large variations between particular Member States. The leaders in the area are the UK and Lithuania, where 62% and 69% of companies, respectively, claim to submit electronic tenders at least sometimes. Finland, Denmark and Sweden also score well. The lowest percentages of such enterprises have been found in France, Malta, Hungary and Poland (3%). Similar findings have been provided by the studies discussed in the next sections of the paper [GHK&Technopolis 2007].

COMPUTERIZATION OF PUBLIC PROCUREMENT PROCESSES IN POLAND – APPLICATION OF SELECTED STATISTICAL METHODS

THE METHODOLOGY FOR QUANTIFYING IT DEVELOPMENT IN PUBLIC PROCUREMENT

It is difficult to evaluate the development of e-procurement because of insufficient informational resources. The statistics provided by the Public Procurement Office (PPO) only include information on the number of tenders involving electronic bidding and the Dynamic Purchasing System (DPS). In 2006, electronic bidding accounted for only 0.3% of all awarded contracts, with the rate falling to 0.05% in 2007 and 2008 and then rising to 0.11% in first quarter 2009. The Dynamic Purchasing System is not very popular either – in 2007 DPS was not used at all, in 2008 there were 10 DPS-based tenders, but in the first quarter of 2009, only one announcement of this type was made. Data on other e-procurement instruments are not available [BI UZP 2008, 2009]. This kind of information can be usually acquired in the course of field research. Unfortunately, most of the data provided during interviews and surveys (e.g. electronic) are confidential, which prevents the use of more advanced analytical methods. This fact inspired us to try to construct a synthetic indicator of the development level of public e-procurement (based on discrete variables).

The selection of diagnostic variables is one of the most important and also most difficult issues in constructing synthetic variables. Variables are generally selected using two approaches: one considering variable relevance and one statistical. An overview of the Polish solutions and studies requested by the European Commission suggests that the diagnostic variables should be enhanced with factors given in Table 1. The list is certainly not exhaustive, but it indicates e-procurement instruments and procedures that are currently used in Poland.

Each of the variables is binary (1 meaning that the respondent is knowledgeable of public e-procurement procedures). The obtained synthetic indicator representing a sum of all diagnostic variables and thus measuring the level of computerization of a public procurement system (synt) took values between 0 and 19 points. The higher the indicator's value, the more advanced e-procedures of a public procurement system.

The constructed indicator will be used to break down the awarding entities into two groups:

Gr. 1 – organizations with an insufficient level of e-procurement development (the indicator's values are up to 9 points),

Gr. 2 – organizations with a sufficient level of e-procurement development (the indicator's values range from 10 to 19 points).

An alternative to the above approach is classifying the awarding entities using cluster analysis. Our study will test both the procedures.

Table 1. List of potential diagnostic variables (questionnaire questions)

Symbol	Variable (question)
X ₁	Did you know that the Public Procurement Office (PPO) had the web portal www.portaluzp.gov.pl ?
X ₂	Did you know that the PPO had the website – www.uzp.gov.pl ?
X ₃	Are you aware of the availability of the Internet bulletin TED?
X ₄	Do you use the web portal www.portaluzp.gov.pl ?
X ₅	Do you use the website at www.uzp.gov.pl ?
X ₆	Does your organization use software that supports the awarding of contracts?
X ₇	Does your organization have a buyer profile?
X ₈	Does your organization use electronic means of transferring statements, requests, notifications and information related to contract awarding?
X ₉	Do you think that the PPL obligates the awarding entity to specify the essential terms of contract on the website for bidding other than unlimited bidding?
X ₁₀	Does your organization publish such specifications on the website for bidding other than unlimited bidding?
X ₁₁	Do you think the PPL allows submitting tenders electronically, if approved by the awarding entity?
X ₁₂	Do such tenders need to have a secure electronic signature that is verifiable using a valid qualified_certificate?
X ₁₃	Does your organization accept electronic tenders bearing a secure electronic signature that is verifiable using a valid qualified_certificate to award contracts?
X ₁₄	Do you know the electronic bidding scenario laid out in PPL regulations?
X ₁₅	Has your organization ever used electronic bidding that the previous PPL regulations called an electronic auction?
X ₁₆	Do you know the electronic auction scenario defined by the current PPL regulations?
X ₁₇	Has your organization ever participated in an electronic auction?
X ₁₈	Do you know the dynamic purchasing system scenario (DPS) provided in the PPL regulations?
X ₁₉	Has your organization ever used the dynamic purchasing system scenario?

Source: own study.

MULTIVARIATE STATISTICAL METHODS IN THE EVALUATION OF E-PROCUREMENT DEVELOPMENT

The development of e-procurement in Poland has been analyzed using mainly the multivariate methods. The multivariate analysis allows dealing with many variables at the same time. Among the methods used by the analysis, the most useful seem to be cluster analysis (to classify the tenderers) and logistic regression (to model the development of e-procurement).

In the *logistic regression models*, the dependent variable is dichotomous and its value is 1 when the probability of success is θ , or 0 when the probability of failure is $1-\theta$. A logistic regression model examines the relationship between one or more independent variable (called risk factors) and the log odds of binary outcome variable. This relation is described by the formula:

$$P(Y = 1) = \frac{1}{1 + \exp(-(B_0 + B_1 \cdot x_1 + \dots + B_k \cdot x_k))} \quad (1)$$

where x_1, x_2, \dots, x_k are independent variables (continuous or categorical). Odds ratios are commonly employed to measure the strength of the partial relationship between one predictor and the dependent variable (in the context of the other predictor variables). A model is constructed based on an iterative maximum likelihood procedure [Wuensche 2007].

Cluster analysis is an exploratory data analysis tool that aims at sorting different objects into groups in a way that the degree of association between two objects is maximal if they belong to the same group and minimal otherwise [Larose 2006, p.151]. A typical approach to cluster analysis first compiles a table of relative similarities or differences between all objects and then uses this information to put objects into groups. The table of relative similarities is called a proximities matrix. The method of combining objects into groups is called a clustering algorithm. The aim is to combine similar objects into separate groups. After identifying the inter-object distances, the next procedure of cluster analysis divides the objects into groups according to the distances (one possibility being the squared Euclidian distance). This is sum across variables (where $i = 1, \dots, v$) of the squared difference between the score on variable i for the one case (X_i) and the score on variable i for the other case (Y_i). Following that, all the distances between the entities (cases and clusters) are recomputed and then two entities separated by the smallest distance are grouped together. When one of the entities is a cluster, or both, the averaged squared Euclidian distance between the members of both entities is computed. This procedure continues until all cases are formed into one giant cluster. It is for the researcher to decide when the procedure should stop and to accept a solution having k clusters (where $k = 1, \dots, N$) [Walesiak, Gatnar 2009, pp. 407-420].

EMPIRICAL RESULTS

Between January and May 2008, an empirical survey requested by the Public Procurement Office (PPO) was conducted. It aimed at evaluating the status quo and the prospects for public procurement computerization in Poland (the survey's results are presented in detail in [Borowicz 2008]). The survey was designed to investigate the use of broadly understood ICT tools in public procurement. This paper presents the results of interviews conducted with a representative, random sample of 701 awarding entities. Because the questionnaire used in the survey catered

for PPO's informational needs, this fact largely narrowed the range of statistical methods that could be applied to analyze respondents' answers. Therefore, most variables represent nominal-scale measures. Moreover, the paper presents the results of multi-dimensional analysis of e-procurement development. Methods enabling the synthetic evaluation of e-procurement development (by constructing a synthetic indicator to evaluate the level of computerization in the surveyed entities) as well as classification of entities by the level of sophistication of their e-procurement systems (using the constructed indicator and cluster analysis) will be thoroughly tested. Besides, an attempt is being made to identify the determinants of e-procurement system development in the awarding entities (using logistic regression).

Because of the practical purpose of the conducted survey (it aimed at identifying the causes of the very weak interest in electronic procedures as well as actions that could curb unfavourable trends), most of its questions requested respondents' opinions on electronic PP procedures, so using them to construct the indicator was pointless. Unfortunately, the limited number of potential variables in the survey and the nominal level of their measurement prevented us from using most formal methods of variable selection. Consequently, the variables were selected using the substantial approach and the indicator of ICT development in public procurement was constructed using all potential diagnostic variables (Table 1). All calculations were run based on the SPSS 14.0 software.

The empirical investigation showed that the surveyed organizations had relatively good knowledge of the electronic instruments used in public procurement – especially of the PPO's web portal and website that the awarding entities use willingly. However, the possibility of submitting tenders electronically apart, the practical use of other instruments is insufficient and even insignificant in the case of electronic bidding, electronic auctioning and Dynamic Purchasing System (despite their being known well). Consequently, the overall evaluation of ICT level in public procurement is not very high.

The level of ICT used by the awarding entities seems to be mediocre as measured by the synthetic indicator (both \bar{x} and Me are close to 9 against the maximum score being 19) – the index values for the sample range from 3 to 15 points. It is noteworthy that the distribution of the constructed variable is 'stabilized' (low skewness and kurtosis of distribution), which is confirmed by the analysis of normality plot and box-plot. This justifies the application of the parametric methods of analysis. The constructed indicator's reliability is relatively high ($\alpha=0.560$), but the performed factor analysis indicates that mid-scales can be distinguished. Because of the limited number of factors, the next part of the analysis uses an indicator that has been constructed in line with the presented procedure. An additional argument in favour of this approach is high compatibility of awarding entities' classifications based on the synthetic variable and on the cluster analysis (Table 2).

Both the methods (using the same list of diagnostic variables/items) produced similar classifications of awarding entities as far as the computerization of public procurement is concerned. The value of Kappa Cohen's consistency ratio confirms the both the classifications are very much alike ($\kappa=0.744$ with p close to 0).

Table 2. The similarity of classification of the awarding entities performed using the synthetic indicator and cluster analysis with respect to the development of e-procurement

		<i>clusters</i>		Total	
		unsufficient	sufficient		
<i>synt gr</i>	unsufficient develop- ment	Number %	205 88.7	24 14.2	229 57.3
	sufficient development	Number %	26 11.3	145 85.8	171 42.8
Total		Number %	231 100.0	169 100.0	400 100.0

Source: own calculations.

The development of electronic procedures within public procurement is driven by many factors – administrative, legislative, economic and psychological. The scope of the questionnaire surveys prevents a correspondingly broad analysis. The main barriers as indicated by the respondents are too complicated regulations laid out in the PPL and the related laws, as well as professional and practical training in this field falling short of the actual needs [Wiktorowicz, Roszko-Grzegorek 2009]. The presented analysis aims to identify factors that have an effect on the development of e-procedures in public procurement. The development of ICT tools in public procurement will be evaluated by means of logistic regression analysis – separately for two dependent variables: *synt_gr* (formed by dividing *synt* variable variants into two classes, ie. Gr. 1 and Gr. 2) and a cluster (resulting from the classification provided by cluster analysis). The multi-dimensional analysis of relations was carried out by distributing the factors among three groups, i.e.:

- (1) factors related to the awarding entity's location (Table 4): *REG* – region by GDP (economic development – ED), *SIZE* – the size of the urban area;
- (2) major characteristics of the awarding entity (Table 5): *N* – the number of contracts awarded in 2007 (between the thresholds requiring mandatory PPL application in Poland and in the EU), *L* – the number of employees working in the public procurement department, *TYPE* – the type of the awarding entity;
- (3) respondents' characteristics (Table 6): *IT* – IT skills, *EXP* – professional experience in public procurement, *EDU* – level of education.

We expected e-procurement to be the best developed in awarding entities that initiate large-scale public procurement procedures more often, based in regions at a higher level of economic development (such as Mazowieckie or Slaskie), especially in large urban areas. A more common use of ICT in public procurement

could be facilitated by high skills of employees. The results of the applied modeling procedure proved that our expectations were right (see Tables 3-6).

Table 3. Estimation results of the logistic regression equation parameters (eq. 1a i 1b)

Independent variables	Dependent variable							
	<i>synt gr (1a)</i>				<i>clusters (1b)</i>			
	B	Wald	p	Exp(B)	B	Wald	p	Exp(B)
REG		4.719	0.094			1.129	0.569	
high ED	0.584	4.259	0.039	1.793	0.269	0.916	0.339	1.309
medium ED	0.233	0.701	0.403	1.263	0.071	0.066	0.797	1.074
SIZE		7.367	0.025			6.428	0.040	
country	-	5.284	0.022	0.368	-1.067	6.379	0.012	0.344
small-or medium-	-	0.008	0.928	0.974	-0.384	1.681	0.195	0.681
Constant	-	2.248	0.134	0.580	-0.048	0.018	0.894	0.953

* Reference group: for *RG* – region with low economic development, for *SIZE* – large urban areas

Source: own calculations.

The estimates of models 1a) and 1b) provided in Table 3 show that the size of the urban area where the awarding entity is located significantly influences the evaluation of ICT used by the entity in the public procurement area. In organizations located outside large urban areas the use of ICT is much more limited ($p=0.022$ for model 1a) and $p=0.012$ for model 1b)), with deflection reaching as much as 65%. The level of economic development in the awarding entity's district does not exert a statistically significant influence, but a comparison of models 1a) and 1b) reveals significant deflections between organizations based in the well-developed regions and in the less successful ones ($p=0.039$). Notwithstanding, both models show, as expected, that in the richer regions the odds of organization's achieving a higher level of computerization are larger (about 30-80% depending on the model) compared with the reference group.

Table 4. Estimation results of the logistic regression equation parameters (eq. 2a and 2b)

Independent variables	Dependent							
	<i>synt gr (2a)</i>				<i>clusters (2b)</i>			
	B	Wald	p	Exp(B)	B	Wald	p	Exp(B)
<i>N</i>	0.026	12.766	0.000	1.026	0.018	7.610	0.006	1.018
<i>L</i>		4.102	0.251			0.550	0.908	
1 person	-20.624	0.000	0.999	0.000	-20.700	0.000	0.999	0.000
2 – 5 persons	-20.185	0.000	0.999	0.000	-20.653	0.000	0.999	0.000
6 – 10 persons	-20.077	0.000	0.999	0.000	-20.238	0.000	0.999	0.000
<i>TYPE</i>		7.292	0.063			10.970	0.012	
<i>ZOZ</i>	0.468	0.875	0.350	1.597	0.311	0.419	0.518	1.364
government	0.541	1.128	0.288	1.718	0.904	2.984	0.084	2.471
others	-0.373	1.831	0.176	0.688	-0.443	2.680	0.102	0.642
Constant	19.975	0.000	0.999	0.004	20.336	0.000	0.999	0.007

* Reference group: for *L* – over 10 persons, for *TYPE* – local administration

Source: own calculations.

As shown by data presented in Table 4, organization's involvement in public procurement activities as measured by the number of tenders is important. Organization's chances of developing e-procurement tools are determined quite strongly by its type, but local administration is not statistically significantly different from the reference group (particularly high and positive differences are noted for the central government administration). Compared with other factors, such as the aforementioned number of tenders and the type of organization, the rate of employment proved to be insignificant (p is close to 1).

Interesting conclusions can be drawn from the analysis of results presented in Table 5. A comparison of the previous equations shows that ICT development in the organizations is mainly decided by the human factor and not by solutions provided within the PPL. Organizations with experienced and well-educated staff are more likely to develop IT solutions in the public procurement area.

Table 5. Estimation results of the logistic regression equation parameters (eq. 3a i 3b)

Variables	Dependent							
	<i>synt gr (3a)</i>				<i>clusters (3b)</i>			
	B	Wald	p	Exp(B)	B	Wald	p	Exp(B)
<i>IT</i>		6.195	0.103			3.979	0.264	
low	-0.955	4.108	0.043	0.142	-1.039	1.346	0.246	0.354
medium	-0.500	0.667	0.414	0.607	-0.308	0.242	0.623	0.735
high	-0.206	0.109	0.741	0.814	0.045	0.005	0.943	1.046
<i>EXP</i>		12.804	0.005			16.911	0.001	
below 1 year	-0.259	7.212	0.007	0.284	-1.548	9.300	0.002	0.213
1 – 2 years	-0.818	3.689	0.055	0.441	-0.437	1.163	0.281	0.646
2 - 5 years	-0.579	5.887	0.015	0.560	-0.801	10.592	0.001	0.449
<i>EDU</i>		5.795	0.215			10.661	0.031	
secondary or lower	-0.128	5.368	0.021	0.324	-1.338	7.055	0.008	0.262
incomplete tertiary	-0.028	3.429	0.064	0.358	-1.087	3.657	0.056	0.337
undergraduate	-0.678	1.555	0.212	0.508	-0.400	0.505	0.477	0.670
graduate	-0.856	3.737	0.053	0.425	-1.145	6.225	0.013	0.318
Constant	1.261	2.998	0.083	3.530	1.300	3.026	0.082	3.669

* Reference group: for *IT* – very high, for *EXP* – over 5 years; for *EDU* - doctorate/ PP postgraduate
Source: own calculations.

The analysis is completed with equations that were built using all variables from the list of potential factors (Table 6). The estimates were similar for both response variables: unlike the other factors, the type of awarding entities, the number of the awarded contracts and employees experience in PP activity proved significant. The only difference between versions (a) and (b) is the location of the awarding entity – both regions (4a model) and the size of the urban area (4b model) turned out to be important. Therefore, the characteristics are analogous to those presented in models 1-3.

Table 6. Estimation results of the logistic regression equation parameters (eq. 4a and 4b)

Dependent: <i>synt_gr</i> (4a)					Dependent: <i>clusters</i> (4b)				
Model 4a	B	Wald	p	Exp(B)	Model 4b	B	Wald	p	Exp(B)
<i>REG</i>		7.792	0.020	2.233	<i>SIZE</i>		8.038	0.018	
high ED	0.803	7.261	0.007	1.454	country	1.289	7.579	0.006	0.275
medium ED	0.374	1.528	0.216	1.029	small- and medium-	0.694	4.775	0.029	0.500
<i>N</i>	0.028	15.749	0.000	1.809	<i>N</i>	0.01	8.031	0.005	1.018
<i>TYPE</i>		8.199	0.042		<i>TYPE</i>		13.142	0.004	
<i>ZOZ</i>	0.593	1.350	0.245	1.610	<i>ZOZ</i>	0.22	0.208	0.648	1.256
government	0.476	0.827	0.363	0.660	government	0.57	1.067	0.302	1.771
administra-					administra-	1			
others	-0.415	2.189	0.139		others	0.700	5.770	0.016	0.496
<i>EXP</i>		8.536	0.036		<i>EXP</i>		13.209	0.004	
below 1 year	-0.957	4.241	0.039	0.384	below 1	1.219	5.945	0.015	0.296
1 – 2 years	-0.548	1.522	0.217	0.578	1 – 2 years	0.122	0.083	0.773	0.885
2 - 5 years	-0.573	5.223	0.022	0.564	2 - 5 years	0.770	9.070	0.003	0.463
Constant	-0.639	3.433	0.064	0.528	Constant	0.86	3.927	0.048	2.375

* See symbols in Tables 4-7

Source: own calculations.

CONCLUSION AND FUTURE RESEARCHES

The procedures of multi-dimensional statistical analysis are only partly useful in examining the survey's results, because of the nominal level of measurement of the investigated phenomena. The measurement methodology proposed in this paper allows comprehensive evaluation of computerization processes in public procurement. The use of logistic regression made it possible to identify certain factors determining progress in the area of e-procurement. Compared with other factors, the type of the awarding entity and its location, as well as the number of awarded contracts and employees' experience in PP turned out significant. Awarding entities based in cities have the best-developed e-procurement facilities, especially local government units and Public Healthcare Establishments. Their involvement in public procurement activity as measured by the number of awarded contracts and the number of PP-dedicated personnel is relatively high. They can develop ICT solutions more intensively because their staff is supported by skilled IT specialists, employees dealing with public procurement are trained to upgrade their ICT skills and large-scale informational campaigns on e-procedures are more readily available.

The development of e-procurement may be a factor boosting innovativeness – in both enterprises competing for contracts and in the awarding entities. Our future research will investigate how public procurement influences innovation activities in enterprises, especially in the SMEs. We will continue our efforts to test the development of e-procurement using statistical methods.

BIBLIOGRAPHY

- BI UZP (Biuletyn Informacyjny UZP) No 12/2008 and 2/2009
- Borowicz A. (ed.), Starzyńska W., Kubiak A., Wiktorowicz J., Roszko-Grzegorek E., Wendołowski L., Popławski S., Kaniewicz S., Plackowski K., Michalski W. (2008) Informatyzacja polskiego systemu zamówień publicznych – stan obecny, oczekiwania, bariery, UZP, Warszawa
- Carsa (2006) Compliance Verification in Electronic Public Procurement. Final report, EC, Brussels
- GHK & Technopolis (2007) Evaluation of SME Access to Public Procurement Markets in the EU. Final Report, EC, Brussels
- Larose D.T. (2006) Odkrywanie wiedzy z danych. Wprowadzenie do eksploracji danych, PWN, Warszawa
- Prince Y.M., van der Zeiden P.Th. (ed.) (2004) The access of SMEs to public procurement contracts. Final report, EC, Vienna
- Ramboll Management (2004) Impact Assessment Action Plan on e-Procurement – Part 1, EC, Brussels
- Ramboll Management (2004) Electronic Public Procurement in EU Member States: Country Reviews, EC, Brussels
- Sprawozdanie z funkcjonowania systemu zamówień publicznych w 2008 roku (2009) UZP, Warszawa 2009
- Ustawa z dnia 29 stycznia 2004 r. Prawo zamówień publicznych, Dz.U. Nr 223, poz. 1655 z późn. zm.
- Walesiak M., Gatnar E. (2009) Statystyczna analiza danych z wykorzystaniem programu R, PWN, Warszawa
- Wiktorowicz J., Roszko-Grzegorek E. (2009) Procedury elektroniczne w zamówieniach publicznych, Wydawnictwo AE we Wrocławiu, złożony do druku
- Wuensche K.E. (2007) Statistics Lessons, East Carolina University, Greenville 2007

Application of selected statistical methods to ICT use in the Public Procurement System in Poland

Abstract: The paper discusses the application of multivariate analysis tools to the development of e-procurement and to the determinants of the process using data derived from a survey of the awarding entities. In particular, statistical methods were used to measure, classify and model the development of computerized public procurement.

Keywords: public procurement, logistic regression

ŁĄCZENIE DANYCH Z DYNAMICZNYM OBCIĘCIEM. WYNIKI WSTĘPNE¹

Paweł Strawiński

Katedra Statystyki i Ekonometrii UW
e-mail: pstrawinski@wne.uw.edu.pl

Streszczenie: Problem ewaluacji działań jest problemem brakujących danych. Zatem ważnym zagadnieniem jest konstrukcja adekwatnej kontrfaktycznej grupy porównawczej. W artykule zaproponowane modyfikację procedury pozwalającej uznać obserwację za dobrze dobrany stan kontrfaktyczny. Proponowana zmiana polega na wprowadzeniu nowego mechanizmu sprawdzania czy jednostka z grupy kontrolnej jest bliskim odpowiednikiem jednostki z grupy porównawczej. W ostatniej części artykułu zaprezentowano wyniki symulacji badającej właściwości statystyczne proponowanego rozwiązania. Uzyskane wyniki wskazują, że metoda pozwala osiągnąć lepsze rezultaty niż dotychczas wykorzystywane techniki

Słowa kluczowe: łączenie danych, matching, propensity score, caliper, ewaluacja.

WSTĘP

W literaturze dotyczącej oceny różnego typu działań o charakterze zorganizowanych programów w ostatnich latach jako narzędzie badawcze wykorzystywane są metody quasi-eksperymentalne. Polegają one na odtworzeniu na podstawie danych nie eksperymentalnych warunków eksperymentu.

Badacz obserwuje efekt działania programu i dokonuje oceny jego skuteczności na podstawie oceny pewnych wielkości. W danej chwili rozpatrywana jednostka może znajdować się w jednym z dwóch alternatywnych stanów, ale nie może być w obu jednocześnie. Może być uczestnikiem programu, bądź w nim nie uczestniczyć. Gdyby dane pochodziły z eksperymentu wystarczyłoby porównać wynik jednostek z grupy eksperymentalnej i grupy kontrolnej. Zatem problem oceny programów jest problemem brakujących danych [Heckman i in. 1997].

¹ Badanie finansowane ze środków na naukę w latach 2008-2009.

Ocena wpływu działania na obserwowany rezultat zawiera spekulację w jaki sposób jednostka by funkcjonowała, gdyby nie została poddana oddziaływaniu [Caliendo, Kopeinig 2008]. W literaturze statystycznej i ekonometrycznej opisanych jest szereg metod pozwalających na utworzenie grupy porównawczej. Występują one pod nazwą metod łączenia. Najpopularniejszą jest metoda wykorzystująca wartości wektora *propensity score*. Wykorzystywane są różne warianty tej metody różniące się sposobami łączenia obserwacji w pary, warunkami uznania stopnia podobieństwa obserwacji za wystarczający, oraz sposobami obliczenia efektu oddziaływania. Każda z implementacji z jednej strony umożliwia wyznaczenie kontrfaktycznego zachowania jednostki, z drugiej, pomimo tego, że metody są asymptotycznie równoważne [Hirano i in. 2003], prowadzi do uzyskania innego wyniku.

W dalszej części artykułu zostaną w skrócie omówione wady i zalety dotychczas wykorzystywanych podejść do konstrukcji grupy porównawczej i wyznaczania na jej podstawie stanu kontrfaktycznego. Głównym celem artykułu jest udoskonalenie jednej z metod stosowanych do łączenia danych. Proponowana zmiana polega na wykorzystaniu zmodyfikowanego mechanizmu sprawdzania czy jednostka z grupy kontrolnej jest bliskim odpowiednikiem jednostki z grupy poddanej oddziaływaniu. Następnie za pomocą symulacji porównamy dotychczas stosowane w badaniach metody z zaproponowaną modyfikacją. Celem dodatkowym jest zbadanie właściwości statystycznych, szczególnie małopróbkowych, różnych metod łączenia danych oraz identyfikacja źródeł wariancji.

METODY ŁĄCZENIA DANYCH

Przeprowadzając analizę ekonomiczną działania pewnej polityki jako jego efekt obserwowana jest dla każdej jednostki i wartość zmiennej wynikowej. Wynosi ona Y_{1i} jeżeli jednostka jest poddana działaniu, lub Y_{0i} jeżeli polityka na nią nie oddziałuje. Jeżeli można by było jednocześnie obserwować obie wartości dla każdej poddanej badaniu jednostki problem związany z oceną programu zniknąłby.

W wielu pracach empirycznych dokonuje się oceny efektów polityki opisując ją jako program. Celem jest pomiar przeciętnego efektu zastosowania programu. Niech zmienna P opisuje stan uczestnictwa w programie i przyjmuje wartość $P_i=1$, gdy jednostka i jest uczestnikiem programu (należy do grupy eksperymentalnej), a $P_i=0$ oznacza, że znajduje się poza programem. Wówczas przeciętny efekt działania programu (ang. *average treatment effect*) można zapisać jako:

$$Y_{ATE} = \frac{1}{N} \left[\sum_{i=1}^N (P_i Y_{1i} + (1 - P_i) Y_{0i}) \right] \quad (1)$$

Efekt programu obliczony w sposób opisany przez wzór (1) jest efektem dla przeciętnej jednostki pochodzącej z badanej próby, bez uwzględnienia czy jest ona poddana oddziaływaniu czy nie. W praktyce jako wielkość efektu przyjmuje się różnicę w średnim wyniku dla grupy eksperymentalnej i wszystkich jednostek nie-

uczestniczących w eksperymencie. Może to spowodować niedokładność, ponieważ jednostki z grupy eksperymentalnej mogą w sposób systematyczny odróżniać się od pozostałych, co przy pomiarze nie jest uwzględnione.

W praktyce badawczej dąży się do obliczenia innej miary, niosącej więcej informacji o skuteczności działania. Jest nią przeciętny efekt oddziaływania programu na jednostki poddane oddziaływaniu (ang. *average treatment on treated*). Jest to efekt ograniczony do grupy, na którą oddziaływano.

$$Y_{ATT} = E(Y_{1i} - Y_{0i} | P_i = 1) = E(Y_{1i} | P_i = 1) - E(Y_{0i} | P_i = 1) \quad (2)$$

Pokazuje on ilościowy efekt oddziaływania programu wobec jednostek w nim uczestniczących. Jest on równy przeciętnej zmianie wartości zmiennej wynikowej wskutek wystąpienia oddziaływania. Informuje ona jaki jest zysk bądź strata z faktu uczestniczenia w programie dla jego beneficjentów. Należy zaznaczyć, że wielkość $E(Y_{1i}|P_i=1)$ jest oczekiwaną wartością zmiennej wynikowej dla beneficjentów programu i jest wartością obserwowaną, natomiast $E(Y_{0i}|P_i=1)$ jest wielkością dla nieobserwowanego stanu kontrfaktycznego.

Łączenie obserwacji polega na obliczeniu miary podobieństwa, a następnie określeniu według wybranego algorytmu, która obserwacja z grupy kontrolnej jest najbardziej zbliżona do obserwacji z grupy eksperymentalnej. Celem łączenia obserwacji jest skonstruowanie poprawnej grupy porównawczej zastępującej brakujące dane, bez konieczności przyjmowania dodatkowych założeń *a-priori* [Blundell i Costa Diás 2000]. Jego zadaniem jest odtworzenie warunków eksperymentu poprzez taki wybór jednostek do grupy kontrolnej, by była jak najbardziej zbliżona pod względem obserwowanych charakterystyk do grupy poddanej oddziaływaniu. Metody łączenia koncentrują się na sposobach konstrukcji grupy referencyjnej wykorzystując techniki nieparametryczne w celu transformacji rzeczywistych danych zebranych podczas badania przekrojowego w zbiór danych quasi-eksperymentalnych. Dzięki wykorzystaniu metod nieparametrycznych maksymalne i efektywne wykorzystywane są dostępne informacje zawarte w danych empirycznych. Poprawna konstrukcja sztucznej grupy odniesienia zapewnia zredukowanie obciążenia i zwiększenie precyzji uzyskanych wyników [Rubin 1973]. Podobnie jak szereg innych elementów teorii ekonometrii, jej korzenie i pierwsze zastosowania tkwią w biologii i naukach medycznych.

Wśród metod łączenia danych można wyróżnić dwa podejścia. Historycznie pierwszą metodą było łączenie według wartości współzmiennych (ang. *covariates matching*). Współzmiennymi nazywa się zmienne używane w procedurze łączenia. Ich nazwa wywodzi się stąd, że różne ich wartości determinują zarówno jeden ze stanów oddziaływania programu, jak również wartość zmiennej mierzącej efekt ilościowy programu. Zatem ich wartość zmienia się wspólnie z wartością zmiennej określającej stan oddziaływania.

Łączenie według wartości współzmiennych polega na łączeniu w pary obserwacji o identycznych lub zbliżonych wartościach współzmiennych. Jeżeli war-

tości zmiennych zostaną przedstawione jako punkty przestrzeni wielowymiarowej, to statystyczna metoda łączenia polega na podziale tej przestrzeni na klatki według wartości współzmiennych. Następnie łączone są nie obserwacje a utworzone klatki w pary. Niestety, jeżeli wśród współzmiennych występują zmienne ciągle albo jest duża liczba współzmiennych to metoda okazuje się niepraktyczna w zastosowaniu, gdyż duża liczba klatek pozostaje pusta, lub klatki są zbyt duże. Zmniejsza to efektywność łączenia i zwiększa wariancję estymatora wielkości oddziaływania.

Problem znalezienia elementu porównawczego w bliskim sąsiedztwie jest problemem wielowymiarowym. Metodą pozwalającą na uniknięcie konieczności porównywania wszystkich cech badanych jednostek jest użycie zamiast cech pewnej miary podobieństwa obserwacji. Najpopularniejsza i najszerzej stosowaną jest wektor *propensity score* [Rosenbaum i Rubin 1983]. *Propensity score* jest wartością pochodnej funkcji prawdopodobieństwa doboru obserwacji do programu obliczona względem wektora cech jednostki. Wartość wektora dla jednostki jest równoważna *a-priori* prawdopodobieństwu zdarzenia, że jednostka należy do grupy eksperymentalnej poddanej działaniu [Coniffe i in. 2000].

Metoda łączenia danych oparta jest na dwóch założeniach [Rosenbaum i Rubin 1983]. Pierwsze z nich jest nazywane warunkową niezależnością (ang. *Conditional Independence Assumption*). Zakłada się, że badacz dysponuje zbiorem czynników warunkujących W , nazywanych współzmiennymi, takich, że w każdej warstwie² zdefiniowanej przez podzbiór $w \in W$ rozkład obserwowanych cech jednostek jest taki sam w grupie na którą oddziaływano i w grupie kontrolnej.

$$(Y_0, Y_1) \perp P \mid W \quad (4)$$

Oznacza to, że realizacje zmiennej wynikowej są warunkowo względem zastosowanych w procedurze łączenia współzmiennych W niezależne od statusu uczestnictwa w programie. Znajomość stanu uczestnictwa w programie warunkowo względem cech w nie powinna determinować wartości zmiennej wynikowej. Drugim założeniem jest, że wszystkie jednostki poddane działaniu programu posiadają swój odpowiednik znajdujący się poza programem.

$$0 < \Pr(P_i = 1 \mid W) < 1 \quad (5)$$

Oznacza to, że dla każdej wartości ze zbioru zmiennych warunkujących W , i dla każdego uczestnika programu, istnieje z niezerowym prawdopodobieństwem co najmniej jedna jednostka znajdująca się poza programem o identycznych cechach. Ponadto, cechy jednostek nie determinują w sposób jednoznaczny czy jednostka należy do grupy eksperymentalnej czy do grupy kontrolnej.

Z założeń (4) i (5) wynika, że gdy znana jest wartości wektora zmiennych warunkujących w , to oczekiwana wartość zmiennej wynikowej jest warunkowo niezależna od statusu uczestnictwa w programie. Zatem obliczając różnicę w śred-

² Warstwa jest odpowiednikiem klatki.

nich wartościach pomiędzy grupami można uzyskać oszacowanie przeciętnego efektu oddziaływania na jednostki poddane oddziaływaniu.

Jeżeli przydzielenie do grupy eksperymentalnej lub grupy kontrolnej jest warunkowo niezależne od obserwowanych cech jednostki, czyli spełnione jest założenie (4), to obciążenie powodowane przez porównanie rezultatu w dwóch grupach wynikające z różnych wartości współzmiennych może zostać wyeliminowane poprzez ich dopasowanie [Hirano i in. 2003]. Rosenbaum i Rubin [1983] pokazali, że zamiast cech W można zastosować wektor *propensity score* $PS(w)$.

$$PS(w) = \Pr(P = 1 | w) \quad (6)$$

Użycie wektorów redukuje problem wielowymiarowy do jednowymiarowego. Zatem analogicznie do (4) zachodzi

$$(Y_0, Y_1) \perp P | PS(w) \quad (7)$$

W takim przypadku uzyskana wartość oszacowania będzie nadal nieobciążonym estymatorem efektu, jednak w ten sposób uzyskany estymator nie musi być efektywny [Heckman i in. 1997]. Nieefektywność metody *propensity score* pojawia się w przypadku ewaluacji danych pochodzących z eksperymentu [Hahn 1998]. Jest ona związana z niepotrzebnym zastosowaniem przekształcenia danych.

Łączenie danych z wykorzystaniem wybranej techniki w większym stopniu wpływa na charakterystyki grupy kontrolnej niż eksperymentalnej. Grupa eksperymentalna ulega modyfikacji tylko w przypadku, gdy dla poszczególnych jednostek należących do tej grupy nie udaje się znaleźć bliskiego odpowiednika w grupie kontrolnej. Z kolei charakterystyki grupy kontrolnej przed i po wykonaniu łączenia różnią się w sposób znaczący. Celem łączenia jest wybranie ze zbioru potencjalnych jednostek kontrolnych tych najbardziej zbliżonych do jednostek z grupy eksperymentalnej. Zatem grupa kontrolna po łączeniu powinna być bardziej zbliżona pod względem charakterystyk użytych do łączenia do grupy eksperymentalnej niż przed zastosowaniem procedury łączenia

Jest to ogólna metoda posiadająca wiele implementacji, różniących się sposobem obliczenia miary podobieństwa. Po pierwsze, w metodzie *propensity score* można wybrać różne postaci funkcyjne dla prawdopodobieństwa poddania oddziaływaniu. Po drugie, można zastosować jedną z dostępnych alternatywnych miar podobieństwa obserwacji. Typowy estymator oparty o połączone obserwacje przybiera formę [Smith, Todd 2005]:

$$\frac{1}{N} \sum [Y_{1i} - E(Y_{0i} | P_i = 1)] \quad (8)$$

gdzie: $E(Y_0 | P_i = 1) = \sum W(i, j) Y_{0j}$ jest oszacowaniem stanu kontrfaktycznego, a $W(i, j)$ jest macierzą odległości między jednostkami i oraz j .

Po obliczeniu miary podobieństwa jednostek są one łączone za pomocą wybranego przez badacza algorytmu. Łączenie może odbywać się według zasady 1 do 1 (w pary), 1 do wielu (do każdej jednostki poddanej oddziaływaniu dobiera się więcej niż jedną jednostkę z grupy kontrolnej) lub wykorzystując jądro estymator gęstości (dla każdej obserwacji obliczana jest wartość dystrybuanty i dobierany jest odpowiednik z drugiej grupy o identycznej lub najbardziej zbliżonej wartości). W ostatnim kroku obliczane są różnice w zmiennej wynikowej dla połączonych obserwacji. Te różnice są oszacowaniami szukanych efektów.

Problem dobierania obserwacji porównawczej pochodzącej z najbliższego sąsiedztwa (ang. *nearest neighbour matching*) jest wielowymiarowy. Tradycyjnym i najbardziej intuicyjnym sposobem jego rozwiązania jest wskazanie, dla każdej jednostki z grupy poddanej oddziaływaniu, jej najbliższego sąsiada z grupy kontrolnej. Definiujemy zbiór A_i w taki sposób, że dokładnie jeden sąsiad jednostki i do niego należy

$$A_j = \left\{ j \mid j \in \{1..n\}: \min \|w_i - w_j\| \right\} \quad (8)$$

gdzie $\|\cdot\|$ jest metryką. Wówczas macierz wag $W(i,j)$ jest kwadratowa i składa się wyłącznie z zer i jedynek. Jedynek jest przypisywana najbliższemu sąsiadowi, a zera pozostałym jednostkom grupy kontrolnej.

Estymator najbliższego sąsiada działa dobrze, jeżeli W_i oraz W_j są określone na tym samym zbiorze, lub część wspólna tych zbiorów jest stosunkowo duża. Badacz musi podjąć decyzję zrobić z niedokładnie połączonymi obserwacjami (Lee, str. 89). Jeżeli duża liczba słabo połączonych par zostanie odrzucona, wówczas zmniejszy się liczebność grupy kontrolnej, i dla pewnych obserwacji z grupy eksperymentalnej nie będzie dobrej pary, zatem zostaną one pominięte w analizie. Zredukuje to problem zbilansowania, ale równocześnie zmniejszy się efektywność, z uwagi na pominięcie istotnej części informacji. Z drugiej strony uwzględnienie wszystkich par, bez względu na jakość połączenia pogorszy zbilansowanie.

Aby zapobiec porównywaniu elementów bardzo różniących się między sobą wprowadza się poprawkę tzw. *caliper matching* [Cochran i Rubin, 1975]. Działanie poprawki można porównać z działaniem przesłony w aparacie fotograficznym. Jeżeli uwaga jest skoncentrowana na wybranym punkcie, to inne punkty leżące od niego daleko zostają przesłonięte. Procedura usuwa obserwacje, dla których nie ma bliskiego odpowiednika.

$$A_j = \left\{ j \mid j \in \{1..n\}: \min \|w_i - w_j\| < \delta \right\} \quad (9)$$

Do zbioru A_i należą takie punkty j , dla których odległość od najbliższego odpowiednika jest nie większa niż δ . Zwiększa to precyzję estymatora. Gdy rozkład prawdopodobieństwa uczestnictwa w programie jest równomierny w grupie eksperymentalnej i kontrolnej obciążenie usuwa skrajnie źle dopasowane jednostki zmniejszając obciążenie estymatora. Gdy rozkłady w obu grupach znacznie różnią

się, stała wartość obciążenia inaczej traktuje połączenia jednostek o niskim prawdopodobieństwie znalezienia się w grupie eksperymentalnej a inaczej o wysokim. Te pierwsze częściej podlegają obciążeniu. W efekcie po eliminacji źle połączonych par obciążenie estymatora może wzrosnąć. Jest to szczególnie prawdopodobne, gdy zachodzi zależność między prawdopodobieństwem uczestnictwa w eksperymencie a jego wynikiem.

W związku z tym uważamy, że należy uzależnić wartość odciążenia (*caliper*) od wartości wektora *propensity score* dla grupy eksperymentalnej. Wówczas zmodyfikowany mechanizm wyboru potencjalnych jednostek do grupy kontrolnej wygląda następująco

$$A_j = \{j \mid j \in \{1..n\}: \min |w_i - w_j| < \delta w_i\} \quad (10)$$

W ten sposób wiążemy wartość odciążenia z oszacowaną wartością wektora *propensity score*. Postulujemy, aby wartość odciążenia δ była ustaloną frakcją wartości wektora. Wówczas dla obserwacji o niskim prawdopodobieństwie znalezienia się w grupie eksperymentalnej (wartość wektora bliska 0) będziemy wymuszali bardzo dokładne określenie odpowiednika w stanie kontrfaktycznym. Z kolei, gdy prawdopodobieństwo przydzielenia do programu będzie dążyć do 1 dynamiczny mechanizm obciążenia będzie w niewielkim stopniu różnił się od standardowego. W rezultacie, w obliczeniach zostanie pominiętych więcej połączonych jednostek dla których wartość oszacowanego prawdopodobieństwa znalezienia się w grupie eksperymentalnej jest niewielka.

EKSPERYMENT SYMULACYJNY

W celu zbadania właściwości statystycznych proponowanej poprawki przeprowadziliśmy symulację według następującego schematu

1. Ustalenie wielkości próby.
2. Generowanie charakterystyk jednostek, oraz dwóch składników losowych.
3. Na podstawie wylosowanych wartości wykorzystując znane równanie o ustalonych współczynnikach generowane są prawdopodobieństwa przynależności do grupy eksperymentalnej.
4. Obliczana jest zmienna wynikowa jako funkcja deterministyczna charakterystyk, indykatora uczestnictwa w grupie eksperymentalnej oraz niewielkiego składnika losowego.
5. Obliczany jest przeciętny efekt oddziaływania za pomocą trzech algorytmów: łączenie 1 do 1, łączenie 1 do 1 z obciążeniem, oraz łączenie 1 do 1 z dynamicznym obciążeniem.
6. Kroki 1-6 powtarzane są 100000 razy

Budując proces generujący dane założono, że prawdziwa oczekiwana wartość oddziaływania wynosi 1. W efekcie eksperymentu uzyskano wyniki zaprezentowane w Tabeli 1.

Wartości parametrów dobrano tak, aby 25% obserwacji znalazło się w grupie eksperymentalnej. Uzyskana przeciętna liczba par wynosząca 241 wynika z losowej natury eksperymentu i obrazuje niedokładność symulacji. Jak widać, dynamiczny sposób uznawania pary za dobrze połączone pozwala w niewielkim stopniu, ale jednak zmniejszyć obciążenie estymatora efektu oddziaływania. W pierwszej chwili wydawało się, że ceną za zmniejszenie obciążenia oszacowania jest mniejsza precyzja uzyskanego wyniku. Ponieważ literatura wskazuje, że estymator ATT ma rozkład normalny, ale tylko asymptotycznie, postanowiliśmy dokładnie przeanalizować uzyskane wyniki. W rezultacie zauważyliśmy, że przyczyną mniejszej dokładności oszacowań może być mniejsza liczba połączonych par. Klasyczny sposób odcięcia powoduje nieuwzględnianie ok. 10% par przy obliczaniu wartości ATT, a metoda dynamiczna odrzuca ok. 46% par, jako niedokładne odpowiedniki.

Tabela 1. Wyniki eksperymentu

Metoda	ATT	se(ATT)	liczba par
Propensity score	0,997	0,176	241
Propensity score z odcięciem	0,997	0,126	220
Propensity score z dynamicznym odcięciem	0,998	0,218	134

Źródło: obliczenia własne

Należy zaznaczyć, że dynamiczny sposób odcięcia odrzuca większą liczbę par, jednak nie należy rozpatrywać tego faktu w kategorii słabości, a zalety metody. Dzięki temu oszacowanie przeciętnego efektu jest uzyskiwane na podstawie lepiej zbilansowanej próby.

Aby rozstrzygnąć problem w sposób jednoznaczny postanowiliśmy przeprowadzić dodatkową symulację. Zbadaliśmy, w jaki sposób wariancja estymatora ATT jest uzależniona od liczby uzyskanych połączonych par. W tym celu dla prób o różnej liczebności przeprowadziliśmy proste łączenie 1 do 1 i obliczyliśmy dla każdej próby estymator ATT i jego odchylenie standardowe. Eksperyment powtórzyliśmy 100000 razy. Wyniki eksperymentu pomocniczego przedstawiono w tabeli 2³.

Ewidentnie widać, że wraz ze wzrostem liczby jednostek w grupie kontrolnej, która odpowiada liczbie połączonych par, błąd standardowy oszacowania maleje w tempie wykładniczym. Przedstawia do w innym świetle wyniki poprzedniego eksperymentu. Przy spadku liczby połączonych par z 241 do 134, 25% odchylenia standardowego można przypisać zmniejszonej liczbie par. W rezultacie należy

³ Z racji na szczupłość miejsca prezentujemy tutaj fragment wyników. Na prośbę osób zainteresowanych wysłamy pełne wyniki.

uznać, że odchylenie standardowe estymatora dynamicznym obciążeniem jest identyczne do odchylenia standardowego estymatora ATT. Tradycyjny mechanizm obciążenia przy tej liczebności próby redukuje odchylenie standardowe oszacowania o 30%. Zatem należy podjąć trudny wybór: czy lepsze jest bardziej precyzyjne oszacowanie o większej wariancji czy mniej precyzyjne o mniejszej wariancji. My opowiadamy się za tym pierwszym.

Tabela 2. Wariancja ATT a wielkość próby

Próba	ATT	se(ATT)	Eksperymentalna	Kontrolna
100	1.004	0.564	24	76
200	0.998	0.397	48	152
300	0.997	0.322	72	228
400	0.997	0.279	96	304
500	0.998	0.249	121	379

Źródło: obliczenia własne

Asymptotycznie wszystkie techniki łączenia danych oparte o wektory *propensity score* powinny dawać takie same rezultaty, gdyż wraz ze wzrostem liczebności próby metoda zbliża się do porównywania obserwacji dokładnie połączonych. Jednakże w małych próbach wybór algorytmu może ważyć na wyniku (Heckman i in. 1997). W takim przypadku należy dokonać wyboru między obciążeniem uzyskanego oszacowania efektu a jego wariancją. Nie ma jednego najlepszego sposobu estymacji sprawdzającego się w każdej sytuacji. Jakość poszczególnych technik i wybór estymatora powinny być uzależnione od statystycznych właściwości zbioru danych.

PODSUMOWANIE

W artykule przedstawiono estymatory przeciętnego efektu oddziaływania, oraz pokazano wykorzystanie techniki łączenia danych wykorzystującej wektor *propensity score*. Celem metody jest odtworzenie na podstawie danych przekrojowych warunków eksperymentalnych poprzez konstrukcję hipotetycznej grupy odniesienia (stanu kontrfaktycznego) pozwalającego na porównanie wyników działania programu z sytuacją, w której nie był on przeprowadzony. Odbywa się to poprzez taki wybór jednostek do grupy kontrolnej by była jak najbardziej zbliżona pod względem obserwowanych charakterystyk do grupy poddanej oddziaływaniu.

Zaprezentowano nową technikę oceny stopnia podobieństwa obserwacji, wykorzystującą zmodyfikowany mechanizm obciążenia. Modyfikacja polega na odnoszeniu wielkości obciążenia do wartości wektora *propensity score* dla jednostki, dla której szukany jest obiekt w stanie kontrfaktyczny. Następnie wykorzystując symulacje komputerowe zbadano właściwości statystyczne metody. Wykazano, że w przypadku znacznych różnic w rozkładach wektora *propensity score* między

grupą poddaną oddziaływaniu a grupą kontrolną dynamiczny mechanizm obciążenia pozwala uzyskać bardziej dokładne oszacowanie. Jednak ceną za mniejsze obciążenie estymatora jest większa wariancja, wynikająca z mniejszej liczby skutecznie połączonych par.

W dalszych pracach zamierzamy po pierwsze uwzględnić w eksperymencie próby o różnej liczebności. Tak jak wariancja estymatora ATT okazała się być nieliniową funkcją liczby połączonych par, także obciążenie estymatora może zależeć w nieliniowy sposób od liczebności próby. Po drugie, zamierzamy tak przeprojektować schemat generowania danych, aby zapewnić stabilną liczbę połączonych par, a przynajmniej chcielibyśmy osiągnąć większy stopień kontroli nad tym parametrem eksperymentu. Po rozwiązaniu tych problemów zamierzamy wykorzystać zmodyfikowaną procedurę w badaniach aplikacyjnych.

BIBLIOGRAFIA

- Blundell R., Costa-Diás M. (2000) "Evaluation Methods for Non-Experimental Data", *Fiscal Studies*, vol. 21/4, str. 427-468.
- Caliendo M., Kopeinig S. (2008) „Some Practical Guidance for the Implementation of Propensity Score Matching”, *Journal of Economic Surveys*
- Cochrane, Rubin (1973) "Controlling Bias in Observational Studies. A Review", *Sankhya*, vol. 35, str 417-466.
- Hahn J. (1998) "On the Role of the Propensity Score in the Efficient Semiparametric Estimation of Average Treatment Effects", *Econometrica*, vol. 66/2, str. 315-331.
- Heckman J., Ichimura H., Todd P. (1997) "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme", *The Review of Economic Studies*, vol. 64/4, str. 605-654.
- Hirano K., Imbens G., Ridder G. (2003) "Efficient Estimation of Average Treatment Effects Using the Estimated Propensity Score", *Econometrica* vol. 71/4, str. 1161-1189.
- Lee M-J. (2005) "Micro-Econometrics for Policy, Program, and Treatment Effects", Oxford University Press.
- Rosenbaum P., Rubin D. (1983) „The Central Role of the Propensity Score in Observational Studies for Causal Effects“, *Biometrika*, vol. 70/1, str 41-55.
- Rubin D. (1973) "Matching to Remove Bias in Observational Studies", *Biometrics*, vol. 29, str. 159-183.
- Smith J., Todd P. (2005) "Does Matching Overcome LaLonde's Critique of nonexperimental estimators?", *Journal of Econometrics*, vol. 125, str. 305-353.
- Strawiński P. (2007) "Przyczynowość, selekcja i endogeniczne oddziaływanie", *Przegląd Statystyczny* nr 4/2007, str. 49-61.

Matching with dynamic caliper. Preliminary results

Abstract: The evaluation problem is a problem of missing data. Henceforth, the important issue is construction of adequate counterfactuals. In this article modification of procedure that accepts observation as a good counterfactual is proposed. The proposed novelty is a change in the mechanism of observation comparison between treated and control group. In the last section the simulation results are presented. The simulation concerns statistical properties of proposed mechanism. The obtained results indicate that our method is capable of providing better results than techniques that are contemporary used.

Keywords: matching, propensity score, caliper, evaluation

KONIUNKTURA NA POLSKIM RYNKU KAPITAŁOWYM WOBEC SYTUACJI NA RYNKACH WIODĄCYCH ¹

Tadeusz Waściński

Katedra Organizacji i Inżynierii Produkcji SGGW
e-mail: twascinski@wp.pl

Grzegorz Przekota

Instytut Ekonomii i Zarządzania PK,
e-mail: gprzekota@wp.pl

Lidia Sobczak

Wydział Zarządzania PW,
e-mail: lidiasobczak@interia.pl

Streszczenie: W przedstawionym artykule „Koniunktura na polskim rynku kapitałowym wobec sytuacji na rynkach wiodących” podjęto próbę określenia siły i kierunku powiązań pomiędzy indeksem WIG Giełdy Papierów Wartościowych w Warszawie a indeksami wiodących giełd światowych: Stany Zjednoczone (S&P 500) oraz Japonia (NIKKEI 225). Analizie poddano dane dzienne wartości wybranych indeksów giełdowych za lata 2000-2008. Wykorzystano tutaj proste metody korelacyjne oraz model z Mechanizmem Korekty Błędem. Wyniki badań wskazują na pełne zintegrowanie polskiego rynku z rozwiniętymi rynkami zagranicznymi. Wysoka siła powiązania uwidoczniła się w ostatnich latach. Reakcja rynku polskiego na sytuację na rynkach zagranicznych następuje w czasie rzeczywistym. Podstawową przyczyną integracji rynków jest ich otwartość oraz rozwój technik komunikacji.

Słowa kluczowe: rynki kapitałowe, szeregi czasowe indeksów, model korekty błędów, analiza korelacji, modele transmisji

¹ Praca sfinansowana ze środków MNiSW Nr projektu N N112 120935

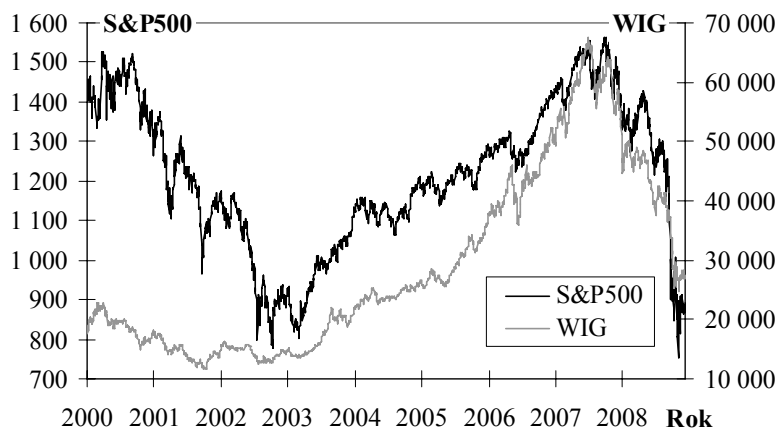
WSTĘP

Szybki postęp technologiczny ostatnich lat jest jedną z przyczyn integracji rynków kapitałowych. Jeśli dodatkowo połączyć to z coraz mniejszą ilością ograniczeń w przepływach kapitału można powiedzieć, iż rynki finansowe są otwarte na napływający kapitał. Dzięki temu potencjalny inwestor może w prosty sposób dokonać dywersyfikacji swojego portfela². Dywersyfikacja portfela na rynku międzynarodowym znacząco przyczynia się do minimalizacji ryzyka³.

Powiązania pomiędzy rynkami kapitałowymi mogą mieć charakter wspólnego trendu, dzieje się tak, gdy papierom o tym samym ryzyku towarzyszy ta sama oczekiwana stopa zwrotu. Fakt ten może dotyczyć nawet rynków o znacząco różnej wielkości. Oczekiwać można, iż impulsy z rynków większych będą kształtować sytuację na rynkach mniejszych⁴. Coraz nowsze badania wskazują na wzrost siły takich powiązań⁵.

W pracy starano się opisać powiązania pomiędzy indeksem WIG Giełdy Papierów Wartościowych w Warszawie a indeksami wiodących giełd światowych: Stany Zjednoczone (S&P 500) oraz Japonia (NIKKEI 225). Przedmiotem zainteresowania była przede wszystkim ocena, z którym z tych dwóch rynków polski rynek kapitałowy jest silniej powiązany, na ile stabilne w czasie są te powiązania oraz rozkład w czasie reakcji polskiego rynku na zmiany wartości indeksów na rynkach zagranicznych.

Rysunek 1. Szeregi czasowe indeksów S&P500 i WIG



Źródło: opracowanie własne na podstawie danych historycznych giełd.

² Porównaj: Solnik, [1974] s. 48-54.

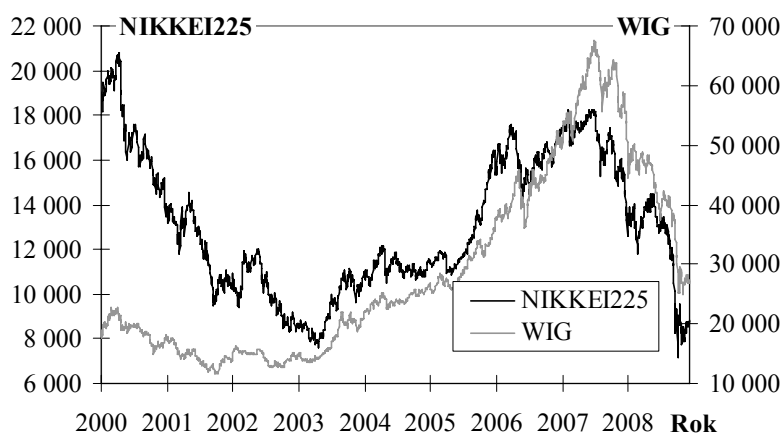
³ Zobacz: Obstfeld [1994] s. 1310-1329.

⁴ Porównaj: Forbes, Rigobon, [2002] s. 2223-2261.

⁵ Zobacz: Bekart, Harvey [1995] s. 403-444.

Analizie poddano dane dzienne wartości wybranych indeksów giełdowych za lata 2000-2008 (rysunek 1 i 2). Dane wybrano w ten sposób, aby notowanie z rynku amerykańskiego i japońskiego poprzedzało notowanie na rynku polskim, zakładając, że polski rynek jako mniejszy będzie reagował na zmiany na rynkach rozwiniętych. Ewentualne stwierdzenie integracji polskiego rynku kapitałowego z rynkami zagranicznymi lub też zaprzeczenie takiej ważnej jest w praktycznym punktu widzenia, gdyż może być skutecznie wykorzystywane w zakresie dywersyfikacji portfela.

Rysunek 2. Szeregi czasowe indeksów NIKKEI225 i WIG



Źródło: opracowanie własne na podstawie danych historycznych giełd.

UWAGI DOTYCZĄCE ZASTOSOWANEGO MODELU

Z punktu widzenia prawidłowej interpretacji parametrów modeli ekonometrycznych ważne jest, aby pomiędzy zmiennymi nie było korelacji pozornych. Procedura estymacji zaproponowana przez Englea i Grangera eliminuje z modelowania takie relacje. Kolejne jej etapy polegają na weryfikacji hipotez dotyczących stacjonarności i kointegracji zmiennych oraz etap końcowy budowa modelu.

Weryfikację stacjonarności szeregów czasowych wybranych indeksów giełdowych przeprowadzono za pomocą testu ADF w dwóch wersjach: bez stałej oraz ze stałą. Hipoteza zerowa brzmi: badany szereg jest szeregiem niestacjonarnym. Sprawdzianem hipotezy zerowej jest iloraz:

$$ADF = \frac{\hat{\delta}}{S(\hat{\delta})} \quad (1)$$

gdzie $\hat{\delta}$ jest oceną parametru δ w równaniu:

$$\Delta y_t = \delta y_{t-1} + \sum_{i=1}^k \delta_i \Delta y_{t-1} + \varepsilon_t \quad (2)$$

otrzymanym metodą MNK, a $S(\hat{\delta})$ jego błędem standardowym otrzymanym dla metody MNK. W równaniu (2) k jest taką ilością opóźnień, która pozwoli na wyeliminowanie autokorelacji. Odrzucenie hipotezy zerowej dla szeregu obserwacji y_t oznacza, iż jest to szereg stacjonarny i oznaczą się go $I(0)$, co oznacza zintegrowanie w stopniu 0.

Brak podstaw do odrzucenia hipotezy zerowej nie oznacza, że szeregu y_t nie można doprowadzić do szeregu stacjonarnego. Może on być zintegrowany w stopniu wyższym niż zero. W takim przypadku należy powyższą procedurę testowania zastosować dla przyrostów obserwacji. Odrzucenie hipotezy zerowej w takim przypadku oznacza, że szereg Δy_t jest stacjonarny, a szereg y_t zintegrowany w stopniu 1, co oznacza się $I(1)$.

Stacjonarność szeregu obserwacji y_t oznacza, iż modelowanie zależności powinno przeprowadzić się dla obserwacji. Natomiast stacjonarność pierwszych przyrostów oznacza, iż modelowanie zależności powinno przeprowadzać się właśnie dla przyrostów. W tym przypadku możliwe jest w jednym modelu pokazanie związków krótko- i długoterminowych. Jednak warunkiem tego jest kointegracja szeregów czasowych.

Dwa szeregi czasowe y_t i x_t są skointegrowane stopnia d, b ($I(d, b)$), jeżeli są zintegrowane tego samego stopnia d oraz istnieje liniowa kombinacja tych zmiennych, która jest zintegrowana stopnia $d-b$. Z praktycznego punktu widzenia najważniejsza jest sytuacja, w której zmienne są skointegrowane w stopniu (1,1). W takich sytuacjach można stosować twierdzenie Grangera⁶.

Twierdzenie Grangera. Jeżeli y_t i x_t są $I(1)$ oraz (y_t, x_t) są skointegrowane w stopniu (1,1), to y_t można przedstawić w postaci modelu z mechanizmem korekty błędem:

$$\Delta y_t = \alpha ECM_{t-1} + \sum_{i=1}^{k-1} \theta_i \Delta y_{t-i} + \sum_{i=0}^{k-1} \gamma_i \Delta x_{t-i} + \varepsilon_t \quad (3)$$

gdzie: ECM – reszty z równania kointegrującego.

Parametr α związany jest z szybkością dostosowywania zmiennej y_t do poziomu długookresowej równowagi z x_t , a parametry θ i γ związane są z dynamiką krótkookresową. Równanie 3 może dodatkowo zawierać stałą. Jeżeli brak jest kointegracji pomiędzy zmiennymi, to w równaniu 3 nie szacuje się parametru α .

⁶ Porównaj: Engle, Granger [1987] s. 1057-1072.

ANALIZA KORELACJI

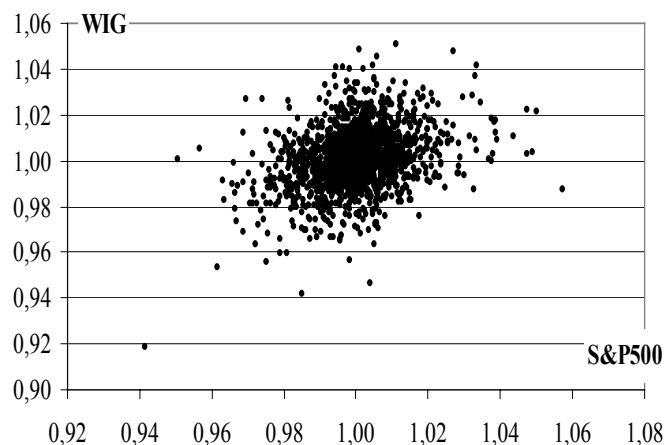
Na rysunku 1 i 2 przedstawiono szeregi czasowe badanych indeksów giełdowych w układzie WIG-S&P500 oraz WIG-NIKKEI225. Wykresy różnią się jedynie w pierwszej części do roku 2003, gdzie tempo spadku było różne. Od roku 2003 wykresy upodabniają się. Tempo wzrostu do połowy roku 2007 było podobne, a od połowy roku 2007 na wszystkich analizowanych rynkach obserwuje się podobne tempo spadku. Znajduje to swoje odzwierciedlenie w wyznaczonych współczynnikach korelacji. W całym analizowanym okresie związek pomiędzy poziomami indeksów w Polsce i w USA oraz Japonii był dodatni, lecz niezbyt silny, natomiast w okresie ograniczonym do lat 2003-2008 był dodatni bardzo silny (tabela 1). Z kolei związek pomiędzy wartościami indeksów S&P500 i NIKKEI225 zarówno w dłuższym jak i krótszym okresie czasu okazał się bardzo silny dodatni.

Tabela 1. Związki korelacyjne pomiędzy analizowanymi indeksami.

Korelacje	2000-2008		2003-2008	
	Poziomy	Stopy zwrotu	Poziomy	Stopy zwrotu
WIG ↔ S&P500	0,4466	0,3100	0,9356	0,2247
WIG ↔ NIKKEI225	0,5152	0,2618	0,9575	0,2537
S&P500 ↔ NIKKEI225	0,9005	0,3545	0,9043	0,4030

Źródło: obliczenia własne.

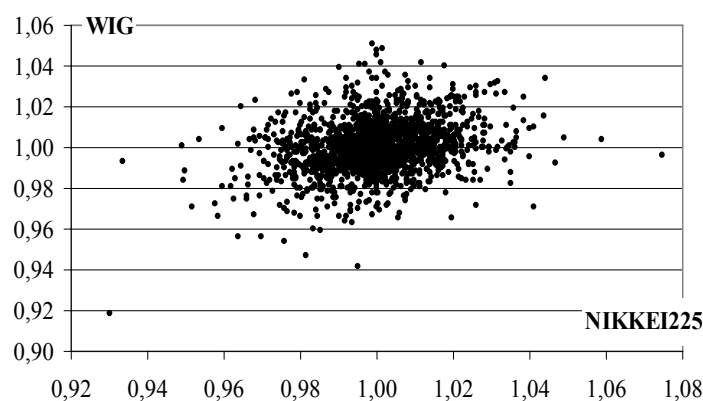
Rysunek 3. Związek korelacyjny stóp zwrotu indeksów WIG i S&P500



Źródło: wykonanie własne.

Inną sytuację obserwujemy w zakresie związku pomiędzy stopami zwrotu. Tutaj niezależnie od przyjętego okresu badania związek korelacyjny jest dodatni, jednak charakteryzuje się słabą siłą (tabela 1). Dotyczy to zarówno związku pomiędzy stopami zwrotu polskiego indeksu ze stopami zwrotu indeksów giełd amerykańskiej i japońskiej oraz związku pomiędzy stopami zwrotu z giełdy amerykańskiej i japońskiej. Wskazują na to wyznaczone współczynniki korelacji (tabela 1), a obrazują tę sytuację w dłuższym okresie czasu rysunki 3 i 4.

Rysunek 4. Związek korelacyjny stóp zwrotu indeksów WIG i NIKKEI225



Źródło: Opracowanie własne.

STACJONARNOŚĆ I KOINTEGRACJA

Zaobserwowane związki korelacyjne mogą być uznane za wiarygodne pod warunkiem występowania kointegracji. Z uwagi na zmianę właściwości powiązań pomiędzy wartościami indeksów giełdowych jaka nastąpiła w ostatnim okresie, czyli zdecydowanie silniejszą zależność w dalszych badaniach ograniczono się do okresu 2003-2008. Ponadto końcowym celem ma być model z Mechanizmem Korekty Błędem dlatego obliczenia prowadzono na szeregach danych zlogarytmowanych.

Szeregi czasowe poziomów indeksów giełdowych WIG, S&P500 oraz NIKKEI225 w okresie 2003-2008 okazały się szeregami niestacjonarnymi (tabela 2). Stacjonarne są pierwsze przyrosty tych indeksów. Wyniki takie uzyskano dla testu ADF bez stałej oraz ze stałą.

Równanie kointegrujące szacowano w postaci $y = a_1x + a_0$ (tabela 3). Parametry równania kointegrującego okazują się statystycznie istotne, ale przede wszystkim reszty z tych równań są stacjonarne. Czyni to wiarygodnym analizę korelacji, a ponadto umożliwia modelowanie związku z jednoczesnym uwzględnieniem zmian długookresowych i krótkookresowych w postaci modelu z mechanizmem korekty błędem.

Tabela 2. Testy stacjonarności

Indeks	ADF bez stałej				ADF ze stałą			
	I(0)		I(1)		I(0)		I(1)	
	ADF	p	ADF	p	ADF	p	ADF	p
WIG	3,6334	0,9999	-28,8661	0,0000	-0,3444	0,9157	-29,1875	0,0000
S&P500	1,9487	0,9881	-33,8092	0,0000	-1,3130	0,6254	-33,9163	0,0000
NIKKEI225	1,8613	0,9854	-31,5271	0,0000	-0,7282	0,8375	-31,6215	0,0000

Objaśnienia: I(0) – testowanie stacjonarności poziomów, I(1) – testowanie stacjonarności pierwszych przyrostów, ADF – wartość empiryczna testu, p – poziom istotności testu.

Źródło: obliczenia własne.

Tabela 3. Test ko integracji

Zmienna niezależna	Zmienna zależna: WIG			
	Model		ADF	
	a1 / poziom p	a0 / poziom p	ADF	poziom p
S&P500	2,7410 0,0000	-9,1090 0,0000	-2,5932	0,0093
NIKKEI225	1,6006 0,0000	-4,8355 0,0000	-3,6143	0,0003

Źródło: obliczenia własne.

Tabela 4. Modele transmisji – zmienna zależna przyrosty wartości indeksu giełdowego WIG

Zmienna niezależna x	ECM_{t-1}	$d(x_t)$	$d(x_{t-1})$	$d(y_{t-1})$	c
S&P500	-0,0068	0,3222	-0,0133	0,0484	0,0010
	0,0389	0,0000	0,7703	0,1323	0,0022
NIKKEI225	-0,0064	0,2315	0,0166	0,0449	0,0010
	0,1057	0,0000	0,5752	0,1588	0,0028

Objaśnienia: górna liczba w komórce – parametr, dolna liczba w komórce – poziom istotności.

Źródło: obliczenia własne.

Według uzyskanych wartości modelowych (tabela 4) istotny wpływ na przyrosty indeksu giełdowego WIG w krótkim okresie czasu wywierają jedynie bieżące przyrosty wartości indeksów giełd zagranicznych. Okazuje się, iż w okresie 2003-2008 zmiana wartości indeksu S&P500 o 1% na otwarciu gene-

rowała średnią zmianę wartości indeksu giełdowego WIG na zamknięciu 0,3222%, natomiast zmiana wartości indeksu NIKKEI225 o 1% na zamknięciu generowała średnią zmianę wartości indeksu giełdowego WIG na zamknięciu o 0,2315%. Wpływ zmian wartości indeksów zagranicznych z dnia poprzedniego oraz zmian indeksu WIG z dnia poprzedniego na bieżącą zmianę wartości indeksu WIG okazuje się nieistotny. Stosunkowo niskie wartości współczynnika ECM informują, iż wartości indeksu giełdowego WIG stosunkowo wolno dostosowują się do długookresowej równowagi z wartościami indeksów giełd zagranicznych. W ciągu jednego dnia redukowana jest tylko niewielka część odchylenia od stanu równowagi.

PODSUMOWANIE

Przeprowadzona analiza wskazuje, że polski rynek w ciągu ostatnich kilku lat stał się rynkiem w pełni zintegrowanym z rozwiniętymi rynkami zagranicznymi. Koniunktura panująca na rynkach zagranicznych ma swoje bezpośrednie przełożenie na koniunkturę na rynku polskim. Siła oddziaływania rynku amerykańskiego i japońskiego jest zbliżona. Warto jednak zwrócić uwagę na fakt, iż rynek polski niezbyt szybko dostosuje się do długookresowej równowagi z rynkami zagranicznymi. Z punktu widzenia inwestora taka sytuacja oznaczać może osiągnięcie pewnych korzyści z dywersyfikacji.

Tak silny wpływ rynków zagranicznych na polski rynek giełdowy bezsprzecznie wiązać można z szybkim postępem technologicznym ostatnich lat. W sytuacji gdy dane z giełd zagranicznych docierają na rynek polski w czasie rzeczywistym musi następować natychmiastowa reakcja na nie. I tak też się dzieje. Informacje z dnia poprzedniego nie odgrywają tak silnej roli jak właśnie informacje bieżące. Napływające informacje rynek dyskontuje natychmiast. Efektem tego są wspólne trendy rynków światowych, które oczywiście nie wykluczają osiągania korzyści z dywersyfikacji. Poza reakcją, o której mowa, każdy rynek ma jednak swoją wewnętrzną specyfikę.

LITERATURA

- Bekart G., Harvey C., Time varying world market integration, *Journal of Finance*, No 50, 1995, s. 403-444.
- Engle R.F., Granger C.W., Cointegration and error correction: representation, estimation, testing, *Econometrica*, No 55, 1987, s. 1057-1072.
- Forbes K.J., Rigobon R., No contagion, only interdependence: measuring stock market comovements, *Journal of Finance*, Vol. 7, No 5, 2002, s. 2223-2261.
- Obstfeld M., Risk taking, global diversification and growth, *American Economic Review*, No 84, 1994, s. 1310-1329.
- Solnik B., Why not diversify internationally?, *Financial Analysts Journal*, No 20, 1974, s. 48-54.

Conditions of polish capital market against the situation on leading markets

Abstract: The paper presents an attempt towards determination of the strength and direction of relations between the WIG index of Stock Exchange in Warsaw and the indexes of leading world stock exchanges in the United States (S&P 500) and Japan (NIKKEI 225). The daily data values of selected stock exchange indexes for the period 2000-2008 were analyzed, using the simple correlation methods and a model with Error Correction Mechanism. The results point out at full integration of Polish market with the developed foreign markets. The high strength of relations has occurred in recent years. The Polish market response to situation on foreign markets occurs in real time. The basic cause of market integration is their open nature and development of communication techniques.

Keywords: capital markets, temporary rows of indexes, Error Correction Model (ECM), coefficients of correlation, models of transmission

BADANIA EFEKTYWNOŚCI PREDYKCYJNEJ MODELU KLASY GMDH OPARTEGO NA ZACHOWANIU UCZESTNIKÓW RYNKU WALUTOWEGO

Antoni Wiliński

Wydział Informatyki, ZUT
e-mail: awilinski@wi.zut.edu.pl

Streszczenie: W pracy przedstawiono wyniki eksperymentu przeprowadzonego w celu predykcyjnym na rynku walutowym. Założono, że rynek nie jest efektywny i daje się z przeszłości wyekstrahować wiedzę o błędach popełnionych przez uczestników wybranej platformy brokerskiej. Dla wykonania predykcji wykorzystano zmodyfikowaną metodę GMDH (Group Method of Data Handling) umożliwiającą sukcesywny wybór nieliniowego modelu wielomianowego najlepiej w danym kroku opisującego rynek. Przedstawiono interesujące wyniki eksperymentu na danych historycznych potwierdzające użyteczność metody. Danymi wejściowymi były zarejestrowane na platformie zachowania inwestorów – rozkład otwartych pozycji i złożonych zleceń. Stąd – rozpatrywane podejście można zaliczyć do modelowania behawioralnego.

Słowa kluczowe: zachowanie inwestorów, finanse behawioralne, modele regresyjne, GMDH, platformy brokerskie, rynek efektywny

WPROWADZENIE

Wśród zagorzałych zwolenników analizy technicznej powszechne jest przekonanie, że każdy obserwowany rynek pozwala przewidzieć jego zmiany na podstawie bogatego arsenału środków umożliwiających wydobycia sugestii zmian z przeszłości. Na drugim biegunie poglądów lokują się hipotezy Fama'y [Fama E.F. 1998] o rynku efektywnym niemożliwym lub prawie niemożliwym do prognozowania wskutek pełnego i natychmiastowego dostępu do informacji przez wszystkich uczestników rynku. Problematyką kojarzenia wszelkich informacji, pochodzących z różnych niezależnych źródeł i mających często heterogeniczną, nieporównywalną naturę zajmowało się wielu analityków. Z pewnością najważniejszym jest

niedawny kandydat do nagrody Nobla wspomniany prof. Eugene Fama, ale także kontestujący jego prace noblista Joseph Stiglitz, czy zwolennicy finansów behawioralnych R.Barberis [Barberis N i in. 1998], R. Thaler, W. DeBondt [DeBondt W. i in. 1985] i J.Piper [Piper J. 2006].

Dość powszechne są także ostatnio poglądy, że najlepsze prawdopodobnie wyniki predykcyjne na poziomie zapewniającym ich praktyczną użyteczność inwestycyjną uzyska się poprzez kojarzenie elementów analizy technicznej i fundamentalnej. Sposobem na osiągnięcie tego celu jest hybrydyzacja metod o różnym charakterze modeli opisujących rzeczywistość związaną z danym rynkiem i różnych technik wykonania decyzji rekomendowanych przez te modele. Sheng-Hsun Hsu [Sheng-Hsun Hsu i in. 2009] i Zhang Yudong [Zhang Yudong i in. 2009] łącząc modelowanie za pomocą sztucznych sieci neuronowych z teorią zbiorów przybliżonych uważają, że to najbardziej perspektywiczny kierunek opanowywania niestacjonarnych procesów, a Hia Jong Teoh [Hia Jong Teoh i in. 2009] przedstawia metodę łączącą sieci neuronowe z GMDH (*Group Method of Data Handling*). S. Dehuri i S. Cho [Dehuri S. 2009] łączą techniki regresyjne z optymalizacją rojem cząstek.

Na tym tle zaprezentowana jest metoda o podobnej próbie połączenia metody GMDH [Ivakhnenko A. i in. 1999, Ivakhnenko A. i in. 2000] z danymi o charakterze behawioralnym wyekstrahowanymi z przestrzeni zachowań uczestników rynku do tablicy przypominającej tablicę informacyjną w teorii zbiorów przybliżonych. Interesujące może być rozpatrywanie bieżących zachowań tych uczestników, nie poprzez obserwację kursu waluty (zresztą to niemożliwe), lecz poprzez badanie przyczyn zmian tego kursu wynikających ze złożonych zleceń i otwartych pozycji na rynku spot. W artykule rozpatrywana jest koncepcja systemu wykorzystującego statystyczne dane o rozkładzie preferencji użytkowników przetworzone następnie w modelu regresyjnym zbudowanym wg zasad GMDH.

CHARAKTERYSTYKA DANYCH

Przykładem platformy prezentującej behawioryzm jej użytkowników jest aplikacja Small Speculant Sentiment Analyzer [Oanda 2008, Wanadoo 2008] przedstawiająca intencje wybranej części użytkowników znanej platformy brokerskiej Oanda.com [Oanda 2008]. Wspomniana aplikacja wizualizuje zamiary kilkudziesięciu tysięcy użytkowników w formie dwóch diagramów, z których górny przedstawia rozkład otwartych pozycji rynku spot, a dolny rozkład zleceń (rys.1). Każdy kwadrant diagramu (krzywa na nim przedstawiona) może być zinterpretowany jako suma liczby decyzji (otwartych pozycji lub zleceń) w funkcji odległości od bieżącej wartości rozpatrywanego instrumentu finansowego - na przedstawionym wykresie na przykładzie eurodolara.

Kolejno, np. diagram w górnej prawej ćwiartce (kwadrancie) oznacza narastającą liczbę (względna, wyrażoną w procentach) pozycji długich otwartych powyżej bieżącej wartości kursu. Wykres przedstawia sytuację na rynku EUR/USD

przy wartości bieżącej kursu $C = 1,5305$. Każda wartość (w %) pozycji długich otwartych w przedziale $\langle x_b, x_b + \delta \rangle$ w całkowitej liczbie wszystkich otwartych pozycji (długich i krótkich). Niech δ oznacza pewną, racjonalnie wybraną, stałą wielkość rozpatrywanego przedziału (np. 100, 200 lub więcej pipsów). Zmienna ta winna umożliwić w miarę dokładne odczytanie udziału otwartych pozycji z wykresu (rys.1).

Przez analogię, odpowiednio (rys. 2):

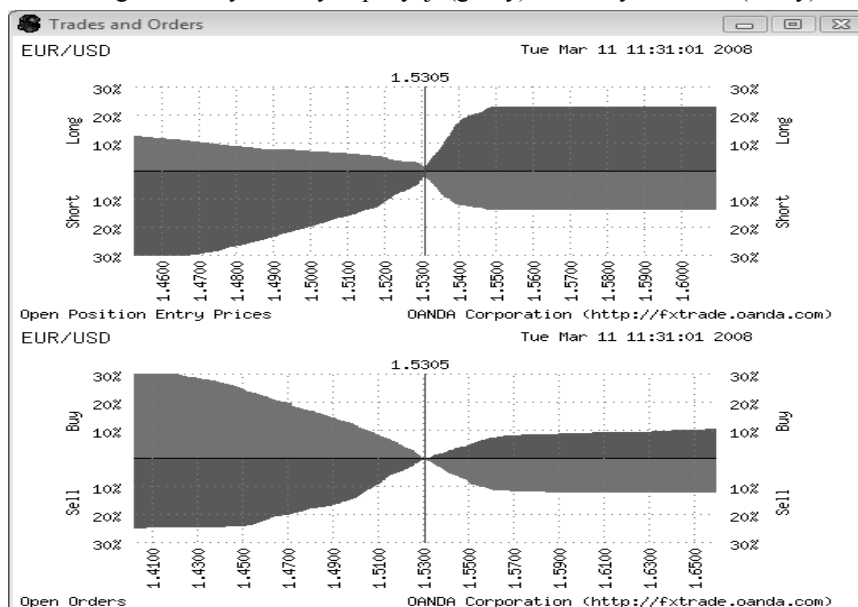
$G_2(x_b, x_b + \delta)$ - oznaczać będzie udział pozycji krótkich otwartych powyżej wartości bieżącej kursu w przedziale $\langle x_b, x_b + \delta \rangle$;

$G_3(x_b - \delta, x_b)$ - oznaczać będzie udział pozycji krótkich otwartych poniżej wartości bieżącej kursu w przedziale $\langle x_b - \delta, x_b \rangle$;

$G_4(x_b - \delta, x_b)$ - oznaczać będzie udział pozycji długich otwartych w przedziale $\langle x_b - \delta, x_b \rangle$, czyli poniżej wartości bieżącej kursu.

Wprowadzenie parametru δ powoduje, że suma czterech rozpatrzonych sum G_1 do G_4 nie będzie równa 100%, lecz pozwoli na powtarzalność obliczeń pól w tych samych warunkach.

Rysunek 1. Diagram liczby otwartych pozycji (górny) i złożonych zleceń (dolny)



Źródło: Oanda, www.oanda.com

Podobnie dla nierozpatrywanego do tej pory diagramu dolnego na rys. 1 niech odpowiednio:

$G_5(x_b, x_b + \delta)$ - oznacza udział zleceń kupna złożonych z limitem położonym w przedziale $\langle x_b, x_b + \delta \rangle$. Zlecenia te mogą wynikać z zamian otwarcia pozycji długiej powyżej bieżącej wartości kursu (są to tzw. zlecenia Buy Stop), lub z zamiaru zamknięcia przegrywającej pozycji krótkiej (zlecenie Stop Loss). Oba te przypadki są rezultatem takiej samej postawy inwestora – jego przekonania, że kurs waluty wzrośnie. W pierwszym przypadku otwarcie pozycji długiej następuje w nadziei dalszego wzrostu, w drugim przypadku zamknięcie pozycji krótkiej (technicznie tożsame z otwarciem pozycji długiej) następuje w obawie także przed dalszym wzrostem kursu i powiększaniem straty.

$G_6(x_b, x_b + \delta)$ - oznacza udział zleceń sprzedaży złożonych z limitem położonym powyżej bieżącej wartości kursu. Są to zlecenia wynikające z zamiaru zamknięcia pozycji długiej wygrywającej (Take Profit) lub otwarcia pozycji krótkiej. Oba wynikają z takiego samego przekonania inwestorów o spadku kursu (po osiągnięciu poziomu wyznaczonego w zleceniu). Inwestor w przypadku posiadania pozycji długiej składa zlecenie typu Take Profit lokując je powyżej bieżącej wartości kursu w przekonaniu że kurs dalej nie będzie rósł. Inny inwestor składa zlecenie otwarcia pozycji krótkiej Sell Limit także w przekonaniu, że kurs po osiągnięciu poziomu określonego w zleceniu zacznie spadać.

$G_7(x_b - \delta, x_b)$ - to udział zleceń sprzedaży złożonych poniżej wartości bieżącej kursu. Zlecenia te wynikają z zamiaru inwestora otwarcia pozycji krótkiej (Sell Stop) lub zamknięcia przerywającej pozycji długiej (Stop Loss). Oba więc wynikają z przekonania, że kurs będzie spadał.

$G_8(x_b - \delta, x_b)$ - to udział zleceń kupna złożonych poniżej wartości bieżącej kursu. Zlecenia takie są składane przez inwestora w celu otwarcia w celu otwarcia pozycji długiej (Buy Limit) lub dla zamknięcia krótkiej pozycji wygrywającej (Take Profit). Oba typy zleceń wynikają z tego samego przekonania inwestorów, że kurs po osiągnięciu poziomu wyznaczonego przez zlecenie, będzie następnie rósł.

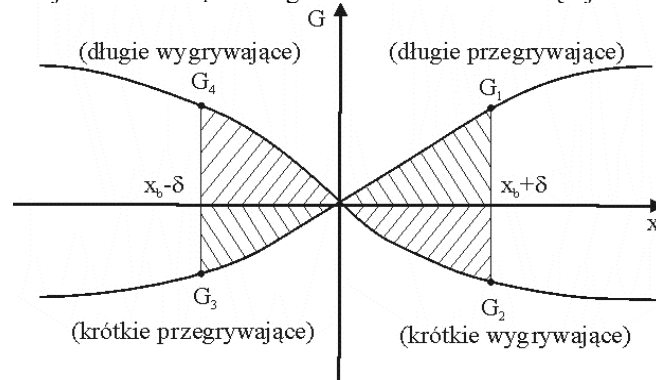
Całki G_5 do G_8 dotyczące zleceń przedstawione są na rys. 3. Przedział δ (odległość od wartości bieżącej kursu) może być oczywiście inny niż na rys.2.

Jeżeli w powyższych definicjach G_i oraz δ wprowadzić pojęcie małego np. kilkupipsowego przyrostu kursu d_x nie mniejszego jednak niż 1 pips oraz pojęcie rozkładu liczby zdarzeń g_j w każdym j -tym przedziale wyznaczonym przez kolejne przyrosty kursu (zdarzenie to zlecenie lub otwarcie pozycji), to ogólna definicja sumy G_i będzie:

$$G_i = \sum_j^k g_j \quad (1)$$

gdzie $k = \delta / d_x$ dla $i = 1, 2 \dots 8$; dla $\delta = 100, 200, 300 \dots$ pipsów.

Każdy indeks i oznacza zupełnie inną funkcjonalność, inny świat - zwycięstw lub porażek, optymizmu lub pesymizmu.

Rysunek 2. Definicja udziałów G_i w odległości δ od wartości bieżącej kursu x_b 

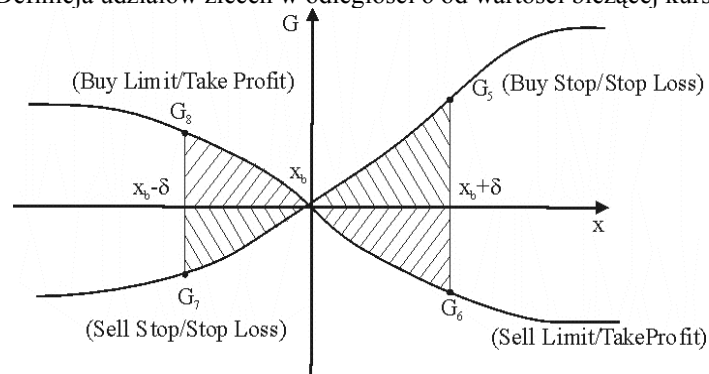
Źródło: Opracowanie własne

Dla pewnego ustalonego arbitralnie, aczkolwiek rozsądnie δ , można przeprowadzić normalizację sum G_i w dwóch grupach - dla $i \in \{1,2,3,4\}$ oraz $i \in \{5,6,7,8\}$, taką, że

$$G_i^n = \frac{G_i}{\sum_1^4 G_i}, \text{ dla } i = 1, 2 \dots 4 \quad (2)$$

$$G_i^n = \frac{G_i}{\sum_5^8 G_i}, \text{ dla } i = 5, 6 \dots 8 \quad (3)$$

Zabiegi powyższe mają na celu wyodrębnienie takiego fragmentu obrazu (rys.1), który w sposób powtarzalny, dla wielu tysięcy pobrań z serwera, zapewni zawsze tę samą procedurę znajdowania proporcji pomiędzy otwartymi pozycjami

Rysunek 3. Definicja udziałów zleceń w odległości δ od wartości bieżącej kursu x_b 

Źródło: Opracowanie własne

w czterech kwadrantach górnej części rys.1 oraz złożonych zleceń w dolnej części tego rysunku a dodatkowo zapewni stałą wartość czterech sum.

Po przeprowadzeniu normalizacji wg (2) i (3) przygotowano do dalszego przetwarzania macierzy danych obserwowanych $X_{M \times N}$ o następującej strukturze:

M - liczba wierszy (pomiarów, obserwacji – tu świec obserwowanych w cyklu jednogodzinnym);

N - liczba kolumn; tu $N = 10$.

W kolumnach macierzy X zapisano odpowiednio:

X_1 - kurs badanej co godzinę pary walutowej EUR/USD;

$X_2 \div X_5$ - kolejno obserwowane na podstawie górnego diagramu (rys.1) i przetworzone zgodnie z (2) wartości udziałów $G_1^4 + G_4^4$, oznaczające względną liczbę pozycji otwartych w rozpatrywanych kwadrantach układu rys.2;

$X_6 \div X_9$ - obserwowane na podstawie dolnego diagramu (rys. 1) i przetworzone wg (3) wartości udziałów złożonych zleceń $G_5^4 + G_6^4$, oznaczające względną liczbę w kwadrantach rys.3.

$X_{10} \equiv Y$ – obserwowana zmienna wyjściowa - różnica pomiędzy bieżącą a poprzednią wartością kursu, stanowiąca podstawę do rozważań predykcyjnych.

Formalnie, $X_{i,10} = X_{i+1,1} - X_{i,1}$. Tabele te (macierz danych X) są bazą dla dalszych rozważań i zbudowania modelu predykcyjnego.

SFORMUŁOWANIE PROBLEMU BADAWCZEGO

Przedstawiona aplikacja internetowa jest, zgodnie z jej sugerowaną nazwą (Small Speculant), źródłem wielu pomysłowych rozwiązań i miejscem ścierania się różnych poglądów spontanicznie powstającej wokół niej społeczności. Wiele z nich koncentruje się, jak to wynika z dyskusji na forum, na próbie realizacji tezy, że skoro ok. 95 % wszystkich drobnych inwestorów przegrywa, to należy poszukiwać wzorca rozkładu udziałów $G_1 \div G_8$ takiego, że prawdopodobna i możliwa będzie porażka tak znacznej części populacji inwestorów reprezentowanych poprzez bieżący rozkład otwartych pozycji i zleceń.

Rozpatrywana tu autorska koncepcja jest całkowicie odmienna od wyżej przedstawionej. Zbieżne są one jedynie w aspekcie założenia, że istnieją pewne wzorce wyrażane poprzez konfiguracje danych, których rozpoznanie umożliwi lub uprawdopodobni skuteczność prognozy. W przypadku, opisanym jako popularny sposób zachowania się większości inwestorów, wzorcem takim mogłaby być sytuacja na diagramie rys.1 – jako informacja wejściowa dla znalezienia odpowiedzi na pytanie, jaka możliwa najbliższa zmiana rynku spowodowałaby porażkę większości graczy. Na diagramie tym dla większości inwestorów gorszy byłby wzrost kursu – spowodowałby on pogłębienie się strat ok. 35% populacji z tytułu otwartych pozycji krótkich i poprawiałby sytuacje ok. 24% przegrywających długich pozycji. W diagramie złożonych zleceń ruch byłby o wiele mniejszy niż dla przypadku otwartych pozycji.

Ponieważ historycznie dzieje się tak, że większość przegrywa, to logicznym (w sensie myślenia indukcyjnego, a nie logicznego) ruchem byłoby wykonanie operacji (np. otwarcie pozycji), wygrywającej przy równoczesnej porażce większości. Na zasadzie przekory należałoby więc otworzyć tu pozycję długą.

W rozpatrywanym modelu zakłada się także występowanie wzorców uzasadniających postawienie następującej tezy. Jeżeli wielokrotnie weryfikowana konfiguracja zmiennych wejściowych $X_2 \div X_9$ pozwala na zbudowanie modelu regresyjnego (niekoniecznie liniowego) takiego, że wg ustalonego kryterium akceptowalne mogą być błędy uczestników wynikające z zastosowania tego modelu na odcinku predykcyjnym w fazie uczenia, to oznaczać to może, że znaleziono wzorzec (model regresyjny), na podstawie którego można zbudować system transakcyjny (oparty na koncepcji prognozy).

Jak zwykle, ważne jest rozróżnienie modelu predykcyjnego od modelu systemu transakcyjnego (inwestycyjnego). Pierwszy jest wyrazem zastosowania wyrafinowanych metod analizy technicznej (badania skutków zmian rynku) w celu prognostycznym, a drugi jest zastosowaniem (wykorzystaniem) wyników pierwszego modelu w warunkach proponowanych przez konkretną platformę inwestycyjną.

Podobnie jak w pracach [6,15,16] do rozwiązania zadań prognostycznych użyta była metoda GMDH. Zastosowany był tu algorytm oparty na nieliniowej regresji [Wilinski A. 2008], w której zmienną obserwowaną był zwrot po zakończeniu kolejnej świecy a zmiennymi wejściowymi kurs waluty i osiem opisanych sum G_i , $i = 1, 2, \dots, 8$.

Przez zwrot rozumie się tu różnice (w pipsach) pomiędzy wartościami kursu w chwili bieżącej i poprzedniej, oddalonymi o okres jednej świecy.

Zanim przystąpiono do zbudowania modelu regresyjnego rozszerzona została macierz danych wejściowych poprzez konkatenację horyzontalną. Macierz dodana z prawej strony macierzy pierwotnej utworzona została z nowych kolumn o długości N , z których każda jest iloczynem wartości w kolumnach z danymi pierwotnymi. Tych kolumn z danymi pierwotnymi uwzględniono 8, gdyż rozważano tylko kolumny reprezentujące sumy G_1 do G_8 .

W związku z tym kolejne kolumny o indeksie większym o liczbę k od dotychczasowej liczby kolumn N kolejno zawierały:

$$X_{N+k} = X_i * X_j, \text{ dla } k = 1, 2, \dots, i = 2, 3 \dots 9, j = i+1, i+2, \dots, 9, \quad (4)$$

gdzie operator „*” oznacza mnożenie tablicowe wektorów wyraz po wyrazie, tzn. wynikiem mnożenia jest kolejny wektor o takim samym wymiarze, którego wyrazy są iloczynami odpowiednich wyrazów dwóch kolumn wejściowych (kolejno dodawanymi iloczynami Hadamarda dwóch kolumn macierzy pierwotnej).

Liczba dodatkowych kolumn wg (4) z uwzględnieniem kwadratów wartości kolumn pierwotnych wyrazów $N_d = N(N-1)/2 + N$. Razem po konkatenacji macierz danych ma wymiar $M \times (N + N_d)$.

Dla zbudowania modeli prognostycznych na podstawie tak przetworzonych danych pierwsza kolumnę (X_1) traktuje się jako zmienną obserwowaną Y (odnosi się ona do zwrotu po kolejnym kroku) a wszystkie pozostałe kolumny - jako zmienne wejściowe (objaśniające). W ten sposób ostateczny model opisujący np. K wierszy zmiennej wyjściowej miał postać formalnie liniową (w rzeczywistości – kwadratową):

$$\hat{y}_{t+1} = \mathbf{X}_t * \mathbf{A}_{i+1} \quad (5)$$

$\begin{matrix} \text{[1x1]} & \cdot & \text{[(L+1)x1]} & \cdot & \text{[(L+1)xL]} \end{matrix}$

dla $L = 3, 4, 5, \dots$ $i = 1, 2, \dots$

W modelu tym zmienna \hat{y}_{t+1} oznacza predykcyjną modelowaną wartość zmiennej obserwowanej, która w chwili bieżącej wynosi y_i . \mathbf{X}_i jest wektorem $(L+1)$ zmiennych wejściowych wybranym wg opisanego dalej algorytmu spośród N_d zmiennych rozszerzonej o wyrazy kwadratowe macierzy \mathbf{X} zaobserwowanym w chwili bieżącej i . Wektor \mathbf{A}_{i+1} jest zestawem $(L+1)$ współczynników równania liniowego

$$\mathbf{A}_{i+1} = \left[\mathbf{a}_{i+1}^0 \quad \mathbf{a}_{i+1}^1 \quad \mathbf{a}_{i+1}^2 \quad \dots \quad \mathbf{a}_{i+1}^L \right] \quad (6)$$

obliczonym w przeszłości na podstawie dostępnych danych historycznych dla wierszy danych rozpoczynających się w chwili $i-K-1$ a kończących się w chwili $i-1$.

Ponieważ zakłada się, że jest dana informacja do chwili i -tej to wykorzystując dane wejściowe do chwili $i-1$ a zmienną obserwowaną do chwili i -tej uzyskuje się możliwość wykonania prognozy o jeden krok wprzód. Liczba K obserwowanych wierszy jest (na razie) przyjmowana arbitralnie i może być także przedmiotem dalszej optymalizacji. Wektor \mathbf{A}_{i+1} obliczany jest z równania regresji wykonanej na K wierszach i L kolumnach.

Niewiadomą w tych rozważaniach jest liczba L współczynników równania jak i struktura modelu (czy do równania wchodzi tylko wyrazy liniowe, kwadratowe czy model jest mieszaniną tych wyrazów).

Problem ten rozwiązuje algorytm oparty na idei Group Method of Data Handling [6,15] dokonujący dla każdego i -tego kroku wielokrotnego wyboru L wyrazów dla $L=3,4,5,\dots,N_d$

Dla każdego obliczanego modelu rozpatrywane jest kryterium dokładności odtworzenia przebiegu zmiennej obserwowanej jako:

$$K_{iL}^J = \frac{1}{K} \sum \left\| \hat{Y}_t - Y_t \right\|^2 \rightarrow \min \quad (7)$$

Indeksy w kryterium K_{iL} oznaczają:

L - jest to indeks wskazujący liczbę wyrazów w modelu - dla modelu zawierającego $(L+1)$ wyrazów (L kolumn danych oraz wektor jedynek przeznaczony w modelu do obliczenia wyrazu wolnego równania regresji);

$J = \{j_1, j_2, \dots, j_L\}$ - określa zbiór indeksów kolumn włączonych do macierzy informacyjnej tworzącej wejścia dla obliczenia modelu \hat{Y}_i .

W ten sposób obliczając L_{1L}^J dla wszystkich kombinacji C_{Nd}^L dla każdego $L = 3, 4, 5, \dots$ znajdujemy K_{1Lmin}^J z odpowiadającym mu wektorem L indeksów kolumn $\{j_1, j_2, \dots, j_L\}$.

Tak znaleziony model użyty jest do wykonania predykcji (5). Jest to predykcja wykonana jeden krok wprzód przy założeniu, że model \hat{Y}_i uznany za najlepszy wg kryterium (11), będzie skuteczny jeszcze jeden krok wprzód i da satysfakcjonujące przybliżenie \hat{y}_{i+1} wg (5). Gdy upłynie czas trwania jednego kroku, będzie można sprawdzić, czy wartość y_{i+1} (której w danej chwili jeszcze nie ma) będzie bliska modelowi \hat{y}_{i+1} .

Wielokrotne powtarzanie tej procedury i ocena jakości modeli \hat{y}_{i+1} była przedmiotem badań.

Należy przy tym zauważyć, że w każdym kroku model \hat{Y}_i budowany jest z użyciem innych zmiennych i z zastosowaniem różnej struktury modelu. Posiada także, co oczywiste, zupełnie inny zestaw współczynników A_{i+1} .

W identyczny sposób można zbudować model (modele) umożliwiające predykcję na dwa kroki wprzód (także więcej). Jeżeli bowiem w jakiejś chwili t_i wykonana jest predykcja na dwa kroki wprzód, czyli dla chwili t_{i+1} i t_{i+2} to również w tej samej chwili znana jest predykcja z poprzedniego kroku dotycząca chwili t_i oraz t_{i+1} . Tak więc w chwili t_i istnieją dwie predykcje dla chwili t_{i+1} . Niech będą one oznaczone odpowiednio p_{i+1}^i i p_{i+1}^{i-1} , gdzie indeks dolny oznacza, dla której chwili predykcja jest wykonana a indeks górny, w której chwili została wykonana. Z różnicy indeksów widać, że pierwsza predykcja p_{i+1}^i to predykcja jeden krok wprzód a predykcja p_{i+1}^{i-1} wykonana jest dwa kroki wprzód.

Przez predykcję niech będzie rozumiane odchylenie zmiennej obserwowanej od wartości w kroku poprzednim np.

$$p_{i+1}^i = \hat{y}_{i+1} - y_i$$

$$p_{i+1}^{i-1} = \hat{y}_{i+1} - \hat{y}_i \quad (8)$$

lecz także

$$p_{i+1}^{i-1} = \hat{y}_i - y_{i-1} \quad (9)$$

Ponieważ w systemach transakcyjnych najważniejsza jest trafność kierunku zmian zmiennej obserwowanej, a nie dokładne określenie jej wartości, to w dalszych rozważaniach istotne będzie, czy zachodzi warunek równości znaków:

$$\text{sign}(p_{i+1}^i) = \text{sign}(p_{i+1}^{i-1}) \quad (10)$$

Jeżeli warunek (10) zachodzi, to interesująca staje się wartość znaku.

Jeżeli $\text{sign}(p_{i+1}^i) = 1$, oznacza to, że prognozowany jest wzrost rozpatrywanego kursu (tu ceny pary walutowej EUR/USD) i należy otworzyć pozycję długą, a gdy $\text{sign}(p_{i+1}^i) = -1$, należy otworzyć pozycję krótką. Jeżeli warunek (16) nie był spełniony, nie wykonywano żadnej operacji.

WYNIKI BADAŃ

Dla tak zdefiniowanej strategii z użyciem modelu GMDH przeprowadzono badania z wykorzystaniem danych zgromadzonych w wyniki dwumiesięcznego automatycznego monitoringu platformy [Wanadoo 2008]. Każda otwarta pozycja była zamykana po jednym kroku. Przyjęto, że spread wynosi 0.0002 (dwa pipsy).

Wykonano wiele eksperymentów dla różnych wartości K (liczby wierszy rozpatrywanych wstecz w modelu regresyjnym), SL (Stop Loss wyrażanego w pipsach dla ograniczenia strat) i różnych horyzontów czasowych. Rozpatrywano także różną liczbę okresów predykcji (od 1 do 3). Jakość systemu transakcyjnego (poprawność doboru parametrów) szacowano za pomocą uproszczonego wskaźnika Sharpe'a (pozbawionego członu odpowiedzialnego za stopę zwrotu wolną od ryzyka).

Niemal zawsze uzyskiwane były wyniki mające praktyczną wartość transakcyjną. Na wykresie przedstawiono wyniki jednej spośród tych symulacji. Wybrano warunki trudne – $SL=0.0050$, co dla świec jednogodzinnych na EUR/USD jest poziomem trudnym do osiągnięcia, ale osiąganym (tu 16 razy). Liczba członów modelu nieliniowego L wahała się od 3 do 4. Dla obliczenia współczynników modelu aproksymującego wg GMDH rozpatrywano $K=45$ wierszy wstecz od chwili bieżącej.

Dla 800 kolejnych kroków uzyskano modelowaną zgodność predykcji w pierwszym i drugim kroku w 429 przypadkach, spośród których 237 (55,2%) otwartych pozycji dało sukces a 192 porażkę.

Po każdym k kroku mierzono zwrot (w pipsach).

$$z_k = (y_k - y_{k-1}) \cdot \text{sign}(p_k^{k-1}) \quad (11)$$

Zapis (11) oznacza, że wartość dodatnią (sukces) osiągnano wówczas, gdy przy otwartej pozycji długiej ($\text{sign}(p_k^{k-1}=1)$) nastąpił wzrost kursu lub przy otwartej pozycji krótkiej ($\text{sign}(p_k^{k-1} = -1)$) nastąpił spadek.

Przedstawiony na rys 4. wykres skumulowanej wartości z_k po s krokach:

$$z_k(s) = \sum_{i=1}^k z_i \quad (12)$$

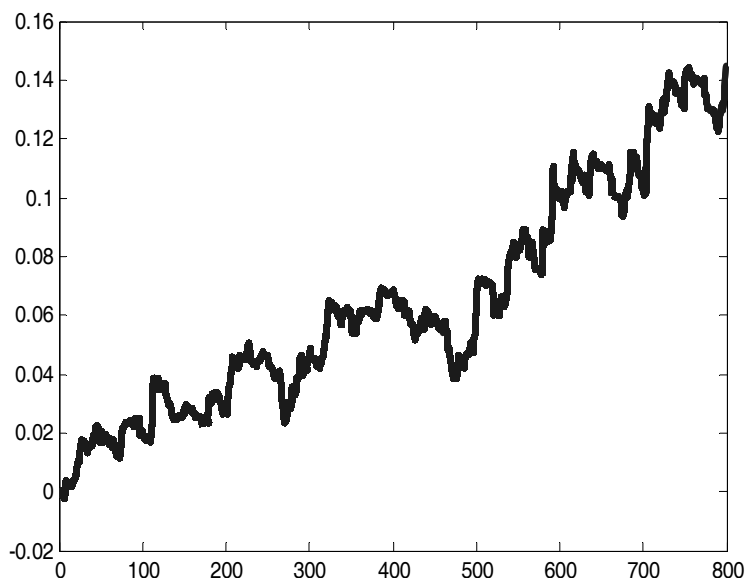
osiąga wartość wskaźnika Sharpe'a równą 0.1443.

Wskaźnik Sharpe'a obliczany był jako:

$$SR = z(k) / \text{std}(z) / k \quad (13)$$

gdzie $z(k)$ – zwrot po k kroku; $\text{std}(z_k)$ – odchylenie standardowe w szeregu czasowym ze zwrotów.

Rysunek 4. Wyniki symulacji – zysk skumulowany dla 800 transakcji



Źródło: Opracowanie własne

To uproszczone formuły, nie uwzględniające stopy zwrotu wolnej od ryzyka, ale tak zdefiniowane używane są tylko w obrębie pracy, dla porównania różnych wariantów symulacji. Wykres narastaniu kapitału z rys. 4 wykazuje wyraźną stabilną tendencję wzrostu wyrażoną ilościowo przez wskaźnik Sharpe'a.

KONKLUZJE

Należy zwrócić uwagę na dobre wyniki wszystkich symulacji (także tych, z uwagi na objętość artykułu, tu niepublikowanych) pomimo niewielkiej przewagi prognoz poprawnych, co do kierunku zmiany ceny, nad błędnymi. Doszukiwać się w tym sukcesie należy pozytywnego wpływu istoty metody – prognozuje ona nie tyle kierunek zmian, ile rzeczywiste wartości kursu po jednym lub dwu krokach. Stąd zapewne błędy (co do kierunku) mają zdecydowanie mniejszy wpływ negatywny na wynik końcowy, niż pozytywny wpływ poprawnych predykcji. Te liczne przegrane są po prostu średnio mniejsze, niż liczne wygrane. Np. dla badań przedstawionych na rys. 4 średnia przegrana wynosiła 6 pipsów a średnia wygrana aż 14 pipsów. Sytuacje te można interpretować jako dobrą ekstrakcję tendencji rynku na

podstawie kontekstu behawioralnego. Zamiary uczestników rynku dają się często poprawnie odwzorować poprzez złożone nieliniowe modele wielomianowe.

Podsumowując wyniki przeprowadzonych symulacji, należy uznać je za bardzo obiecujące. Uzyskane są na największym pod względem obrotów światowym rynku, na najbardziej płynnej parze walutowej. Zagrożeniem dla ewentualnego handlu automatycznego zbudowanego na zaprezentowanej koncepcji może być jedynie stabilność dostępu do danych o zachowaniach inwestorów, zależna wyłącznie od woli zarządzających platformą brokerską.

BIBLIOGRAFIA

- Barberis N., Shleifer A. Vishny R. A model of investor sentiment. *Journal of Financial Economics* 49 (1998), str.307-343
- DeBontd W., Thaler R. Does the stock market overreact? *Finance of Journal* 40(3) 1985, str.793-805
- Dehuri S., Cho S.-B. Multicriterion Pareto based particle swarm optimized polynomial neural network for classification: A review and state-of-the-art. *Computer Science Review* 3(2009) str.19-40.
- Fama E.F. Market efficiency, long-term returns, and behavioral finance. *Journal of Financial Economics* 49 (1998) str.283-306
- Hia Jong Teoh, Tai Liang Chen, Ching Hsue Cheng, Hsing Hui Chu. A hybrid multi-order fuzzy time series for forecasting stock markets. *Expert Systems with Applications* 36 (2009) str. 7888-7897
- Ivakhnenko A., Ivakhnenko G., Mueller J. Self- organization of Neural Network with Active Neurons. *Pattern Recognition and Image Analysis* 1999 v.4/2 str.185-196
- Ivakhnenko A., Ivakhnenko G., Problems of Further Development of the Group Method of Data Handling Algorithms. Part I. *Pattern Recognition and Image Analysis* vol.10 No.2, 2000, str.. 187-194
- Lemke F., Mueller J. A. Self- Self- Organizing Data Mining A Portfolio Trading System. *Journal of Computational Intelligence in Finance*, 1997/05 str. 12-26
- Madala H.R, Ivakhnenko A., *Inductive Learning Algorithms for Complex Systems Modeling*. CRC Press, 1994.
- Park B.J., Oh S.K., Ahn T.C., Pedrycz W. A study on multilayer fuzzy polynomial interference system based on an extended GMDH algorithm. *Proc. 8th IEEE Inf. Conference on Fuzzy Systems*, Seoul 1999.
- Piper J. *The Way to Trade*, Harriman Modern Classic, 2006.
- Sheng-Hsun Hsu, Po-An Hsieh, Ting-Chih Chih, Kuei-Chu Hsu. A two-stage architecture for stock price forecasting by integrating self-organizing map and support vector regression. *Expert Systems with Applications* 36(2009) 7947-7951
- Shleifer A. *An inefficient market: An introduction to behavioral finance*. Clarendon Lectures. Oxford University Press 2000.

Sung-Kwun Oh, Pedrycz W. The design of self-organizing Polynomial Neural Networks. *Information Sciences*, 141 (2002) str. 237-258

Wilinski A. GMDH Based Computational Intelligence Method for Prediction of Financial Markets. "Intellectual Systems of Decision Making and The Problems of Computational Intelligence (ISDMCI'2008)", Eupatoria, Tom 1 cz.1, 2008.

Wiliński A., Czapiewski P. GMDH-based Multiple Prediction of Financial Markets. *Polish Journal of Environmental Studies* vol.18 No 4a, 2008.

Zhang Yudong, Wu Lenan. Stock market prediction of S&P 500 via combination of improved BCO approach and BP neural network. *Expert Systems with Applications* 36(2009) str. 8849-8854.

Oanda, www.oanda.com dostęp kwiecień – listopad 2008.

Wanadoo, www.wanadoo.com dostęp lipiec 2008.

Prediction efficiency investigation of GMDH-class model based on behaviour of currency market users.

Abstract: The paper presents the results of an experiment concerning prediction of the foreign exchange market. It was assumed, that the market is not efficient and that it is possible to extract from the past the knowledge regarding traders' mistakes. A modified version of GMDH method was used for prediction, which allows for successive selection of such nonlinear polynomial model, that describes the market most adequately at a particular moment. Presented results confirm usefulness of the proposed method. Input data was comprised of the information on traders behaviour, registered by the brokerage platform, regarding open positions and orders. Hence, such a solution can be thought of as behavioural modelling.

Keywords: investors' behaviour, behavioural finance, regression models, GMDH, brokerage platforms, efficient market

BADANIE INFORMACYJNEJ EFEKTYWNOŚCI RYNKU W FORMIE SILNEJ NA PRZYKŁADZIE WYBRANYCH FUNDUSZY INWESTYCYJNYCH¹

Dorota Witkowska, Krzysztof Kompa

Katedra Statystyki i Ekonometrii SGGW,

dorota_witkowska@sggw.pl, krzysztof_kompa@sggw.pl

Marta Grabska

marta.grabska@gmail.com

Streszczenie: Celem badań jest ocena efektywności 19 wybranych otwartych funduszy inwestycyjnych działających na polskim rynku. Analizy przeprowadzono dla danych dziennych z okresu od 2.11.2006r. do 2.11.2008r. Badaniem objęto szeregi 3-, 6-, 12- i 24-miesięczne. Do oceny efektywności wykorzystano wskaźniki Sharpe'a, Treynora i Jensena. Badano również selektywność i wycucie rynku za pomocą metod Treynora-Mauzy'ego oraz Henrikssona-Mertona. Za indeks rynku przyjęto indeks WIG, a za instrumenty wolne od ryzyka - bony skarbowe.

Słowa kluczowe: otwarte fundusze inwestycyjne, wskaźniki efektywności inwestycji, selektywność, wycucie rynku

WPROWADZENIE

Większość analiz finansowych opiera się na hipotezie rynku efektywnego² (*Efficient Market Hypothesis*, EMH), zgodnie z którą rynki finansowe szybko i dokładnie odzwierciedlają dostępne publicznie informacje. Sformułowane przez Fama warunki informacyjnej efektywności rynku w praktyce nie są spełnione, zakłada się więc, że do spełnienia warunków efektywności wystarcza aby rynek

¹ Badania prowadzono w ramach realizacji grantu Nr N111 014 32/1227 nt. Badanie efektywności inwestycyjnej otwartych funduszy inwestycyjnych funkcjonujących na polskim rynku kapitałowym.

² Została ona sformułowana w pracy: Fama [1970].

charakteryzował się następującymi cechami³: (1) dużą liczbą uczestników; (2) homogenicznością produktów, co oznacza, że wszystkie akcje charakteryzują się parametrami, umożliwiającymi ich porównanie; (3) sposobem organizacji rynku, zapewniającym dostęp do informacji przez wszystkich uczestników rynku oraz (4) losowym charakterem napływu nowych informacji (z wyjątkiem raportów kwartalnych, których terminy publikacji są znane).

W literaturze wyróżnia się trzy formy efektywności informacyjnej:

- słabą (*weak form*), wg której ceny walorów odzwierciedlają wszystkie istotne informacje zawarte w notowaniach historycznych;
- półsilną lub średnią (*semi-strong form*), kiedy ceny papierów wartościowych odzwierciedlają nie tylko wiadomości, które można odczytać z notowań historycznych, ale także wszystkie inne publiczne informacje (np. zawarte w raportach finansowych spółek, ogłaszane publicznie dane makroekonomiczne lub informacje prasowe);
- silną (*strong form*), wg której zarówno publicznie dostępne dane, jak i wiadomości poufne (tj. takie, które są dostępne tylko wąskiej grupie osób) znajdują szybkie odzwierciedlenie w cenach walorów.

W analizach najczęściej weryfikuje się hipotezę o słabej efektywności informacyjnej rynku, do czego wykorzystuje się między innymi narzędzia analizy technicznej oraz statystyczne testy weryfikujące losowość zmian⁴. W przypadku badania silnej formy efektywności najczęściej przeprowadza się analizy portfeli inwestorów instytucjonalnych, zwłaszcza funduszy inwestycyjnych. Zakłada się bowiem, że profesjonaliści zarządzający funduszami stanowią najlepiej poinformowaną grupę inwestorów⁵.

Celem badań jest ocena efektywności wybranych funduszy inwestycyjnych (FIO) zrównoważonych i aktywnej alokacji działające w Polsce. Badania przeprowadzono w oparciu o dzienne notowania jednostek uczestnictwa pochodzące z 2-letniego okresu. W analizach wykorzystano wskaźniki efektywności Sharpe'a, Treynora i Jensena oraz Treynora-Mazury'ego i Henrikssona – Mertona.

METODY OCENY EFEKTYWNOŚCI INWESTYCJI

Dokonanie prawidłowego pomiaru efektywności portfela funduszu wymaga zastosowania specjalnych mierników, które uwzględniają nie tylko zysk z inwesty-

³ Porównaj: Buczek [2005], s. 17 – 18, Szyszka [2003], s. 17.

⁴ Wyniki analiz efektywności informacyjnej polskiego rynku opisano m.in. w pracach: Buczek [2005], czekaj i in. [2001], Kompa, Matuszewska-Janica [2008], Szyszka [2005], Witkowska, Żebrowska – Suchodolska [2008].

⁵ Jest to powszechnie przyjmowane założenie, o którym wspomina się m.in. w pracy: Czekaj i in. [2001], s. 132 - 137.

cji, ale i ryzyko⁶. Oczekiwany przez inwestora zysk, mierzy się zazwyczaj. oczekiwaną (średnią) dzienną stopą zwrotu R_p obliczoną dla szeregu τ -elementowego:

$$R_p = \frac{1}{\tau} \sum_{t=1}^{\tau} R_{pt} \quad (1)$$

gdzie: R_{pt} - dzienna logarytmiczna stopa zwrotu:

$$R_{pt} = \ln\left(\frac{P_t}{P_{t-1}}\right) \quad (2)$$

gdzie: P_t - wartość jednostki uczestnictwa w dniu t .

Często przyjmowaną miarą ryzyka jest odchylenie standardowe stóp zwrotu S_p obliczane według wzoru:

$$S_p = \sqrt{\frac{1}{\tau-1} \sum_{t=1}^{\tau} (R_{pt} - R_p)^2} \quad (3)$$

gdzie: oznaczenia jak poprzednio.

Inną popularną miarą ryzyka jest tzw. współczynnik beta modelu Sharpe'a⁷:

$$R_{pt} = \alpha_p + \beta_p \cdot R_{mt} + \varepsilon_t \quad (4)$$

lub linii rynku papierów wartościowych SML (*Security Market Line*) modelu równowagi rynku papierów wartościowych CAPM (*Capital Asset Pricing Model*)⁸:

$$R_{pt} - R_{ft} = \alpha_p + \beta_p \cdot (R_{mt} - R_{ft}) + \varepsilon_t \quad (5)$$

gdzie dla każdego okresu t : R_{pt} - stopa zwrotu p -tego funduszu inwestycyjnego, R_{ft} - stopa zwrotu instrumentu wolnego od ryzyka, R_{mt} - stopa zwrotu z indeksu rynku, $R_{mt} - R_{ft}$ - premia za ryzyko tj. dodatkowy dochód powyżej stopy zwrotu instrumentu wolnego od ryzyka, który ma zachęcić inwestorów do podjęcia ryzyka. Modele (4) i (5) tradycyjnie szacuje się MNK⁹.

⁶ Porównaj: Czekaj i in. [2001], s. 132 - 157.

⁷ Model Sharpe'a jest bogato omawiany w literaturze przedmiotu (por. Jajuga, Jajuga [2005], s. 162 - 167, Tarczyński [1997], s. 103 - 111).

⁸ Model CAPM jest omawiany m.in. w pracach: Czekaj i in. [2001], s. 139, Jajuga, Jajuga [2005], s. 167 - 173, Tarczyński [1997], s. 121 - 124, Witkowska i in. [2008], s. 221 - 225.

⁹ Stosowane są również inne metody estymacji porównaj chociażby prace: Fiszedler [2005], Gajdka, Brzeszczyński [2007].

Porównanie wyników inwestycyjnych polega na wyznaczeniu względnej efektywności inwestycji, pozwalającej na ocenę poszczególnych portfeli. W tym celu konieczne jest zastosowanie wskaźników efektywności, na wartość których wpływa zarówno ryzyko, jak i stopa zwrotu. Wyróżnia się trzy podstawowe wskaźniki efektywności¹⁰: Sharpe'a, Treynora i Jensena.

Współczynnik efektywności inwestycji Sharpe'a - WS_p pozwala ocenić wyniki portfela w oparciu o miarę ryzyka całkowitego i stopę zwrotu:

$$WS_p = \frac{R_p - R_f}{S_p} \quad (6)$$

gdzie: R_p – oczekiwana stopa zwrotu portfela funduszu, R_f – stopa zwrotu wolna od ryzyka, S_p – ryzyko całkowite portfela, wyznaczone z (3). Zauważmy, że oczekiwane (średnie) stopy zwrotu dla portfela R_p , indeksu rynku R_m i instrumentu wolnego od ryzyka R_f wyznacza się z relacji (1).

Czym wyższa wartość współczynnika Sharpe'a tym inwestycje charakteryzują się większą efektywnością. Miara Sharpe'a pozwala także na porównanie efektywności wszystkich analizowanych portfeli z efektywnością rynku mierzoną za pomocą wskaźnika WS_m wyznaczonego jako:

$$WS_m = \frac{R_m - R_f}{S_m} \quad (7)$$

gdzie: R_m – stopa zwrotu indeksu rynku (którym może być np. indeks WIG), S_m – ryzyko całkowite rynku, wyznaczone z (3), pozostałe oznaczenia jak w (6).

Porównanie wyników funduszu i rynku umożliwi wyróżnienie funduszy godnych zainteresowania inwestorów. Jeśli bowiem zachodzi: $WS_p > WS_m$, wówczas analizowany fundusz inwestycyjny jest bardziej efektywny niż rynek i opłaca się w niego inwestować.

Podobne podejście do pomiaru efektywności inwestycji reprezentuje wskaźnik Treynora. Obydwie miary różnią się jedynie tym, że Treynor za właściwą miarę ryzyka przyjął ryzyko systematyczne (niedywersyfikowalne) portfela. W celu wyprowadzenia współczynnika, który mierzy premię za ryzyko w odniesieniu do systematycznego ryzyka portfela, wykorzystano równanie linii papierów wartościowych SML. Zakładając zatem, iż rynek jest w równowadze i oczekiwania inwestorów realizują się na poziomie średniej, wskaźnik efektywności WT_p przyjmuje następującą postać:

$$WT_p = \frac{R_p - R_f}{\beta_p} \quad (8)$$

¹⁰ Porównaj: Czekał i in. [2001], s. 133, Tarczyński [1997], s. 154 – 156, Ostrowska [2003], Witkowska i in. [2008], s.225 -230.

gdzie: $\hat{\beta}_p$ - ryzyko systematyczne funduszu inwestycyjnego, oszacowane w (5); pozostałe oznaczenia jak w (6).

Podobnie jak w przypadku wskaźnika Sharpe'a, im wyższa jest wartość wskaźnika Treynora, tym fundusz inwestycyjny jest bardziej efektywny i atrakcyjny dla inwestora. Ujemne wartości wskaźnika Treynora świadczą o złych wynikach osiągniętych przez zarządzającego (por. Czeka [2008], s. 446). W celu ustalenia czy fundusz inwestycyjny osiągnął lepszy wynik od rynku należy porównać wskaźnik Treynora, obliczony dla analizowanego funduszu, ze wskaźnikiem wyznaczonym dla portfela rynkowego (WT_m). Ten ostatni wyznaczany jest analogicznie jak w przypadku miary Sharpe'a zgodnie ze wzorem (7), zastępując jedynie ryzyko całkowite rynku (S_m) ryzykiem systematycznym rynku (β_m). Wówczas za portfele opłacalne uważa się te, których wskaźniki efektywności są wyższe od wskaźnika efektywności rynku.

Inną miarą efektywności jest współczynnik Jensena:

$$\alpha_p = (R_p - R_f) - \hat{\beta}_p \cdot (R_m - R_f) \quad (9)$$

zwany również alfą Jensena, ponieważ α_p jest oceną estymatora wyrazu wolnego modelu (5).

Wartość α_p informuje o skuteczności i efektywności menadżera zarządzającego inwestycją finansową, a także o rezultacie doboru składników do portfela inwestycyjnego (por. Ostrowska [2005], s. 217). Dodatnia wartość tego wskaźnika oznacza, że zwrot z portfela przekracza oczekiwania oszacowane na podstawie modelu CAPM. Jeśli α_p osiągnie wartość mniejszą od zera, wówczas jest to sygnał, że dana inwestycja (portfel) nie jest godna zainteresowania inwestora. Stosując współczynnik Jensena należy pamiętać, że jest on miarą absolutną. Dlatego też, w swej podstawowej postaci nie jest właściwy do relatywnych porównań portfeli o różnych poziomach ryzyka. Dopiero jeśli zostanie podzielony przez współczynnik beta, może stanowić kryterium oceny portfela i być wykorzystywany w analizie porównawczej poszczególnych inwestycji.

W ocenie silnej efektywności uwzględnia się dwie cechy zarządzających portfelami: selektywność (*selectivity*) definiowaną jako umiejętny wybór akcji w celu poprawy efektywności portfela oraz wycucie rynku (*market timing*) polegające na właściwym wyborze momentu zmiany struktury portfela.

O ile w przypadku selektywności ocenia się, właściwą zarządzającym, umiejętność formułowania poprawnych prognoz dotyczących wybranych instrumentów finansowych w oparciu o informację publiczną i poufną (nieдоступną dla przeciętnego inwestora), o tyle przez wycucie rynku rozumie się umiejętność prognozowania

wania zachowania całego rynku, co umożliwia właściwy wybór momentu zmiany struktury portfela (przede wszystkim akcji)¹¹.

Wycucie rynku przez zarządzających instrumentem finansowym bada się z wykorzystaniem metody parametrycznej - wskaźnika Treynora-Mauzy'ego. Należy w tym celu oszacować model postaci (Treynor, Mauzy [1966]):

$$R_{pt} = \alpha_p + \beta_p R_{mt} + \gamma_p R_{mt}^2 + \varepsilon_t \quad (10)$$

gdzie dla każdego t : R_{pt} - stopa zwrotu z portfela, R_{mt} - stopa zwrotu z portfela rynkowego, β_p - miara ryzyka systematycznego portfela, γ_p - miara wycucia rynku przez zarządzającego funduszem.

Dodatnia wartość parametru γ_p świadczy o dobrym wycuciu rynku tzn. zarządzający portfelem dobrze dostosowują jego strukturę do warunków panujących na rynku. Innymi słowy, ze wzrostem stóp zwrotu z portfela rynkowego, zwiększają ekspozycję portfela na ryzyko, a wraz ze spadkiem stóp zwrotu indeksu giełdy, zwiększają udział instrumentów wolnych od ryzyka.

Wskaźnik Henrikssona-Mertona, podobnie jak Treynora-Mauzy'ego, pozwala na badanie wycucia rynku przez zarządzających portfelami. Konstrukcja tego testu, umożliwia oddzielenie od siebie wpływu umiejętności zarządzającego portfelem w zakresie wycucia rynku oraz umiejętności odpowiedniego doboru aktywów. Aby odseparować te dwie kategorie od siebie należy oszacować MNK model (Henriksson, Merton [1981]):

$$R_{pt} - R_{ft} = \alpha_p + \beta_{1p} \cdot (R_{mt} - R_{ft}) + \beta_{2p} y_t + \varepsilon_t \quad (11)$$

gdzie: $R_{mt} - R_{ft}$ - nadwyżka stopy zwrotu z portfela rynkowego nad stopą wolną od ryzyka, $y_t = \max\{0, R_{mt} - R_{ft}\}$, ε_t - składnik losowy modelu, który spełnia założenia modelu CAPM¹².

W modelu (11) istotną rolę pełnią oszacowania parametrów α_p oraz $\beta_{2,p}$. Badanie selektywności polega na weryfikacji hipotezy zerowej: $H_o : \alpha_p = 0$, która oznacza, że zarządzający funduszem nie posiadają umiejętności przewidywania w skali mikro. Z kolei badania umiejętności wycucia rynku polega na weryfikacji hipotezy: $H_o : \beta_p = 0$, z której wynika, że zarządzający nie posiadają umiejętności w zakresie wycucia rynku lub ich nie wykorzystują (por. Henriksson [1984]). Jeżeli $\beta_{2,p} < 0$, to zarządzający posiada umiejętności w zakresie wycucia rynku, jednak technika timing market ma negatywny wpływ na wartość portfela. Czyli wybór momentu dokonania inwestycji oraz czasu jej trwania w oparciu o krótko-

¹¹ Porównaj Czekaj i in. [2001], s. 132.

¹² Porównaj Czekaj i in. [2001], s. 153.

terminowe oczekiwania cenowe zmniejszają wartość portfela. Może to wynikać na przykład z ponoszonego ryzyka związanego z instrumentami pochodnymi.

OPIS DANYCH

Badaniem objęto 19 otwartych funduszy inwestycyjnych zrównoważonych i aktywnej alokacji działających na polskim rynku¹³ (por. tab. 1 i 2). Wybrano fundusze, które - wg Izby Zarządzających Funduszami i Aktywami – miały w listopadzie 2008r. największy udział w rynku aktywów netto funduszy inwestycyjnych¹⁴. Przyjęto, że portfel rynkowy jest reprezentowany przez indeks WIG. Analizy, obejmujące okres od 2.11.2006r. do 2.11.2008r., przeprowadzono dla czterech szeregów czasowych o różnej długości (por. rys. 1), które oznaczono następująco:

- M24 - szereg 24-miesięczny, obejmujący okres: 2.11.2006r. - 2.11.2008,
- M12 - szereg 12-miesięczny, obejmujący okres: 22.10.2007 - 22.10.2008,
- M6 - szereg 6-miesięczny, obejmujący okres: 14.04.2008 - 14.04.2008,
- M3 - szereg 3-miesięczny, obejmujący okres: 16.06.2008r. - 16. 09.2008.

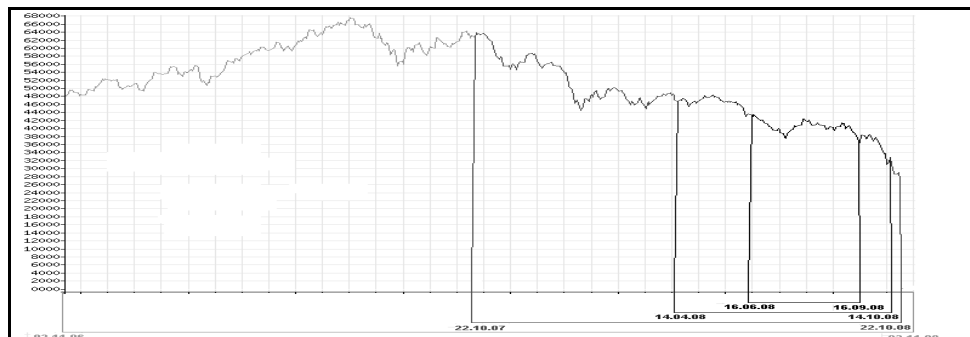
Ocena efektywności inwestycji wymaga wykorzystania stopy wolnej od ryzyka, za którą przyjęto stopę zwrotu z instrumentów finansowych, emitowanych przez Ministerstwo Finansów. Stąd początek każdego okresu wyznaczono jako dzień emisji bonów bądź obligacji skarbowych o określonym czasie wymagalności. Dzienną stopę zwrotu wolną od ryzyka dla poszczególnych szeregów wyznaczono jakoienne oprocentowanie następujących instrumentów finansowych:

- 2-letnich obligacji skarbowych dla szeregu 24-miesięcznego,
- 52-tygodniowych bonów skarbowych dla szeregu 12-miesięcznego,
- 26-tygodniowych bonów skarbowych dla szeregu 6-miesięcznego,
- 13-tygodniowych bonów skarbowych dla szeregu 3-miesięcznego.

Rysunek 1. Okresy badania oraz wykres WIGu w okresie od 2.11.2006 do 2.11.2008r.

¹³ Badanie efektywności funduszy inwestycyjnych działających na polskim rynku przedstawiono m.in. w pracach: Czekaj i in. [2001], Ostrowska [2003], Olbryś [2008].

¹⁴ Por. Analizy Online i raporty Izby Zarządzających Funduszami i Aktywami listopad 2008.



Źródło: Opracowanie własne.

Wszystkie analizowane fundusze zrównoważone, niezależnie od długości rozpatrywanego szeregu, charakteryzowały się ujemną wartością średniej dziennej stopy zwrotu, chociaż istotnie mniejsze od zera były jedynie roczne stopy zwrotu dla wybranych funduszy. Podobnie było w przypadku funduszy aktywnej alokacji, z wyjątkiem funduszu AIG Fundusz Aktywnej Alokacji, który miał nieujemne stopy zwrotu dla 6- oraz 12-miesięcznego szeregu. Dla większości funduszy najmniejsze oczekiwane straty zaobserwowano dla 24-miesięcznych szeregów¹⁵.

Wśród badanych funduszy zrównoważonych najwyższym ryzykiem całkowitym charakteryzował się fundusz Arka BZWBK Zrównoważony FIO dla wszystkich czterech badanych szeregów. Natomiast najniższe ryzyko całkowite zaobserwowano dla różnych funduszy w zależności od długości rozpatrywanego okresu:

- dla szeregu M3 jest to UniFundusze FIO Subfundusz UniKorona Zrównoważony,
- dla M6 jest to Noble Funds FIO Subfundusz Noble Fund Mieszany,
- dla M12 oraz M24 jest to Millenium FIO SubFundusz Zrównoważony.

Ponadto zauważono, że w przypadku 3-, 6-, 12-miesięcznych szeregów, wraz ze wzrostem długości analizowanego okresu wzrastało ryzyko całkowite funduszy.

¹⁵ Nie dotyczy to funduszy, których działalność trwała krócej niż dwa lata lub rok (por. tab. 1 i 2).

Tabela 1. Informacje dotyczące analizowanych funduszy inwestycyjnych zrównoważonych.

Nazwa funduszu inwestycyjnego	TFI	Polityka inwestycyjna funduszu inwestycyjnego	Data rozpoczęcia działalności
Pioneer Zrównoważony FIO	Pioneer Pekao	<=70% akcje, >=30% instrumenty finansowe o charakterze wierzycielskim, część aktywów stanowią depozyty bankowe	30.07.1992
Arka BZ WBK Zrównoważony FIO	BZ WBK AIB	30-70% akcje, pozostałe środki to instrumenty finansowe o charakterze wierzycielskim	07.01.2002
PKO/Credit Suisse Zrównoważony FIO	PKO	<=70% akcje, >=30% to instrumenty finansowe o charakterze wierzycielskim	09.07.1997
ING FIO Zrównoważony	ING	średnio 50% akcje, średnio 50% instrumenty finansowe o charakterze wierzycielskim	09.03.1998
Skarbiec Waga FIO Zrównoważony	Skarbiec	30-70% akcje, pozostałe środki- instrumenty finansowe o charakterze wierzycielskim	09.10.1997
UniFundusze FIO subfundusz UniKorona Zrównoważony	Union Investment	20-80% akcje, pozostałe środki to instrumenty finansowe o charakterze wierzycielskim	09.08.1995
PZU FIO Zrównoważony	PZU	30-70% akcje, pozostałe środki to instrumenty finansowe o charakterze wierzycielskim	18.11.2005
Millennium FIO Subfundusz Zrównoważony	Millennium	40-60% akcje, pozostałe środki to instrumenty finansowe o charakterze wierzycielskim	03.01.2002
Commercial Union FIO subfundusz CU Zrównoważony	Commercial Union	40-60% akcje, 20-60% to instrumenty finansowe o charakterze wierzycielskim	29.12.2006
DWS Polska FIO Zrównoważony	DWS Polska	40-65% akcje, instrumenty oparte o akcje, pozostałe środki to instrumenty finansowe o charakterze wierzycielskim	05.01.1998
SEB FIO Subfundusz SEB Zrównoważonego Wzrostu	SEB	<=70% akcje, pozostałe środki to instrumenty finansowe o charakterze wierzycielskim	01.06.1998
Noble Funds FIO Subfundusz Noble Fund Mieszany	Noble Funds	20-80% papiery udziałowe, <=80% instrumenty finansowe o charakterze wierzycielskim	01.12.2006

Źródło: Opracowanie własne na podstawie: <http://fundusze.wp.pl/wizytowki.html?POD=6>

Tabela 2. Informacje dotyczące analizowanych funduszy inwestycyjnych aktywnej alokacji.

Nazwa funduszu inwestycyjnego	Towarzystwo funduszu inwestycyjnego	Polityka inwestycyjna funduszu inwestycyjnego	Data rozpoczęcia działalności
Pioneer Aktywnej Alokacji FIO	Pioneer Pekao	0-100% akcje, 0-100% instrumenty finansowe o charakterze wierzyielskim oraz depozyty bankowe	18.09.2007
KBC Parasol FIO Subfundusz Aktywny	KBC	30-70% akcje, pozostałe środki to instrumenty finansowe o charakterze wierzyielskim	26.07.2002
BPH FIO Parasolowy BPH Subfundusz Aktywnego Zarządzania	BPH	<=70% akcje, >=30% instrumenty finansowe o charakterze wierzyielskim	06.06.2001
PZU FIO Optymalnej Alokacji	PZU	0-100% akcje, 0-100% instrumenty finansowe o charakterze wierzyielskim, 0-100% instrumenty pochodne	20.12.2007
AIG Fundusz Aktywnej Alokacji	AIG	20-80% akcje, pozostałe środki to instrumenty finansowe o charakterze wierzyielskim, <=10% papiery wartościowe o stałym dochodzie i depozyty bankowe utrzymywane dla celów zarządzania płynnością	03.11.2006
Allianz FIO subfundusz Allianz Aktywnej Alokacji	Allianz Polska	0-80% akcje, średnio 45% akcje i obligacje zamienne, 55% to instrumenty finansowe o charakterze wierzyielskim o terminie wykupu dłuższym niż rok	16.12.2005
SKOK Parasol FIO Subfundusz SKOK Aktywny Zmiennej Alokacji	SKOK	0-100% akcje, pozostałe środki to instrumenty finansowe o charakterze wierzyielskim (obligacje skarbowe)	18.07.2005

Źródło: Opracowanie własne na podstawie: <http://fundusze.wp.pl/wizytowki.html?POD=6>

ANALIZA MODELU SHARPE'A, WSKAŹNIK SHARPE'A, WSKAŹNIK TREYNORA

Model Sharpe'a pozwala zbadać wpływ zmian zachodzących na rynku, które w naszym badaniu odzwierciedla indeks WIG, na kształtowanie się stóp zwrotu funduszy. Oprócz tego umożliwia analizę ryzyka ocenianych funduszy. W związku z tym dla analizowanych funduszy oszacowano MNK modele jednowskaźnikowe na podstawie szeregów czasowych M3 – M24. Oceny parametru beta wszystkich funduszy, niezależnie od długości analizowanego okresu, były statystycznie istotne i zawierały się w przedziale (0; 1), a skorygowany współczynnik determinacji w większości przypadków oscylował wokół 0,90. Oznacza to, że analizowane fundusze były funduszami defensywnymi.

Tabela 3. Zbiór portfeli bardziej efektywnych niż rynek: fundusze zrównoważone

Zbiór portfeli bardziej efektywnych niż Sharpe'a	rynek wyznaczonych za pomocą wskaźnika Treynora
dla szeregu	3- miesięcznego
<ul style="list-style-type: none"> • Arka BZWBK, • PKO/Credit Suisse Zrównoważony FIO, • ING FIO Zrównoważony, • Skarbiec Waga FIO Zrównoważony, • Millenium FIO Subfundusz Zrównoważony, • Commercial Union FIO subfundusz CU Zrównoważony, • DWS Polska FIO Zrównoważony. 	<ul style="list-style-type: none"> • PKO/Credit Suisse Zrównoważony FIO, • ING FIO Zrównoważony, • Commercial Union FIO subfundusz CU Zrównoważony.
dla szeregu	6- miesięcznego
<ul style="list-style-type: none"> • Arka BZWBK, • PKO/Credit Suisse Zrównoważony FIO, • ING FIO Zrównoważony, • Skarbiec Waga FIO Zrównoważony. 	<ul style="list-style-type: none"> • PKO/Credit Suisse Zrównoważony FIO, • Arka BZWBK, • ING FIO Zrównoważony.
dla szeregu	12- miesięcznego
<ul style="list-style-type: none"> • Arka BZWBK, • Skarbiec Waga FIO Zrównoważony, • DWS Polska FIO Zrównoważony • PKO/Credit Suisse Zrównoważony FIO. • ING FIO Zrównoważony. 	Nie istnieje
dla szeregu	24- miesięcznego
<ul style="list-style-type: none"> • Arka BZWBK, • UniFundusze FIO subfundusz Uni-Korona Zrównoważony, • PZU FIO Zrównoważony • DWS Polska FIO Zrównoważony. 	PZU FIO Zrównoważony.

Źródło: Opracowanie własne.

Tabela 4. Zbiór portfeli bardziej efektywnych niż rynek: fundusze aktywnej alokacji

Zbiór portfeli bardziej efektywnych niż rynek wyznaczonych za pomocą wskaźnika Sharpe'a	rynek wyznaczonych za pomocą wskaźnika Treynora
dla szeregu	3- miesięcznego
Zbiór portfeli bardziej efektywnych niż rynek wyznaczonych za pomocą wskaźnika Sharpe'a	Zbiór portfeli bardziej efektywnych niż rynek wyznaczonych za pomocą wskaźnika Treynora
dla szeregu	3- miesięcznego
Pioneer Aktywnej Alokacji FIO, BPH Aktywnego Zarządzania Sub BPH Parasolowy FIO, AIG Fundusz Aktywnej Alokacji, Allianz FIO Subfundusz Allianz Aktywnej Alokacji, SKOK Parasol FIO Subfundusz SKOK Aktywny Zmiennej Alokacji dla szeregu 3- miesięcznego.	Pioneer Aktywnej Alokacji FIO, AIG Fundusz Aktywnej Alokacji, Allianz FIO Subfundusz Allianz Aktywnej Alokacji, SKOK Parasol FIO Subfundusz SKOK Aktywny Zmiennej Alokacji.
dla szeregu	6- miesięcznego
Pioneer Aktywnej Alokacji FIO, Allianz FIO Subfundusz Allianz Aktywnej Alokacji, SKOK Parasol FIO Subfundusz SKOK Aktywny Zmiennej Alokacji.	Pioneer Aktywnej Alokacji FIO, AIG Fundusz Aktywnej Alokacji, Allianz FIO Subfundusz Allianz Aktywnej Alokacji, SKOK Parasol FIO Subfundusz SKOK Aktywny Zmiennej Alokacji.
dla szeregu	12- miesięcznego
Pioneer Aktywnej Alokacji FIO, AIG Fundusz Aktywnej Alokacji, Allianz FIO subfundusz Allianz Aktywnej Alokacji, SKOK Parasol FIO Subfundusz SKOK Aktywny Zmiennej Alokacji.	Pioneer Aktywnej Alokacji FIO, AIG Fundusz Aktywnej Alokacji, Allianz FIO Subfundusz Allianz Aktywnej Alokacji, SKOK Parasol FIO Subfundusz SKOK Aktywny Zmiennej Alokacji.
dla szeregu	24- miesięcznego
Allianz FIO Subfundusz Allianz Aktywnej Alokacji, SKOK Parasol FIO Subfundusz SKOK Aktywny Zmiennej Alokacji.	SKOK Parasol FIO Subfundusz SKOK Aktywny Zmiennej Alokacji

Źródło: Opracowanie własne.

Ocena efektywności inwestycji przeprowadzona na podstawie wskaźników Sharpe'a oraz Treynora polegała na porównaniu wartości mierników wyznaczonych dla funduszy oraz indeksu rynku w analizowanych podokresach. Listę funduszy bardziej efektywnych niż WIG przedstawiono w tab. 3 i 4. Przykładowo dla szeregu 3-miesięcznego stwierdzono, że (tab. 3): PKO/Credit Suisse Zrównoważony FIO, ING FIO Zrównoważony oraz Commercial Union FIO subfundusz CU Zrównoważony charakteryzowały się lepszymi wynikami niż rynek. Fundusze te wypadły korzystnie zarówno biorąc pod uwagę ryzyko systematyczne, jak i całkowite. Funduszami, które przynosiły zyski w przeliczeniu na jednostkę ryzyka całkowitego, zaś straty w przeliczeniu na jednostkę ryzyka rynkowego były: Arka

BZWBK, Skarbiec Waga FIO Zrównoważony, Millenium FIO Subfundusz Zrównoważony, Commercial Union FIO subfundusz CU Zrównoważony, DWS Polska FIO Zrównoważony.

Warto przy tym zauważyć, że Pionier Zrównoważony FIO, SEB Zrównoważonego Wzrostu oraz Noble Fund Mieszany były mniej efektywne niż rynek we wszystkich analizowanych okresach. Natomiast Millenium Fundusz Zrównoważony oraz Commercial Union Zrównoważony osiągnęły lepsze rezultaty niż rynek w ostatnich 3 miesiącach prowadzonej analizy. Przy czym Noble Fund Mieszany i Commercial Union Zrównoważony są najkrócej działającymi funduszami inwestycyjnymi, poddanymi analizie.

Z kolei biorąc pod uwagę fundusze aktywnej alokacji (tab. 4) zauważa się, że dla szeregu M3 lepsze wyniki niż rynek osiągnęły takie fundusze jak: Pioneer Aktywnej Alokacji FIO, AIG Fundusz Aktywnej Alokacji, Allianz FIO Subfundusz Allianz Aktywnej Alokacji i SKOK Parasol FIO Subfundusz SKOK Aktywny Zmiennej Alokacji. Fundusze te wypadły korzystnie z punktu widzenia zarówno wskaźnika Sharpe'a, jak i Treynora. Jedynym funduszem, który przynosił zyski większe niż WIG w całym okresie analizy był: Pioneer Aktywnej Alokacji, aczkolwiek działał on krócej niż 2 lata (bo od września 2007). Warto też zauważyć, że FIO Allianz Aktywnej Alokacji tylko w jednym badanym okresie osiągnął rezultaty gorsze od rynku i to wyłącznie wg oceny współczynnikiem Treynora. Z kolei KBC Parasolowy FIO Aktywny oraz PZU Optymalnej Alokacji nie osiągnęły lepszych od rynkowych rezultatów w żadnym badanym okresie. Przy czym ostatni ze wspomnianych funduszy rozpoczął swoją działalność dopiero w grudniu 2007r.

Zbiór funduszy wyznaczonych przy wykorzystaniu wskaźnika Treynora, zawierał się w zbiorze określonym przez wskaźnik Sharpe'a. Taka sytuacja wynika z konstrukcji obu wskaźników. Wskaźnik Sharpe'a informuje o premii za ryzyko inwestora przypadającej na jednostkę całkowitego ryzyka, mierzonego odchyleniem standardowym. Natomiast wskaźnik Treynora jest zyskiem ponad stopę wolną od ryzyka w przeliczeniu na jednostkę ryzyka systematycznego.

MODEL CAPM, WSKAŹNIK JENSENA

Kolejnym etapem w badaniu efektywności portfeli skonstruowanych przez fundusze inwestycyjne była analiza modelu CAPM. Zastosowanie linii rynku kapitałowego wykazało, że wszystkie analizowane fundusze zrównoważone i aktywnej alokacji były portfelami nieefektywnymi ze względu na oczekiwaną stopę zwrotu i ryzyko całkowite, niezależnie od rozważanych okresów. W dalszej kolejności sprawdzono, za pomocą modelu równowagi rynku papierów wartościowych, czy badane fundusze były prawidłowo wycenione. W obu grupach funduszy można było znaleźć fundusze przeszacowane dla wszystkich badanych okresów (tab. 5).

Tabela 5. Lista funduszy przeszacowanych

Fundusze zrównoważone	Fundusze aktywnej alokacji
Pioneer Zrównoważony FIO, Skarbiec WAGA FIO Zrównoważony, UniFundusze FIO Subfundusz UniKorona Zrównoważony, Millennium FIO Subfundusz Zrównoważony, DWS Polska FIO Zrównoważony, SEB FIO Subfundusz SEB1 Zrównoważonego Wzrostu, Nobel Funds FIO Subfundusz Noble Fund Mieszany.	KBC Parasol FIO Subfundusz Aktywny, BPH Aktywnego Zarządzania SUB BPH Parasolowy FIO, PZU FIO Optymalnej Alokacji.

Źródło: Opracowanie własne.

Tabela 6. Zbiór portfeli, których stopa zwrotu przekraczała oczekiwania pochodzące z modelu CAPM dla szeregów o różnej długości

Fundusze zrównoważone bardziej efektywne niż rynek dla szeregu:			
M3	M6	M12	M24
ING FIO Zrównoważony, Commercial Union FIO Subfundusz CU Zrównoważony	ING FIO Zrównoważony, Arka BZWBK Zrównoważony FIO	Arka BZWBK Zrównoważony FIO	PZU FIO Zrównoważony
Fundusze aktywnej alokacji bardziej efektywne niż rynek dla szeregu:			
M3	M6	M12	M24
SKOK Parasol FIO Subfundusz SKOK Aktywny Zmiennej Alokacji Allianz FIO subfundusz Allianz Aktywnej Alokacji, Pioneer Aktywnej Alokacji, AIG Fundusz Aktywnej Alokacji,	AIG Fundusz Aktywnej Alokacji, SKOK Parasol FIO Subfundusz SKOK Aktywny Zmiennej Alokacji Pioneer Aktywnej Alokacji, Allianz FIO subfundusz Allianz Aktywnej Alokacji,	AIG Fundusz Aktywnej Alokacji, Allianz FIO subfundusz Allianz Aktywnej Alokacji, Pioneer Aktywnej Alokacji, SKOK Parasol FIO Subfundusz SKOK Aktywny Zmiennej Alokacji	SKOK Parasol FIO Subfundusz SKOK Aktywny Zmiennej Alokacji

Źródło: Opracowanie własne.

Niedoszacowanie funduszy zrównoważonych występowało tylko dla pojedynczych szeregów czasowych¹⁶:

- Arka BZWBK Zrównoważony FIO dla szeregów 6-, 12-miesięcznych,
- PKO/Credit Suisse Zrównoważony FIO dla szeregów 3-, 6-miesięcznych,
- ING FIO Zrównoważony dla szeregów 3-, 6-miesięcznych,

¹⁶ Wartości jednostek uczestnictwa wymienionych funduszy, dla pozostałych szeregów, charakteryzowały się przeszacowaniem.

- PZU FIO Zrównoważony dla szeregu 24-miesięcznego,
- Commercial Union FIO subfundusz CU Zrównoważony dla szeregu 3-miesięcznego.

Natomiast spośród funduszy aktywnej alokacji funduszami niedoszacowanymi we wszystkich badanych okresach były:

- Pioneer Aktywnej Alokacji,
- AIG Fundusz Aktywnej Alokacji,
- SKOK Parasol FIO Subfundusz SKOK Aktywny Zmiennej Alokacji.

Allianz FIO subfundusz Allianz Aktywnej Alokacji był funduszem, który był przeszacowany dla 24-miesięcznego szeregu, zaś dla pozostałych charakteryzował się niedoszacowaniem.

Następnym etapem badania było wyznaczenie wartości wskaźnika Jensena, na podstawie którego wskazano inwestycje opłacalne oraz zbudowano ranking funduszy wykorzystując miarę względną $\hat{\alpha}_p / \beta_p$. Stopy zwrotu zaledwie z kilku funduszy przekraczały oczekiwania pochodzące z modelu CAPM (tab. 6). Były to portfele niedoszacowane w danych okresach i – zgodnie z przewidywaniami – były to przede wszystkim fundusze aktywnej alokacji.

Warto przy tym zauważyć, że żaden z funduszy znajdujących się w tabeli 6 nie charakteryzował się statystycznie istotnym parametrem $\hat{\alpha}_p$. Parametr alfa był statystycznie istotnie mniejszy od zera dla PKO Zrównoważony FIO, Millenium FIO Subfundusz Zrównoważony dla szeregów M12 i M24, Noble Funds FIO Subfundusz Noble Fund Mieszany dla szeregu M3 oraz KBC Parasol FIO SubFundusz Aktywny dla 12 i 24-miesięcznych szeregów.

MODELE TREYNORA-MAZUY'EGO ORAZ HENRIKSSONA-MERTONA

W celu zbadania selektywności oraz wycucia rynku przez zarządzających FIO, oszacowano modele Treynora-Mazuy'ego i Henrikssona-Mertona. Oceny estymatorów parametrów γ_p oraz β_{2p} zamieszczono w tab. 7 i 8.

Tabela 7. Oszacowania parametru gamma modelu Treynora-Mazuy'ego

Nazwa funduszu	γ_p dla szeregu:			
	M3	M6	M12	24M
Fundusze zrównoważone				
Pioneer Zrównoważony FIO	0,0578	-0,1355	-0,6961	-0,2740
Arka BZWBK Zrównoważony FIO	0,0846	0,2527	0,0401	-0,5253
PKO/Credit Suisse Zrównoważony FIO	0,8713	-0,1996	-1,3022	-0,7248
ING FIO Zrównoważony	0,1560	0,2845	0,4407	0,2910
Skarbiec WAGA FIO Zrównoważony	-0,5148	0,6150	-0,2160	-0,0013

UniFundusze FIO Subfundusz UniKorona Zrównoważony	-0,4190	-0,4082	-0,5945	-0,7437
PZU FIO Zrównoważony	-0,04450	-0,6298	-0,4901	-0,5819
Millennium FIO Zrównoważony	0,3604	-0,8395	-0,5866	-0,2695
Commercial Union FIO Subfundusz CU Zrównoważony	-0,5730	-0,4755	-0,7175	fundusz nie istniał
DWS Polska FIO Zrównoważony	0,2045	-1,1153	-1,2947	-1,2006
SEB FIO Subfundusz SEB1 Zrównoważonego Wzrostu	0,8376	-1,2186	-0,8687	-0,8939
Noble Funds FIO Subfundusz Fund Mieszany	-0,1374	-0,0540	0,0834	fundusz nie istniał
Fundusze aktywnej alokacji				
Pioneer Aktywnej Alokacji FIO	-0,2803	0,4254	-0,3020	fundusz nie istniał
KBC Parasol FIO Subfundusz Aktywny	-0,4871	-0,3294	-0,6919	-0,0424
BPH Aktywnego Zarządzania Sub BPH Parasolowy FIO	-0,2480	-1,6000	-0,4139	0,0046
PZU FIO Optymalnej Alokacji	1,7281	0,7569	fundusz	nie istniał
AIG Fundusz Aktywnej Alokacji	0,5067	4,3113	4,4848	fundusz nie istniał
Allianz FIO Subfundusz Allianz Aktywnej Alokacji	0,0499	-1,7020	-2,4579	-2,0262
SKOK Parasol FIO Subfundusz SKOK Aktywny Zmiennej Alokacji	0,7305	-0,1111	-0,3193	-0,25793

Źródło: Opracowanie własne.

Uwaga: Parametry statystycznie istotne dla poziomu istotności 0,05 oznaczono pogrubioną czcionką.

Spośród 19 analizowanych funduszy, tylko dla czterech z nich ocena estymatora parametru γ_p była dodatnia we wszystkich analizowanych okresach. Były to fundusze: ING FIO Zrównoważony i Arka BZWBK Zrównoważony FIO oraz AIG Fundusz Aktywnej Alokacji i PZU FIO Optymalnej Alokacji. Można zatem uznać, że zarządzający tymi właśnie funduszami mieli dobre wyczucie rynku i odpowiednio zwiększali bądź zmniejszali ekspozycje na ryzyko rynkowe w zależności od warunków panujących na rynku. Wniosek ten został potwierdzony statystyczną weryfikacją tylko w stosunku do funduszy: ING FIO Zrównoważony i AIG Fundusz Aktywnej dla 12-miesięcznego szeregu.

Dla większości analizowanych modeli oceny estymatora parametru γ_p były ujemne. Fundusze zrównoważone: UniFundusze FIO Subfundusz UniKorona Zrównoważony, PZU FIO Zrównoważony oraz Commercial Union FIO Subfundusz CU Zrównoważony, a także jeden fundusz aktywnej alokacji - KBC Parasol FIO Subfundusz Aktywny charakteryzowały się ujemnymi wartościami parametru γ_p dla wszystkich badanych szeregów (tab. 7). Fundusz BPH Aktywnego Zarzą-

dzania Sub BPH Parasolowy FIO dla większości podokresów charakteryzował się złym wyczuciem rynku, jednak dla najdłuższego szeregu czasowego fundusz był dobrze zarządzanym portfelem.

Występowały także fundusze, które w większości okresów wykazywały negatywny wpływ techniki market timing (tab. 7), jednak w pojedynczych okresach były one dobrze zarządzanymi portfelami. Przykładem takich funduszy są: Pioneer Zrównoważony FIO, PKO/Credit Suisse Zrównoważony FIO, Skarbiec WAGA FIO Zrównoważony, Millennium FIO Zrównoważony, DWS Polska FIO Zrównoważony, Noble Funds FIO Subfundusz Fund Mieszany i SEB FIO Subfundusz SEB1 Zrównoważonego Wzrostu, a także: SKOK Parasol FIO Subfundusz SKOK Aktywny Zmiennej Alokacji, Allianz FIO Subfundusz Allianz Aktywnej Alokacji, Pioneer Aktywnej Alokacji FIO. Dla większości tych funduszy, parametr γ_p przyjmował dodatnie (aczkolwiek nieistotnie różne od zera) wartości dla 3-, bądź 6-miesięcznych szeregów. Oznacza to, że zarządzający rozpatrywanymi funduszami w niewielkim stopniu umieli wyczuć rynek w analizowanych okresach i nie dostosowali struktury portfela do warunków panujących na giełdzie.

W dalszym postępowaniu zweryfikowano hipotezy o istotności parametru α_p modelu Henrikssona-Mertona. Okazało się, że w przypadku funduszy: PZU FIO Zrównoważony, Allianz FIO Subfundusz Allianz Aktywnej Alokacji i SKOK Parasol FIO Subfundusz SKOK Aktywny Zmiennej Alokacji dla 24-miesięcznych szeregów przyjęto hipotezę o dodatniej wartości tegoż parametru, co potwierdza umiejętności przewidywania w skali mikro. Innymi słowy zarządzających tymi funduszami wykazali się umiejętnościami w zakresie doboru papierów wartościowych, tak aby otrzymywać dobre wyniki inwestycyjne. Następnie zbadano wyczucie rynku, które polegało na weryfikacji hipotez w odniesieniu do parametru β_{2p} . W tab. 8 oznaczono symbolem ~ brak podstaw do odrzucenia hipotezy zerowej, co świadczy o tym, że dany portfel nie charakteryzował się dobrym wyczuciem rynku. W sytuacji przyjęcia hipotezy alternatywnej podano jej postać.

Tabela 8A. Testowanie statystycznej istotności parametru β_{2p} w teście Henrikssona-Mertona dla funduszy zrównoważonych

Nazwa funduszu	M3	M6	M12	24M
Pioneer Zrównoważony FIO	~	~	~	~
Arka BZWBK Zrównoważony FIO	~	~	~	~
PKO/Credit Suisse Zrównoważony FIO	~	~	~	~
ING FIO Zrównoważony	~	~	~	~
Skarbiec WAGA FIO Zrównoważony	~	~	~	~
UniFundusze FIO Subfundusz UniKorona Zrównoważony	~	~	~	~

PZU FIO Zrównoważony	~	$\beta_{2p} < 0$	~	$\beta_{2p} < 0$
Millennium FIO Zrównoważony	~	$\beta_{2p} < 0$	$\beta_{2p} < 0$	~
Commercial Union FIO Subfundusz CU Zrównoważony	~	$\beta_{2p} < 0$	$\beta_{2p} < 0$	fundusz nie istniał
DWS Polska FIO Zrównoważony	~	~	~	$\beta_{2p} < 0$
SEB FIO Subfundusz SEB1 Zrównoważonego Wzrostu	~	~	~	~
Noble Funds FIO Subfundusz Fund Mieszany	~	~	~	fundusz nie istniał

Źródło: Opracowanie własne.

Jak widać w tab. 8 dla ostatniego okresu M3 żaden z funduszy nie był dobrze zarządzany. Statystycznie istotna dodatnia wartość parametru β_{2p} wystąpiła tylko dla Funduszu AIG Aktywnej Alokacji w przypadku 12-miesięcznego okresu analizy. Dla takich funduszy jak: PZU FIO Zrównoważony dla 6- i 24-miesięcznych szeregów, Millennium FIO Zrównoważony dla 6- i 12-miesięcznych szeregów, Commercial Union FIO Subfundusz CU Zrównoważony dla 6- i 12-miesięcznych szeregów oraz DWS Polska FIO Zrównoważony dla szeregu M24 przyjęto hipotezę o ujemnej wartości parametru β_{2p} . Podobnie jak w przypadku funduszy aktywnej alokacji: KBC Parasol FIO Subfundusz Aktywny dla szeregu M12, BPH Aktywnego Zarządzania Sub BPH Parasolowy FIO dla szeregu M6, Allianz FIO Subfundusz Allianz Aktywnej Alokacji dla 12- i 24-miesięcznych szeregów. Oznacza to, że wprawdzie zarządzający posiadali umiejętność wyboru momentu zawarcia transakcji, ale wpłynęło to negatywnie na wartość portfela.

Zastosowanie dwóch modeli oceny efektywności doprowadziło do różnych wniosków. Różnice mogły wynikać z faktu, że model Henrikssona-Mertona badanie umiejętności wycucia rynku dzieli na ocenę umiejętności selektywności oraz umiejętności wycucia market timing. Natomiast model Treynora-Mauzy'ego nie prowadzi do identyfikacji oraz odseparowania tych czynników.

Tabela 8B. Testowanie statystycznej istotności parametru beta w teście Henrikssona-Mertona dla funduszy aktywnej alokacji

Nazwa funduszu	M3	M6	M12	24M
Pioneer Aktywnej Alokacji FIO	~	~	~	fundusz nie istniał
KBC Parasol FIO Subfundusz Aktywny	~	~	$\beta_{2p} < 0$	~
BPH Aktywnego Zarządzania Sub BPH Parasolowy FIO	~	$\beta_{2p} < 0$	~	~
PZU FIO Optymalnej Alokacji	~	~	fundusz	nie istniał
AIG Fundusz Aktywnej Alokacji	~	~	$\beta_{2p} > 0$	fundusz nie istniał
Allianz FIO Subfundusz Allianz Aktywnej Alokacji	~	~	$\beta_{2p} < 0$	$\beta_{2p} < 0$
SKOK Parasol FIO Subfundusz SKOK Aktywny Zmiennej Alokacji	~	~	~	~

Zródło: Opracowanie własne.

Wśród analizowanych funduszy nie znaleziono takich, które charakteryzowały się zarówno dobrym wyczuciem rynku wg Treynora-Mazuy'ego oraz umiejętnością market timing i selektywnością wg Henrikssona-Mertona. Natomiast można wskazać fundusze, które charakteryzowały się brakiem obu tych własności. Specyficznym funduszem okazał się PZU FIO Zrównoważony (dla 6 i 24-miesięcznych szeregów), którego zarządzający charakteryzowali się słabym stopniem wyczucia rynku wg Treynora-Mazuy'ego oraz umiejętnościami odpowiedniego doboru aktywów oraz wyczucia rynku wg Henrikssona-Mertona, która wpływała negatywnie na wartość portfela, czyli próby efektywnego zarządzania przyniosły odwrotny efekt do pożądanego.

ZAKOŃCZENIE

Celem badania była próba oceny silnej efektywności informacyjnej, którą przeprowadzono w oparciu o upraszczające założenie, że zarządzających funduszami inwestycyjnymi, reprezentującymi inwestorów instytucjonalnych, można posądzać o posiadanie informacji niepublicznych i poufnych. Warto przy tym zauważyć, że teoria rynku efektywnego nie wyklucza występowania dodatniej selektywności „u niektórych menedżerów w niektórych okresach”¹⁷. Słabością przeprowadzonego badania może być okres analizy, który nie dość, że krótki to przypadł na okres głównie trendu spadkowego na Giełdzie Papierów Wartościowych w Warszawie. Jednakże zarabianie na giełdzie w okresie hossy nie wymaga szczególnych umiejętności. Dlatego wydaje się, że rzeczywistym sprawdzianem umiejętności ekspertów finansowych jest właśnie bessy na rynku.

¹⁷ Czekaj i in. [2001], s. 133.

Z przedstawionych wyników badań można wnioskować, że fundusze inwestycyjne zrównoważone i aktywnej alokacji nie były w analizowanym okresie zarządzane w sposób efektywny. Chociaż wyciągnięte wnioski nie tylko nie dają jednoznacznej odpowiedzi, ale czasem są ze sobą sprzeczne. Wynika to z faktu, że wykorzystane metody oceny efektywności inwestycji opierają się na różnych kryteriach. Wskaźniki Sharpe'a i Treynora pozwalają stwierdzić czy inwestycje funduszy były bardziej efektywne niż rynek. W przypadku mierników Jensena ocena dotyczy stopy zwrotu z inwestycji w porównaniu do oczekiwań wynikających z modelu CAPM. Modele Treynora-Mazuy'ego oraz Henrikssona-Mertona pozwalają na odpowiedź na pytanie czy zarządzający posiadali umiejętność przewidywania w skali mikro bądź umiejętność w zakresie wyczcucia rynku.

LITERATURA

- Buczek S.B. (2005) Efektywność informacyjna rynków akcji. Teoria a rzeczywistość, Szkoła Główna Handlowa w Warszawie, Warszawa.
- Czekaj J. (2008) Rynki, instrumenty i instytucje finansowe, Wydawnictwo Naukowe PWN, Warszawa.
- Czekaj J., Woś M., Żarnowski J. (2001) Efektywność giełdowego rynku akcji w Polsce, Wydawnictwo Naukowe PWN, Warszawa.
- Fama E.F.: Efficient capital market: (1970) A review of theory and empirical work, *Journal of Finance*, vol. 25.
- Fiszeder P. (2005) Estymacja współczynników beta na podstawie wielorównaniowego modelu GARCH. *Acta Universitatis Nicolai Copernici, Ekonomia XXXVI*, Uniwersytet Mikołaja Kopernika, Toruń.
- Gajdka J., Brzeszczyński J.: (2007) Estymacja parametru β przy użyciu modeli klasy ARCH, [w:] Tarczyński W. [red.], Rynek kapitałowy. Skuteczne inwestowanie. Część I, Wydawnictwo Naukowe Uniwersytetu Szczecińskiego, Szczecin.
- Henrikson R., Merton R.: On Market Timing and Investment Performance. II Statistical Procedures for Evaluating Forecasting Skills, *Journal of Business* 54, No. 4, s. 513 – 533, 1981.
- Henrikson R., Market Timing and Mutual Fund Performance. An Empirical Investigation, *Journal of Business* 57, s. 73 -96.
- Jajuga K., Jajuga T. (2005) Inwestycje. Instrumenty finansowe. Ryzyko finansowe. Inżynieria finansowa, Wydawnictwo Naukowe PWN, Warszawa.
- Kompa K., Matuszewska – Janica A. (2008) Charakterystyki opisowe i efektywność informacyjna wybranych instrumentów notowanych na GPW, W. Tarczyński (red.), Rynek Kapitałowy. Skuteczne inwestowanie. *Studia i prace Wydziału Nauk Ekonomicznych i Zarządzania, Uniwersytet Szczeciński*, Nr 9, Szczecin, s. 614 – 629.

Olbryś J. (2008) Parametryczne testy umiejętności wycucia rynku- porównanie wybranych metod na przykładzie OFI Akcji, Binderman Z. (red.), Metody ilościowe w badaniach ekonomicznych Nr 9, cz. I: Analiza rynków finansowych, modele ekonometryczne, Szkoła Główna Gospodarstwa Wiejskiego w Warszawie, Warszawa.

Ostrowska E. (2003) Efektywność funduszy inwestycyjnych na polskim rynku finansowym – wskaźniki Sharpe’a, Treynora i Jensena, [w] Prace naukowe Akademii Ekonomicznej we Wrocławiu, Inwestycje finansowe i ubezpieczenia – tendencje światowe a polski rynek, red. nauk. Jajuga K., Ronka-Chmielowiec W., Wydawnictwo Akademii Ekonomicznej we Wrocławiu, Wrocław.

Ostrowska E. (2005) Banki na rynkach kapitałowych, PWE, Warszawa.

Szyska A.: Efektywność Giełdy Papierów Wartościowych w Warszawie na tle rynków dojrzałych, Akademia Ekonomiczna w Poznaniu, Poznań 2003.

Tarczyński W. (1997) Rynki kapitałowe – Metody ilościowe. Tom 2, Agencja Wydawnicza Placet, Warszawa.

Treynor J. L., Mazuy K. (1966) Can Mutual Funds Outguess the Market?, Harvard Business Review, 44, s. 131 – 136

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

Witkowska D., Żebrowska – Suchodolska D. (2008) Badanie słabej formy efektywności informacyjnej GPW, W. Tarczyński (red.), Rynek Kapitałowy. Skuteczne inwestowanie. Studia i prace Wydziału Nauk Ekonomicznych i Zarządzania, Uniwersytet Szczeciński, Nr 9, Szczecin, s. 155 - 165.

Investigation of the Strong Form Efficient Market Hypothesis: the Example of Selected Investment Funds

Abstract: The aim of the research is evaluation of the efficiency of 19 selected open-end investment funds that have been operating at the Polish market. Analysis is based on daily data and covers the period from November, 2, 2006 to November, 2, 2008. Investigation is provided for the time series that contain observations from 3, 6, 12 and 24 months. The risk-free instruments are represented by selected treasury bonds and the market index is represented by the WIG (Warsaw Stock Exchange Index). In the research we apply Sharpe, Treynor and Jensen ratios as well as we investigate selectivity and market timing employing Treynor-Mauzy and Henriksson-Merton models.

Keywords: open-end investment funds, investment efficiency evaluation, selectivity, market timing

UOGÓLNIONA MIARA DOPASOWANIA W MODELU LINIOWYM

Wojciech Zieliński

Katedra Ekonometrii i Statystyki, SGGW
Nowoursynowska 159, PL-02-767 Warszawa
wojtek.zielinski@statystyka.info

Streszczenie: W modelu regresji liniowej stosowane są dwie miary wpływu zmiennych niezależnych na zmienną zależną: współczynnik determinacji oraz współczynnik korelacji cząstkowej. W pracy zaproponowane jest uogólnienie tych miar i skonstruowana jest miara wpływu grupy zmiennych niezależnych z wyłączeniem wpływu pozostałych zmiennych niezależnych.

Słowa kluczowe: miara dopasowania, współczynnik determinacji, korelacja cząstkowa, analiza regresji

Załóżmy, że obserwacje Y_i pewnej zmiennej losowej możemy przedstawić w postaci

$$Y_i = \beta_1 x_{1i} + \dots + \beta_p x_{pi} + \varepsilon_i, \quad i = 1, \dots, n \quad (*)$$

gdzie x_{ki} , $k = 1, \dots, p$, $i = 1, \dots, n$, są znane, β_1, \dots, β_p są nieznanymi współczynnikami regresji, natomiast ε_i , $i = 1, \dots, n$, są zmiennymi losowymi o rozkładach normalnych o zerowej wartości oczekiwanej i wariancji σ^2 .

Jednym z pytań stawianych w analizie modeli (*) jest pytanie o wpływ zmiennych niezależnych na cechę Y . W zastosowaniach wykorzystuje się w zasadzie dwie miary: miarę łącznego wpływu wszystkich zmiennych niezależnych oraz miarę wpływu pojedynczych zmiennych (z eliminacją wpływu pozostałych). Jest to powszechnie znany współczynnik determinacji oraz współczynnik korelacji cząstkowej. Ścisłe określenia tych pojęć można znaleźć w bardzo bogatej literaturze poświęconej analizie regresji (w spisie literatury podanych znaleźć kilka takich książek). Podane miary niestety nie udzielają odpowiedzi na pytania o wpływ grup wybranych zmiennych niezależnych na zmienną zależną, np. jak zmierzyć wpływ zmiennej x_1 oraz x_2 z wyłączeniem pozostałych zmiennych? W dalszym ciągu zaproponowana jest taka miara. Pokazano również, że współczynnik determinacji

oraz współczynnik korelacji cząstkowej są szczególnymi przypadkami tej ogólniejszej miary.

W dalszych rozważaniach znacznie wygodniej jest się posługiwać zapisami macierzowymi. Model liniowy (*) w zapisie macierzowym ma postać

$$Y = X\beta + \varepsilon$$

gdzie $Y' = [Y_1, \dots, Y_n]$ jest wektorem obserwacji, $\beta' = [\beta_1, \dots, \beta_p]$ jest wektorem nieznanymi parametrów, $\varepsilon' = [\varepsilon_1, \dots, \varepsilon_n]$ jest wektorem błędów losowych oraz

$$X = \begin{bmatrix} x_{11} & \cdots & x_{p1} \\ x_{12} & \cdots & x_{p2} \\ \vdots & \cdots & \vdots \\ x_{1n} & \cdots & x_{pn} \end{bmatrix}$$

jest macierzą eksperymentu. Zgodnie z poczynionymi założeniami wektor losowy Y ma n -wymiarowy rozkład normalny:

$$Y \sim N_n(X\beta, \sigma^2 I_n)$$

Niech teraz $X = [X_1 : X_2]$, gdzie macierze X_1 oraz X_2 są macierzami o wymiarach odpowiednio $n \times (p - q)$ i $n \times q$. Niech wektor parametrów β będzie podzielony odpowiednio na dwa podwektory β_1 i β_2 zgodnie z podziałem macierzy X . Interesuje nas zmierzenie wpływu na zmienną zależną zmiennych zawartych w macierzy X_2 . Niech $\hat{\beta}$ oznacza estymator wektora β w modelu z macierzą X

$$\hat{\beta} = (X'X)^{-1}X'Y$$

natomiast niech $\hat{\beta}_1$ będzie estymatorem najmniejszych kwadratów wektora β w modelu z ograniczeniami $\beta_2 = 0$ (Zieliński 2007):

$$\hat{\beta}_1 = \hat{\beta} - (X'X)^{-1}A[A'(X'X)^{-1}A]^{-1}A'\hat{\beta}$$

Tutaj $A = [0_{n \times (p-q)} : I_q]$ jest macierzą ograniczeń, tzn. $\beta_2 = 0$ jest równoważne temu, że $A\beta = 0$ (0 oznacza, w zależności od kontekstu, wektor lub macierz zerową, I_q oznacza macierz jednostkową wymiaru q).

Wektor obserwacji Y możemy zapisać jako

$$Y - X\hat{\beta}_1 = (X\hat{\beta} - X\hat{\beta}_1) + (Y - X\hat{\beta} + X\hat{\beta}_1)$$

Zauważmy, że wektory $X\hat{\beta} - X\hat{\beta}_1$ i $Y - X\hat{\beta} + X\hat{\beta}_1$ są ortogonalne. A zatem

$$\|Y - X\hat{\beta}_1\|^2 = \|X\hat{\beta} - X\hat{\beta}_1\|^2 + \|Y - X\hat{\beta} + X\hat{\beta}_1\|^2$$

gdzie $\|\cdot\|^2$ oznacza kwadrat długości wektora (w normie euklidesowej).

Jako miarę wpływu zmiennych zawartych w macierzy X_2 przyjmujemy

$$\mathfrak{R}^2(X_2) = \frac{\|X\hat{\beta} - X\hat{\beta}_1\|^2}{\|Y - X\hat{\beta}_1\|^2}$$

Twierdzenie. Zmienna losowa

$$\frac{\mathfrak{R}^2(X_2)}{1 - \mathfrak{R}^2(X_2)} \cdot \frac{n - q}{q}$$

ma niecentralny rozkład F z $(q, n - q)$ stopniami swobody i parametrem niecentralności $\beta' A [A' (X' X)^{-1} A]^{-1} A \beta$, gdzie $A = [\mathbf{0}_{n \times (p-q)} : I_q]$.

Dowód. W dowodzie korzystamy z dobrze znanych faktów dotyczących rozkładów prawdopodobieństwa form liniowych i kwadratowych wektorów losowych o wielowymiarowym rozkładzie normalnym.

Łatwo sprawdzić, że

$$\frac{\mathfrak{R}^2(X_2)}{1 - \mathfrak{R}^2(X_2)} = \frac{\|X\hat{\beta} - X\hat{\beta}_1\|^2}{\|Y - (X\hat{\beta} - X\hat{\beta}_1)\|^2} = \frac{\hat{\beta}' A' [A (X' X)^{-1} A'] A \hat{\beta}}{Y' (I_n - A' [A (X' X)^{-1} A'] A) Y}$$

Macierz $I_n - A' [A (X' X)^{-1} A'] A$ jest idempotentna, więc w mianowniku mamy zmienną losową o centralnym rozkładzie chi-kwadrat z $(n - q)$ stopniami swobody. Ponieważ wektor losowy $\hat{\beta}$ ma rozkład $N_p(\beta, \sigma^2 (X' X)^{-1})$, więc w liczniku mamy zmienną losową o niecentralnym rozkładzie chi-kwadrat z q stopniami swobody i parametrem niecentralności $\beta' A' [A (X' X)^{-1} A'] A \beta$. Licznik i Mianownik są niezależnymi zmiennymi losowymi i stąd wynika teza twierdzenia.

Pokażemy teraz, że współczynnik determinacji oraz współczynnik korelacji cząstkowej są szczególnymi przypadkami miary $\mathfrak{R}^2(X_2)$.

Współczynnik determinacji jest miarą wpływu wszystkich zmiennych niezależnych x_1, \dots, x_m na zmienną Y w modelu

$$Y_i = \beta_0 + \beta_1 x_{1i} + \dots + \beta_m x_{mi} + \varepsilon_i, \quad i = 1, \dots, n.$$

W zapisie macierzowym

$$\mathbf{X} = \begin{bmatrix} 1 & x_{11} & \cdots & x_{m1} \\ 1 & x_{12} & \cdots & x_{m2} \\ \vdots & \vdots & \cdots & \vdots \\ 1 & x_{1n} & \cdots & x_{mn} \end{bmatrix}, \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_m \end{bmatrix}.$$

Macierz \mathbf{X} i wektor parametrów można podzielić na $\mathbf{X}_1 = \mathbf{1}_n$,

$$\mathbf{X}_2 = \begin{bmatrix} x_{11} & \cdots & x_{m1} \\ x_{12} & \cdots & x_{m2} \\ \vdots & \cdots & \vdots \\ x_{1n} & \cdots & x_{mn} \end{bmatrix}, \quad \boldsymbol{\beta}_1 = [\beta_0], \quad \boldsymbol{\beta}_2 = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_m \end{bmatrix}.$$

W tym modelu mamy $p = m + 1$, $q = m$ oraz

$$\mathbf{X}'\mathbf{X} = \begin{bmatrix} n & \mathbf{1}'_n\mathbf{X}_2 \\ \mathbf{1}_n\mathbf{X}'_2 & \mathbf{X}'_2\mathbf{X}_2 \end{bmatrix}, \quad (\mathbf{X}'\mathbf{X})^{-1} = \begin{bmatrix} M & \mathbf{k}' \\ \mathbf{k} & L \end{bmatrix}$$

gdzie $L^{-1} = \mathbf{X}'_2\mathbf{X}_2 - \frac{1}{n}\mathbf{X}'_2\mathbf{1}_n\mathbf{1}'_n\mathbf{X}_2$, $\mathbf{k}' = -\frac{1}{n}\mathbf{1}'_n\mathbf{X}_2L$ oraz $M = \frac{1}{n} + \mathbf{k}'L\mathbf{k}$ (ogólny wzór na odwrotność blokowej macierzy 2×2 można znaleźć w np. Rao 1982, Zieliński 2007). Estymator najmniejszych kwadratów wektora $\boldsymbol{\beta}$ w pełnym modelu ma postać

$$\hat{\boldsymbol{\beta}} = \begin{bmatrix} M & \mathbf{k}' \\ \mathbf{k} & L \end{bmatrix} \begin{bmatrix} \mathbf{1}'_n \\ \mathbf{X}'_2 \end{bmatrix} \mathbf{Y} = \begin{bmatrix} M\mathbf{1}'_n + \mathbf{k}'\mathbf{X}'_2 \\ \mathbf{k}\mathbf{1}'_n + L\mathbf{X}'_2 \end{bmatrix} \mathbf{Y}$$

tzn.

$$\hat{b}_0 = (M\mathbf{1}'_n + \mathbf{k}'\mathbf{X}'_2)\mathbf{Y}, \quad \hat{\boldsymbol{\beta}}_2 = (\mathbf{k}\mathbf{1}'_n + L\mathbf{X}'_2)\mathbf{Y}$$

Estymator wektora $\boldsymbol{\beta}_1$ w modelu z ograniczeniami określonymi macierzą $\mathbf{A} = [\mathbf{0}_m : \mathbf{I}_m]$ wyraża się wzorem

$$\begin{aligned} \hat{\boldsymbol{\beta}}_1 &= \left(\mathbf{I}_{m+1} - \begin{bmatrix} M & \mathbf{k}' \\ \mathbf{k} & L \end{bmatrix} \begin{bmatrix} \mathbf{0}'_m \\ \mathbf{I}_m \end{bmatrix} \left(\begin{bmatrix} \mathbf{0}_m & \mathbf{I}_m \end{bmatrix} \begin{bmatrix} M & \mathbf{k}' \\ \mathbf{k} & L \end{bmatrix} \begin{bmatrix} \mathbf{0}'_m \\ \mathbf{I}_m \end{bmatrix} \right)^{-1} \begin{bmatrix} \mathbf{0}_m & \mathbf{I}_m \end{bmatrix} \right) \hat{\boldsymbol{\beta}} \\ &= \begin{bmatrix} 1 & -\mathbf{k}'L^{-1} \\ \mathbf{0}_m & \mathbf{0}_{m \times m} \end{bmatrix} \hat{\boldsymbol{\beta}} = \begin{bmatrix} \hat{b}_0 - \mathbf{k}'L^{-1}\hat{\boldsymbol{\beta}}_2 \\ \mathbf{0}_m \end{bmatrix} = \begin{bmatrix} \frac{1}{n} \mathbf{1}'_n \mathbf{Y} \\ \mathbf{0}_m \end{bmatrix} = \begin{bmatrix} \bar{Y} \\ \mathbf{0}_m \end{bmatrix} \end{aligned}$$

gdzie $\bar{Y} = \frac{1}{n} \sum Y_i$. Ponieważ $\mathbf{X}\hat{\boldsymbol{\beta}}_1 = Y\mathbf{1}_n$, więc $(\hat{Y} = \mathbf{X}\hat{\boldsymbol{\beta}})$

$$\|\mathbf{X}\hat{\boldsymbol{\beta}} - \mathbf{X}\hat{\boldsymbol{\beta}}_1\|^2 = (\hat{Y} - Y\mathbf{1}_n)'(\hat{Y} - Y\mathbf{1}_n) = \sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2$$

oraz

$$\|Y - X\hat{\beta}_1\|^2 = (Y - \bar{Y} \mathbf{1}_n)'(Y - \bar{Y} \mathbf{1}_n) = \sum_{i=1}^n (Y_i - \bar{Y})^2$$

Współczynnik determinacji, wyrażany zwyczajowo w procentach, określony jest więc w następujący sposób

$$D = \mathfrak{R}^2(\mathbf{X}_2) \cdot 100\% = \frac{\Sigma(\hat{Y}_i - \bar{Y})^2}{\Sigma(Y_i - \bar{Y})^2}$$

Z twierdzenia otrzymujemy

Wniosek. Zmienna losowa

$$\frac{D}{100 - D} \cdot \frac{n - m}{m}$$

ma niecentralny rozkład F z $(m, n - m)$ stopniami swobody i parametrem niecentralności $\beta_2' \mathbf{X}_2' (\mathbf{I}_n - \mathbf{1}_n \mathbf{1}_n') \mathbf{X}_2 \beta_2$.

Współczynnik korelacji cząstkowej jest miarą wpływu jednej ze zmiennych niezależnych na zmienną Y w modelu

$$Y_i = \beta_0 + \beta_1 x_{1i} + \dots + \beta_m x_{mi} + \varepsilon_i, \quad i = 1, \dots, n.$$

Wyprowadzimy odpowiedni wzór dla zmiennej x_1 . Dla uproszczenia zapisu model zapisujemy w postaci

$$Y_i = \beta_0 + \beta_2 x_{2i} + \dots + \beta_m x_{mi} + \beta_1 x_{1i} + \varepsilon_i, \quad i = 1, \dots, n.$$

W zapisie macierzowym

$$\mathbf{X} = \begin{bmatrix} 1 & x_{21} & \dots & x_{m1} & x_{11} \\ 1 & x_{22} & \dots & x_{m2} & x_{12} \\ \vdots & \vdots & \dots & \vdots & \vdots \\ 1 & x_{2n} & \dots & x_{mn} & x_{1n} \end{bmatrix}, \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_2 \\ \vdots \\ \beta_m \\ \beta_1 \end{bmatrix}.$$

Macierz \mathbf{X} i wektor parametrów można podzielić na

$$\mathbf{X}_1 = \begin{bmatrix} 1 & x_{21} & \dots & x_{m1} \\ 1 & x_{22} & \dots & x_{m2} \\ \vdots & \vdots & \dots & \vdots \\ 1 & x_{2n} & \dots & x_{mn} \end{bmatrix} \quad \mathbf{x}_2 = \begin{bmatrix} x_{11} \\ x_{12} \\ \vdots \\ x_{1n} \end{bmatrix} \quad \boldsymbol{\beta}_1 = \begin{bmatrix} \beta_0 \\ \beta_2 \\ \vdots \\ \beta_m \end{bmatrix} \quad \boldsymbol{\beta} = [\boldsymbol{\beta}_1]$$

W tym modelu

$$\mathbf{X}'\mathbf{X} = \begin{bmatrix} \mathbf{X}'_1 \mathbf{X}_1 & \mathbf{X}'_1 \mathbf{x}_2 \\ \mathbf{x}'_2 \mathbf{X}_1 & \mathbf{x}'_2 \mathbf{x}_2 \end{bmatrix} \quad (\mathbf{X}'\mathbf{X})^{-1} = \begin{bmatrix} \mathbf{M} & \mathbf{k} \\ \mathbf{k}' & L \end{bmatrix}$$

gdzie

$$L = \frac{1}{\mathbf{x}'_2 \mathbf{x}_2 - \mathbf{x}'_2 \mathbf{X}_1 (\mathbf{X}'_1 \mathbf{X}_1)^{-1} \mathbf{X}'_1 \mathbf{x}_2}$$

$$\mathbf{k}' = -L(\mathbf{X}'_1 \mathbf{X}_1)^{-1} \mathbf{X}'_1 \mathbf{x}_2 \mathbf{X}_1, \quad \mathbf{M} = (\mathbf{X}'_1 \mathbf{X}_1)^{-1} - L \mathbf{k} \mathbf{k}'.$$

Ograniczenia opisane są za pomocą wektora $[\mathbf{0}'_m : 1]$. Po wykonaniu odpowiednich obliczeń otrzymujemy, że miarą wpływu zmiennej x_1 na Y jest

$$\mathfrak{R}^2(x_2) = \frac{(\mathbf{x}'_2 (\mathbf{I}_n - \mathbf{X}_1 (\mathbf{X}'_1 \mathbf{X}_1)^{-1} \mathbf{X}'_1) \mathbf{Y})^2}{(\mathbf{Y}' (\mathbf{I}_n - \mathbf{X}_1 (\mathbf{X}'_1 \mathbf{X}_1)^{-1} \mathbf{X}'_1) \mathbf{Y}) (\mathbf{x}'_2 (\mathbf{I}_n - \mathbf{X}_1 (\mathbf{X}'_1 \mathbf{X}_1)^{-1} \mathbf{X}'_1) \mathbf{x}_2)}$$

Współczynnik korelacji cząstkowej między Y a x_1 określony jest jako $\sqrt{\mathfrak{R}^2(x_2)}$ i oznaczany $R_{Y(x_1)(x_2, \dots, x_m)}$. Znak tego współczynnika jest oczywiście zgodny ze znakiem $\mathbf{x}'_2 (\mathbf{I}_n - \mathbf{X}_1 (\mathbf{X}'_1 \mathbf{X}_1)^{-1} \mathbf{X}'_1) \mathbf{Y}$. Z twierdzenia otrzymujemy następujący wniosek

Wniosek. Zmienna losowa

$$\frac{R_{Y(x_1)(x_2, \dots, x_m)}^2}{1 - R_{Y(x_1)(x_2, \dots, x_m)}^2} \cdot (n - m - 1)$$

ma niecentralny rozkład F z $(1, n - m - 1)$ stopniami swobody i parametrem niecentralności $\beta_2^2 \mathbf{x}'_2 (\mathbf{I}_n - \mathbf{X}_1 (\mathbf{X}'_1 \mathbf{X}_1)^{-1} \mathbf{X}'_1) \mathbf{x}_2$.

Dodatkowa zmienna. Niech $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$. Przypuśćmy, że do modelu włączamy jeszcze jedną zmienną niezależną, tzn. $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{z}\gamma + \boldsymbol{\eta} = \mathbf{W}\boldsymbol{\delta} + \boldsymbol{\eta}$, gdzie $\mathbf{W} = [\mathbf{X} : \mathbf{z}]$, $\boldsymbol{\delta} = [\boldsymbol{\beta} : \gamma]'$ oraz $\boldsymbol{\eta}$ jest wektorem błędów losowych. Tutaj \mathbf{z} jest wektorem n -wymiarowym reprezentującym włączaną zmienną, natomiast γ jest nieznanym współczynnikiem regresji. Wyznamy współczynnik $\mathfrak{R}^2(\mathbf{W})$ i porównamy go z $\mathfrak{R}^2(\mathbf{X})$.

Estymatorem najmniejszych kwadratów wektora $\boldsymbol{\delta}$ jest oczywiście $\hat{\boldsymbol{\delta}} = (\mathbf{W}'\mathbf{W})^{-1} \mathbf{W}'\mathbf{Y}$. Mamy

$$\mathbf{W}'\mathbf{W} = \begin{bmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{z} \\ \mathbf{z}'\mathbf{X} & \mathbf{z}'\mathbf{z} \end{bmatrix}$$

oraz

$$(\mathbf{W}'\mathbf{W})^{-1} = \frac{1}{c} \begin{bmatrix} c(\mathbf{X}'\mathbf{X})^{-1} + \mathbf{a}\mathbf{a}' & -\mathbf{a} \\ -\mathbf{a}' & 1 \end{bmatrix}$$

gdzie $\mathbf{a} = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{z}$ oraz $c = \mathbf{z}'(\mathbf{I}_n - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}')\mathbf{z}$. Zatem

$$\mathbf{d} = \begin{bmatrix} \mathbf{b} - \frac{\mathbf{z}'(\mathbf{I}_n - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\mathbf{Y}}{\mathbf{z}'(\mathbf{I}_n - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\mathbf{z}}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{z} \\ \frac{\mathbf{z}'(\mathbf{I}_n - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\mathbf{Y}}{\mathbf{z}'(\mathbf{I}_n - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\mathbf{z}} \end{bmatrix}$$

Ponieważ

$$\mathbf{W}\mathbf{d} = \mathbf{X}\mathbf{b} + \frac{\mathbf{z}'(\mathbf{I}_n - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\mathbf{Y}}{\mathbf{z}'(\mathbf{I}_n - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\mathbf{z}}(\mathbf{I}_n - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\mathbf{z}$$

więc

$$\|\mathbf{W}\mathbf{d}\|^2 = \|\mathbf{X}\mathbf{b}\|^2 + \frac{(\mathbf{z}'(\mathbf{I}_n - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\mathbf{Y})^2}{\mathbf{z}'(\mathbf{I}_n - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\mathbf{z}}$$

Zatem $\mathfrak{R}^2(\mathbf{W}) \geq \mathfrak{R}^2(\mathbf{X})$, tzn. model z dodatkową zmienną nie gorzej odtwarza zmienną zależną niż model bez niej.

LITERATURA

- Borkowski B., Dudek H., Szczesny W. 2003: Ekonometria, wybrane zagadnienia, PWN, Warszawa.
- Draper N. R., Smith H. 1973: Analiza regresji stosowana, PWN, Warszawa.
- Rao C. R. 1982: Modele liniowe statystyki matematycznej, PWN, Warszawa.
- Seber G. A. F. 1977: Linear Regression Analysis, Wiley, New York.
- Zieliński W. 2007: Teoretyczne podstawy ekonometrycznych jednorównaniowych modeli liniowych, Wydawnictwo SGGW, Warszawa.

Generalized measure of fit in a linear model

Abstract: In a model of linear regression two measures of fit are used: coefficient of determination and coefficient of partial correlation. A generalization of those measures is proposed and a measure of influence of a group of independent variables excluding an influence of remaining variables is constructed.

Keywords: measure of fit, determination coefficient, partial correlation, linear regression