MANAGERIAL ECONOMICS

SEMIANNUAL

vol. 17 no. 1 ▪ 2016

Wydział Zarządzania Akademii Górniczo-Hutniczej

EKONOMIA
MENEDŻERSKA

PÓŁROCZNIK

tom 17 nr 1 ■ 2016

WYDAWNICTWA AGH

KRAKÓW 2016

Faculty of Management
AGH University of Science and Technology

MANAGERIAL ECONOMICS

SEMIANNUAL
vol. 17 no. 1 ■ 2016

AGH UNIVERSITY OF SCIENCE AND TECHNOLOGY PRESS KRAKOW 2016

Articles published in the semiannual Managerial Economics have been peer reviewed by reviewers appointed by the Editorial Board. The procedure for reviewing articles is described at the following web address: http://www.managerial.zarz.agh.edu.pl

Language Editor: *Bret Spainhour*

Statistical Editor: *Anna Barańska*

Editorial support: *Magdalena Grzech*

Cover and title page design: *Zofia Łucka*

DOI: http://dx.doi.org/10.7494/manage

 ISSN 1898-1143 (paper version),
ISSN 2353-3617 (on-line version)

Number of copies 45. The printed version of the journal is the primary one.

Redakcja Wydawnictw AGH
al. A. Mickiewicza 30, 30-059 Kraków
tel. 12 617 32 28, tel./faks 12 636 40 38
e-mail: redakcja@wydawnictwoagh.pl; http://www.wydawnictwa.agh.edu.pl

CONTENTS

Managerial Economics
2016, vol. 17, no. 1, pp. 7–20
http://dx.doi.org/10.7494/manage.2016.17.1.7

Anna Czapkiewicz*, Marta Stachowicz**

The long-run relationship between the stock market and main macroeconomic variables in Poland

1. Introduction

In the financial literature, there are some studies concerning long-term relationships between financial markets. For example, Chen et al. (2002) investigated the dynamic interdependence of the major stock markets in Latin America. Using data from 1995 to 2000, they found that there is one cointegrating vector that appears to explain the dependencies in prices of the the stock market indices of Argentina, Brazil, Chile, Colombia, Mexico, and Venezuela. On the basis of daily data from the period 1993–2002, Voronkova (2004) showed the existence of cointegration between European developed markets and the stock markets of Czech Republic, Hungary, and Poland. Additionally, Syriopoulos (2007) indicated that the long-term linkages between emerging CEE markets (Czech Republic, Hungary, Poland, and Slovakia) and developed markets (Germany and the US) are stronger than among the CEE countries themselves.

Furthermore, researchers also studied the relationship between stock markets and exchange rates. Each of these variables play crucial roles in influencing the development of a country's economy (Nieh and Lee, 2001); therefore, this subject has drawn the attention of investors and policy makers. This relationship can be used by practitioners to predict future trends for each other. Among the macroeconomic indicators, exchange rates are able to influence stock prices through trade effect (Geske and Roll, 1983). Moore (2007) investigated the effect of the euro on stock markets for Hungary, Poland, and the UK, and also the co-movement

* AGH University of Science and Technology in Krakow, Faculty of Management, Department of Applications of Mathematics in Economics, e-mail: gzrembie@cyf-kr.edu.pl
** Student, AGH University of Science and Technology in Krakow, Faculty of Applied Mathematics.

of stock prices with the euro-zone using the daily stock price indices. The result reveals that the exchange rate is a more-important factor for Hungary than for Poland. The existence of long-run equilibrium relationships among stock prices, industrial production, real exchange rates, interest rates, and inflation in the United States was investigated by Kim (2003). Applying Johansen's cointegration analysis for monthly data for the period 1974–1998, they found that the S&P 500 stock price is positively related to industrial production yet negatively related to the real exchange rate, interest rate, and inflation. Analysis of the error-correction mechanism revealed that the stock price, industrial production, and inflation adjust to correct disequilibrium among the five variables, while variance decompositions indicate that the stock price is driven by innovations in the interest rate.

What is more, Rjoub (2012) studied the long-run relationship between Turkish stock prices, the exchange rates, and US stock prices to reflect world trends. He found that the floating exchange rate has a negative impact on the Turkish stock market; therefore, a decline in the value of the Turkish national currency is expected to stimulate domestic economic activity. Abdelaziz et al. (2008) used cointegration analysis to investigate the long-run dynamics among domestic stock prices, the global market index represented by the US stock market, oil prices, and the real exchange rates in four oil-exporting Middle East countries. A long-run equilibrium relationship was found among stock prices, real exchange rates, and oil prices for three countries: Egypt, Oman, and Saudi Arabia.

The issue of relationships between stock prices and some economic variables was taken into consideration by Fifield et al. (2002), Bhattacharya and Mukherjee (2006), Brahmasrene and Jiranyakul (2007), Humpe and Macmillan (2007), Mahmood and Dinniah (2009), and Barbic and Condic-Jurkic (2011), among others.

Capital markets play a vital role in achieving sustainable economic growth. Stock markets are particularly important to economic development. Traditional models of the economy suggest that a link exists between stock market performance, exchange rate behavior and the levels of export and import. Dornbusch and Fischer (1980) suggest that exchange rate fluctuations affect the competitiveness of firms, as changes in the exchange rate affect the value of the earnings. The depreciation of local currency makes exporting goods attractive and leads to an increase in stock prices, as growing foreign demand increases a firm's revenues (which positively affects its) value.

The aim of this paper is to investigate the long-run dynamics among some macroeconomic variables, the EUR/PLN exchange rate, and stock prices on the Warsaw Stock Exchange (WSE). For this purpose, a cointegration analysis has been performed using the procedure from Johansen (1988). In this paper, we search the long-term relationship between the stock market, EUR/PLN exchange rate, export volume, and rate of inflation measured by the Consumer Price Index (CPI).

The collected data set exhibits some problems with seasonal and cyclical variations. Such variables may be seasonally adjusted by applying the various types of filtering described in Lutkepohl (2004) and Otsu (2009). Loef and Franses (2000) proved empirically that the seasonal cointegration models yield to more-accurate forecasts only for longer horizons, while the nonseasonal models outperform the other methods in shorter horizons. Lutkepohl (2005) showed that different treatments of seasonal data affect the system and change the impulse response functions.

The rest of the paper is organized as follows: in the next section, we give a short description of the methodology applied in the paper. In Section 3, we present and analyze in detail the data that we use in the empirical study. A short summary concludes the paper.

2. Methodology

Model VEC

The aim of a cointegration analysis is to determine if there are any common stochastic trends between the considered variables. If there is a common stochastic trend, the variables will share a common long-run equilibrium.

Let $Y_t = (Y_{1t}, \ldots, Y_{Kt})$ denotes a vector of K variables. After a determination, an order of integration of Y_t the Johansen's methodology (1988) is applied to investigate the cointegration between the considered time series.

When Y_t is integrated of the order $I(1)$, the VECM model is built, which has a following form:

$$\Delta Y_t = \Pi Y_{t-1} + \Gamma_1 \Delta Y_{t-1} + \ldots + \Gamma_{p-1} \Delta Y_{t-p+1} + \varepsilon_t \tag{1}$$

When **rank** $(\Pi) = r$ and $r < K$, then there are r cointegration relations, and matrix Π can be expressed as $\Pi = \alpha\beta^T$.

Each component of $\beta^T Y_{t-1}$ is stationary and defines a long-run equilibrium relationship.

Testing for the cointegration between components of Y_t is equivalent to testing the rank of matrix Π. To test for the cointegration rank, Johansen proposes two different pairs of hypotheses, which are:

$$H_0\text{: } rank\ (\Pi) = r \text{ against } H_1\text{: } r < rank\ (\Pi) \le K \tag{2}$$

and

$$H_0\text{: } rank\ (\Pi) = r \text{ against } H_1\text{: } rank\ (\Pi) \le r + 1 \tag{3}$$

A test statistic testing the first hypothesis is referred to as the trace statistic, whereas a test statistic used in the second hypothesis is referred as the maximum eigenvalue statistic.

Let λ_1, ..., λ_s be eigenvalues of the Π matrix put in ascending order. The trace statistic and the maximum eigenvalue statistic have following formulas:

$$\lambda_{trace}\left(r,K\right) = -T\sum_{i=r+1}^{K}\log(1-\hat{\lambda}_i) \tag{4}$$

and

$$\lambda_{\max}\left(r,r+1\right) = -T\log(1-\hat{\lambda}_{\max}) \tag{5}$$

Both tests are performed step-by-step until the null hypothesis can no longer be rejected.

Impulse Response Function

The previous section discussed the interrelations between the variables of a system. However, knowledge about the response of one variable to an impulse of another variable seems to be more interesting. Therefore, in what follows, the impulse-response analysis is carried on. Let model VEC be presented as the following VAR model:

$$Y_t = \mu + A_1Y_{t-1} + A_2Y_{t-2} + \ldots + A_{p-1}Y_{t-p+1} + \varepsilon_t \tag{6}$$

where A_1, …, A_{p-1} are autoregressive parameter matrices and ε_t is a vector of residuals. It is assumed that $\varepsilon_t \sim N(0,\Sigma)$.

Matrix Σ can be decomposed as $\Sigma = W\Sigma_e W^T$, where Σ_e is the diagonal matrix. Then, process Y_t can be written equivalently as follows:

$$Y_t = \mu^* + A_0^*Y_t + A_1^*Y_{t-1} + A_2^*Y_{t-2} + \ldots + A_{p-1}^*Y_{t-p+1} + e_t \tag{7}$$

where $A^*_i = W^{-1}A_i$
or as its moving average representation:

$$Y_t = \mu^* + \Psi_0^*e_t + \Psi_1^*e_{t-1} + \Psi_2^*e_{t-2} + \ldots \tag{8}$$

where $e_t = W^{-1}\varepsilon_t$, $\Psi^*_0 = W$ and $\Psi^*_k = \Psi_k W$ for $k = 1, 2, \ldots$

Matrices Ψ_1, Ψ_2 ... come from the moving-average representation of (6). This means that element of the matrix $\Psi^*_{ij,k}$ represents the influence of the $e_{j,t}$ disturbance on $Y_{i,t+k}$.

3. Empirical study

Data

We study the long-term relationship between the Warsaw stock market represented by the WIG20 index, EUR/PLN exchange rate, export volume, and rate of inflation (CPI). The WIG20 prices and monthly EUR/PLN exchange rates were collected from the stooq.pl database, the CPI data is collected from the Central Statistical Office of Poland (GUS), and the Export data is derived from the Eurostat base. In empirical analysis, the logarithm of monthly closing price of WIG20 is considered, the CPI was obtained by comparing the prices of the considered month to the prices of the period where the index was equal to 100. The Export data is given in million EUR and represents the value of goods exported in t months.

The period under study covers the time from January to December, 2015. This period contains different market phases; particularly, it covers the period of the global financial crisis in 2008, which is visible in data spikes on the times-series plots (Figure 1, Figure 2).

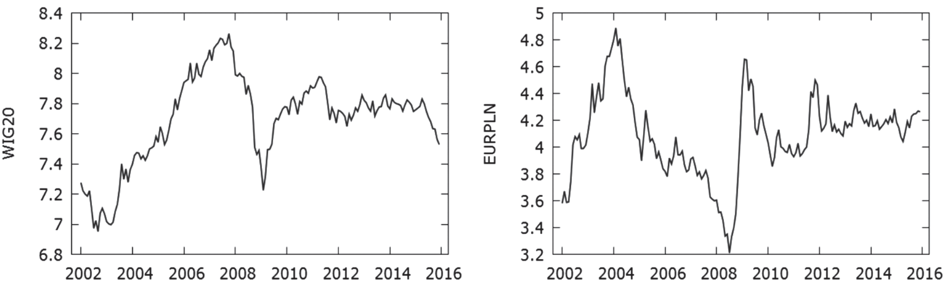

Figure 1. The plots of monthly prices of WIG20 and the monthly EUR/PLN exchange rate

Source: author's own research

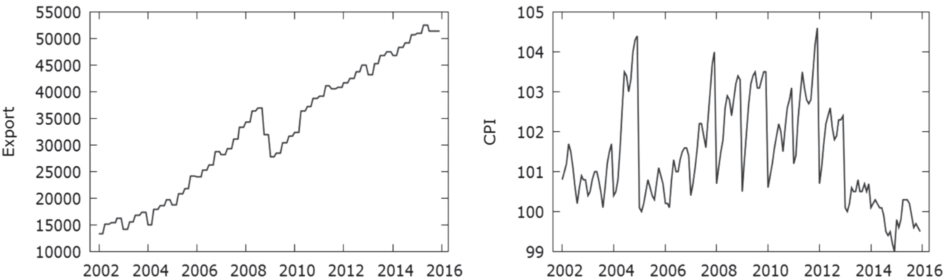

Figure 2. The plots of monthly export in million EUR and monthly Consumer Price Index

Source: author's own research

Firstly, all time series denoted as WIG20, EUR/PLN, CPI, and Export have been tested to be stationary or not. The results of an ADF test (including trend) indicate that variables WIG20, EUR/PLN, and CPI are $I(1)$, whereas Export is $I(0)$. These results are confirmed by a KPSS test. However, the vector consisting of all analyzed variables is $I(1)$ (Lutkepohl, 2005), so the one from the assumptions of possibility of the VECM building is fulfilled.

The data plots (Figure 1 and 2) suggest the presence of seasonality and outliers in some of the variables.

Therefore, all variables under study were tested for the existence of seasonal patterns, level shifts, or outliers. The data has been filtered by employing the X-12-ARIMA algorithm. This filter is used to detect and adjust outliers and other distorting effects. According to proper test results, seasonal moving-average filtering was applied only for the CPI variable, whereas the other variables were corrected for outliers and level-shift existence.

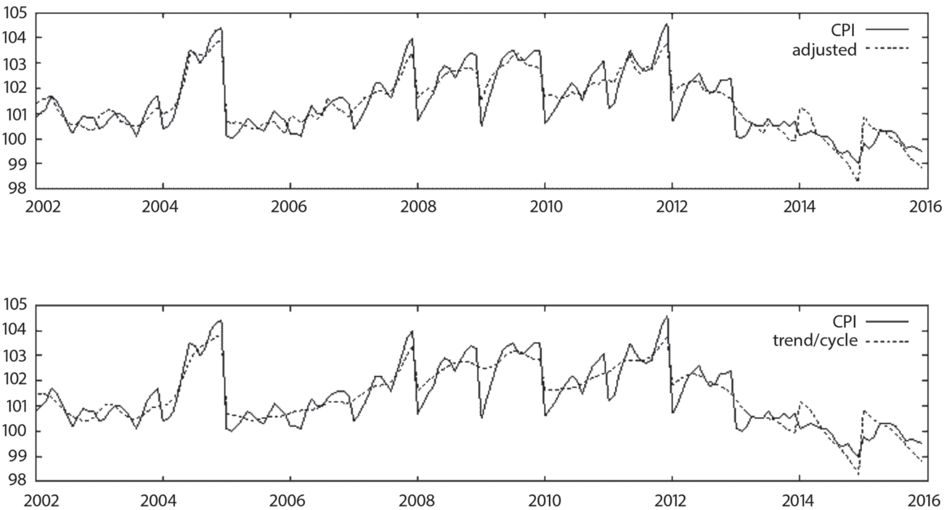

Figure 3. Filtered CPI data

Figures 3 and 4 present the seasonally adjusted series and estimated trends for the CPI and corrected EUR/PLN rate compared with the original data plots. The seasonally adjusted series also incorporate corrections for irregular fluctuations represented by sampling and non-sampling errors or strikes. We can observe significant differences between the adjusted and original series for the CPI, as this was additionally filtered in order to remove seasonality.

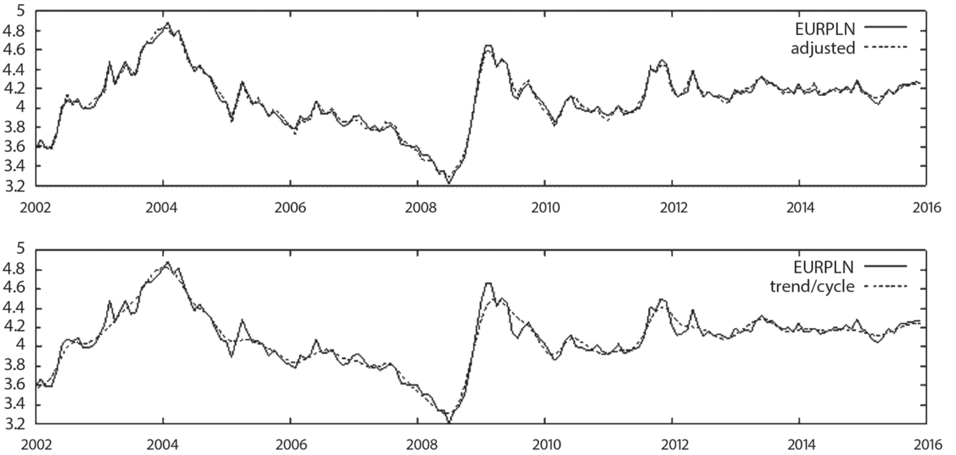

Figure 4. Filtered EUR/PLN data

Model VEC estimation

In order to choose the optimal lag length in model VEC, some information criteria have been used: Akaike information criterion (AIC), Hannan–Quinn (H-Q) information criterion (HQC), Schwartz criterion (SC), and Final Prediction Error criterion (FPE). All the information criteria prefer the model with one lag; so, taking into account the data frequency, we assume one lag length in the VEC model.

Next, the Johansen trace test with linear trend and test of maximum eigenvalue with linear trend have been applied to verify the existence of a long-run relationship between the studied time series. The obtained results coming from both tests indicate the existence of one cointegrating relationship between the index of the stock market and the considered economic variables. Applying the Johansen trace test, we obtained *p-value*=0.00 for *rank*(Π)=1 and *p-value*=0.06 for *rank*(Π)=2, whereas applying the Johansen maximum eigenvalue test, we obtained *p-value*=0.05 for *rank*(Π)=1 and *p-value*=0.13 for *rank*(Π)=2. So, we obtained one long-term relationship between variables WIG20, Export, CPI, and EUR/PLN. The results are presented in Table 1.

The cointegrating equation normalized with respect to the WIG20 variable has the following form:

$$\begin{aligned}\mathrm{WIG20}_t = -0.046867\ \mathrm{trend} - 0.57937\ \mathrm{EURPLN}_t + \\ 0.00022358\ \mathrm{Export}_t + 0{,}22022\mathrm{CPI}_t + \varepsilon_t\end{aligned} \tag{9}$$

The equation above represents the long-run equilibrium for WIG20 and the other variables.

Table 1

Parameters of cointegrating relationship

Variable	β	α
WIG20	1.00	0.04
EURPLN	0.58	−0.03
Export	−0.00	464.77
CPI	−0.22	0.12
Trend	0.05	–

Source: authors' calculations

As the β coefficient for WIG20 has been normalized to 1, the other beta parameters can be interpreted as the level of impact of the system variable related to this parameter on the stock market behavior. We can notice that the Export variable has neglected impact on the stock market, since parameter β is close to zero. According to our expectations, the CPI variable impacts positively on WIG20; however, the level of this impact is rather moderate. The EUR/PLN exchange rate has the biggest impact on WIG20 fluctuations. The results confirmed the fact that the exchange rate plays a crucial role to influence stock prices through trade effect (Geske and Roll, 1983). Since the parameter beta is negative, this implies that EUR/PLN has the opposite impact on the stock index value.

Parameter α in each equation reflects the speed of variable adjustment to the equilibrium relation. The sign of this parameter depends on the direction of the variables' disturbance from the equilibrium relation. If the variables are above the equilibrium, a negative α denotes that they get back to the equilibrium in the next time periods. When α is positive, the variables move away from the equilibrium state. A bigger α (in absolute value) is equivalent to the faster convergence of variables to the equilibrium.

From the results obtained, we can conclude, that Export (having a very low impact on WIG20) departs the equilibrium very quickly, while CPI moves away from this state in a slower phase. EUR/PLN relatively quickly adjusts itself to the equilibrium relationship. As α corresponding to WIG20 is very close to zero, the variable achieves the equilibrium relationship with a very low speed.

Table 2 presents the estimated short-term relationships between the system variables. We can conclude that the WIG20 depends on its lagged values and on the lag of the EUR/PLN interest rate. The EUR/PLN exchange rate does not depend only on the lag of CPI. The Export is being explained only by the lag of EUR/PLN, while CPI is not being explained by the lags of any of the variables creating the system.

Table 2

Coefficients of short-term relation

	Δ**Wig20**$_t$	Δ**CPI**$_t$	Δ**EURPLN**$_t$	Δ**Export**$_t$
Δ**Wig20**$_{t-1}$	−0.24**	−0.69	0.29*	−1796.53
Δ**CPI**$_{t-1}$	−0.00	0.01	0.01	101.53
Δ**EURPLN**$_{t-1}$	−0.10**	0.10	0.22**	−1651.32*
Δ**Export**$_{t-1}$	0.00***	0.00	−0.00***	464.77

Source: authors' calculations

Impulse response function

Figure 5 presents the results of the responses of each variable to one standard deviation shock coming from other variable. We observe that shocks have a permanent impact on each considered variable. In each case, the impulse in one variable causes an immediate increase or decrease in the value of another variable. The effect of shock disappears after some time, and a new level of equilibrium is established. Firstly, let us consider the WIG20 response to a shock coming from other variables. A one-standard-deviation shock to WIG20 causes that, in the beginning, it decreases slightly; later, it stabilizes at a new level of equilibrium lower than the initial state.

WIG20 responds to a one-standard-deviation shock coming from the EUR/PLN exchange rate in a different manner. In the beginning, it decreases a little; later, it grows to reach maximum; and finally, it also decreases and stabilizes at a new level of equilibrium higher than the initial state.

The WIG20 responds to a one-standard-deviation shock coming from Export or CPI quite differently than in the previous two cases. When an impulse comes from the Export variable, the stock index price increases a little in the beginning; later, it decreases to reach a new level of equilibrium lower than the initial state. But when the impulse comes from CPI variable, the stock index price decreases to reach a new level of equilibrium that is also lower than the initial state.

The rest of the analysis cases of impulse response of a given variable to a shock coming from other variables are also presented in Figure 2. We can notice that EUR/PLN response to a shock coming from WIG20 first rapidly increases and finally stabilizes at a new level of equilibrium higher than the initial state. The EUR/PLN variable responds to a one-standard-deviation shock coming from EUR/PLN quite differently: soon after the appearance of the shock, it rapidly increases, then decreases, and finally stabilizes at a new level of equilibrium much lower than the initial state. The EUR/PLN variable after a one-standard-deviation shock coming from CPI increases and stabilizes at a new level of equilibrium much higher than the initial state.

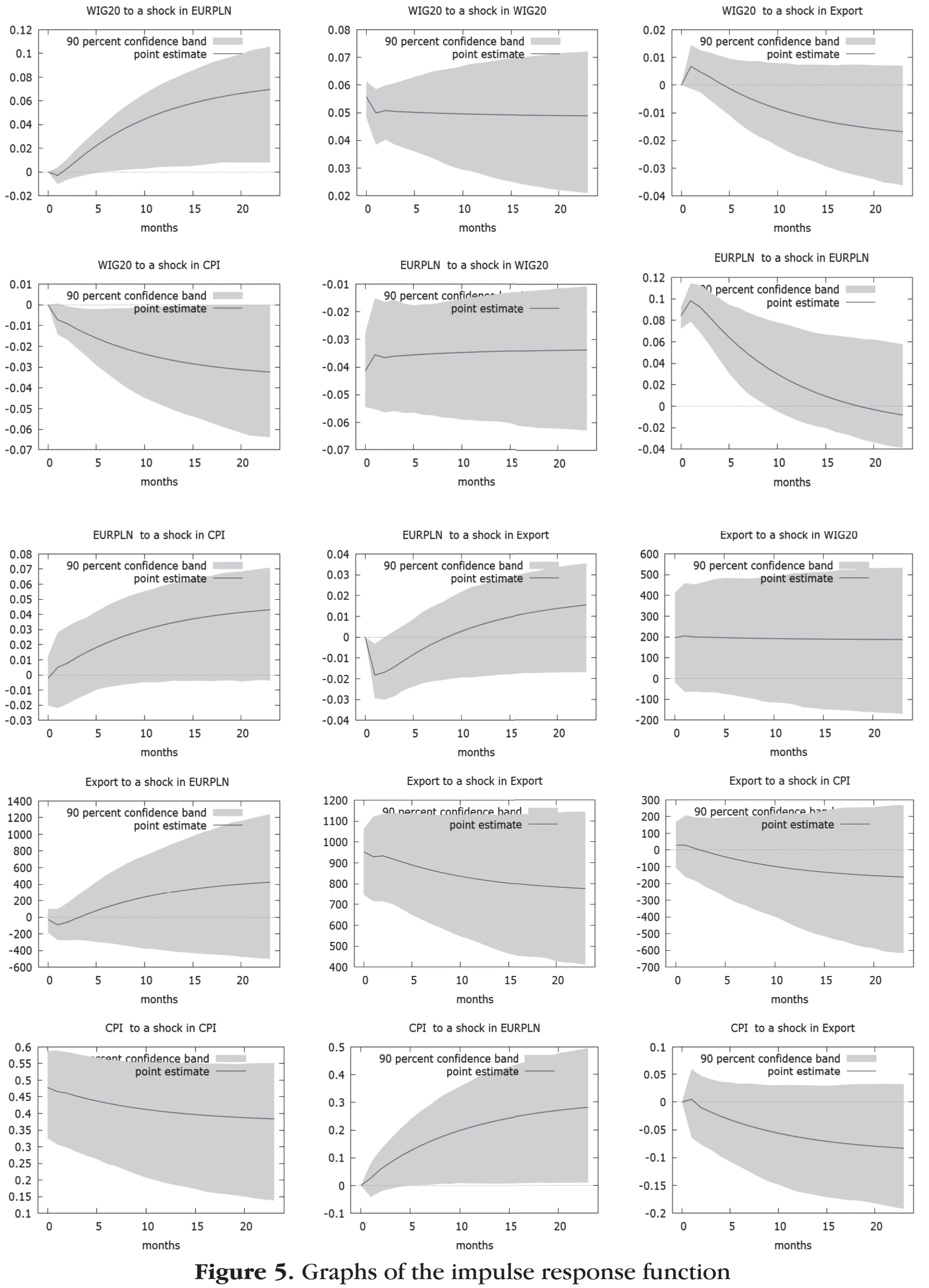

Figure 5. Graphs of the impulse response function

Source: authors' own research

After a one-standard-deviation shock coming from Export, EUR/PLN rapidly decreases, then increases, and finally stabilizes at a new level of equilibrium also much higher than the initial state.

The Export variable reaction is not noticeable after a shock coming from WIG20. But after a shock coming from EUR/PLN, it increases to new level of equilibrium higher than the initial state, whereas after a shock coming from Export or CPI, it stabilizes at a new level of equilibrium a little lower than the initial state.

The CPI variable behaves similarly to the Export variable after a shock coming from other variables. When the shock comes from EUR/PLN (similar to WIG20), its reaction stabilizes at a new level of equilibrium higher than the initial state, whereas when a shock comes from Export or CPI variables, its reaction stabilizes at a level of equilibrium lower than the initial state.

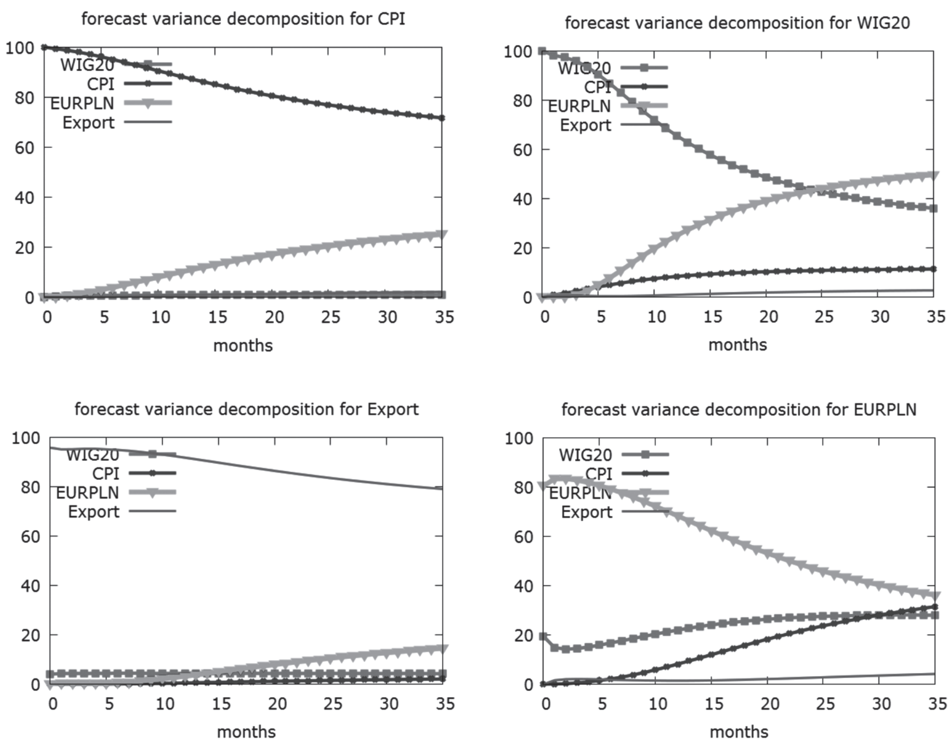

Figure 6. The variance decompositions for WIG20 (left upper panel) and the variance decomposition for Export (right upper panel) for EUR/PLN (left lower panel) and CPI (right lower panel)

Source: authors' own research

Variance decomposition

The forecast error variance decompositions show which shocks are most important in explaining the variance of a given variable through time. They measure the fraction of the forecast error variance of an endogenous variable that can be attributed to orthogonalized shocks to itself or to another endogenous variable.

In Figure 6, the decompositions have been plotted for each variable over a 36-month horizon. The behavior of WIG20 is mainly explained by the shock in the WIG20 value itself (the impact decays from around 100% to 40% after 23 months) and impulses from EUR/PLN in approximately 40% after 20 months, while the rest of its variability is explained by impulses from CPI. In the case of Export, we observe that most of its variability comes from Export itself as well as low impact from EUR/PLN that increases to almost 20% after 30 months.

The variability of the EUR/PLN exchange rate is explained by EUR/PLN shocks in EUR/PLN mainly, but WIG20 and CPI impulses contribute to its variability significantly: each explains almost 20% of the variance. The main drivers for CPI volatility are CPI (around 80%) and EUR/ PLN. The shocks coming from two other variables have a minor impact on CPIs levels.

4. Conclusion

In this paper, we studied the long-term relations between the WIG20 price, EUR/PLN exchange rate, export volume, and rate of inflation CPI. For these purposes, the VEC model was applied. The cointegration analysis has been performed using the Johansen procedure.

The results indicated the existence of one cointegrating relationship between the WIG20 price, EUR/PLN exchange rate, export volume, and rate of inflation CPI. However, the export variable had a neglected impact on the stock market in the long-term relationship. The biggest impact on the stock market had the EUR/PLN exchange rate. The impact of CPI on the stock market is rather moderate. We also concluded that CPI moves away from the equilibrium state with a rather-slow rate, whereas the EUR/PLN comes to the equilibrium state at a similar pace.

Additionally, the impulse response functions obtained from considered VEC model give us a picture of how shock coming from one variable impacts the other variables.

The analysis of the impulse-response function revealed that shocks have a permanent impact on each of the variables considered. In each case, an impulse in one variable causes an immediate increase or decrease in the value of another variable. The effect of shock disappears after some time, and a new level of equilibrium is established. For each variable, the effect of shock decreases after approximately 12 to 15 months.

Taking into account the results obtained from the variance decompositions, we can indicate which shocks are most important in explaining the variability of variables through time. This analysis confirmed the fact that the stock market responds most strongly to the impulse coming from the EUR/PLN rate. However, the behavior of WIG20 is explained mainly by a shock in the WIG20 value itself, while only the rest of its variability is mostly explained by impulses from the EUR/PLN rate.

References

[1] Abdelaziz, M., Chortareas, G. and Cipollini, A. (2008) 'Stock Prices, Exchange Rates, and Oil: Evidence from Middle East Oil-Exporting Countries', Social Science Research Network, vol. 44, pp. 1–27.

[2] Barbic, T. and Condic-Jurkic, I. (2011) 'Relationship between macroeconomic funadamentals and stock market indices in selected CEE countries', Ekonomski Pregled, vol. 62, pp. 113–133.

[3] Bhattacharya, B. and Mukherjee, J. (2006) 'Indian Stock Price Movements and the Macroeconomic Context – A Time-Series Analysis', *Journal of International Business and Economics*, vol. 5, pp. 88–93.

[4] Brahmasrene, T. and Jiranyakul, K. (2007) 'Cointegration and Causality between Stock Index and Macroeconomic Variables in an Emerging Market', *Academy of Accounting and Financial Studies Journal*, vol. 11, pp. 17–30.

[5] Chen, G.M., Firth, M. and Rui, O. M. (2002) 'Stock market linkages: Evidence from Latin America', *Journal of Banking and Finance*, vol. 26, pp. 1113–1141.

[6] Dornbusch, R. and Fischer, S. (1980) 'Exchange rates and current account', *American Economic Review*, vol. 70, pp. 960–971.

[7] Fifield, S.G.M., Power, D.M. and Sinclair, C.D. (2002) 'Macroeconomic Factors and Share Returns: An Analysis using Emerging Market Data', *International Journal of Finance and Economics*, vol. 7, pp. 51–62.

[8] Geske, R. and Roll, R. (1983) 'The fiscal and monetary linkage between stock returns and inflation', *Journal of Finance*, vol. 38, pp. 7–33.

[9] Humpe, A. and Macmillan, P. (2007) 'Can macroeconomic variables explain long term stock market movements? A comparison of the US and Japan', CDMA Working Paper No. 07/20.

[10] Johansen, S. (1988) 'Statistical Analysis of Cointegration Vectors', *Journal of Economic Dynamics and Control*, vol. 12, pp. 231–254.

[11] Johansen, S. (1991) 'Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models', *Econometrica*, vol. 59, pp. 1551–1580.

[12] Kim, K. (2003) 'Dollar exchange rate and stock price: evidence from multivariate cointegration and error correction model', *Review of Financial Economics*, vol. 12, pp. 301–313.
[13] Loef, M. and Franses, P.H. (2000) 'On Forecasting Cointegrated Seasonal Time Series', SSE/EFI Working Paper Series in Economics and Finance, no. 350.
[14] Lutkepohl, H.(1991) Introduction to multiple time series analysis, Berlin and New York: Springer-Verlag.
[15] Lutkepohl, H. (2005) New Introduction to multiple time series analysis, Berlin and New York: Springer-Verlag.
[16] Lutkepohl, H. (2004) Applied Time Series Econometrics, Cambridge: Cambridge University Press.
[17] Mahmood, W.M. and Dinniah, M.N. (2009) 'Stock Returns and Macroeconomic Infl uences: Evidence from the Six Asian-Pacific Countries', International Research, *Journal of Finance and Economics*, vol. 30, pp. 154–164.
[18] Moore, T. (2007) 'The Euro stock markets in Hungary, Poland, and UK', *Journal of Economic Integration*, vol. 22(1), pp. 69–70.
[19] Nieh, C. and Lee, C. (2001) 'Dynamic relationship between stock prices and exchange rates for G-7 countries', *The Quarterly Review of Economics and Finance*, vol. 41, pp. 477–790.
[20] Otsu, T. (2009) Seasonal cycle and filtering, Seijo: Seijo University Economic.
[21] Rjoub, H. (2012) 'Stock prices and exchange rates dynamics: Evidence from emerging markets', *African Journal of Business Management*, vol. 6(13), pp. 4728–4733.
[22] Syriopoulos, T. (2007) 'Dynamic linkages between emerging European and developed stock markets: Has the EMU any impact?', *International Review of Financial Analysis*, vol. 16(1), pp. 41–60.
[23] Voronkova, S. (2004) 'Equity Market Integration in Central European Emerging Markets: A Cointegration Analysis with Shifting Regimes', *International Review of Financial Analysis*, vol. 13(5), pp. 633–647.

Managerial Economics
2016, vol. 17, no. 1, pp. 21–58
http://dx.doi.org/10.7494/manage.2016.17.1.21

Henryk Gurgul*, Łukasz Lach*

Comparative advantage of the EU in global value chains: How important and efficient are new EU members in transition?[1]

1. Introduction

In economic analysis, the measures of competitiveness are expressed in terms of shares of gross output in world export markets. In recent years, these types of measures have become questioned, as one could observe the increasing fragmentation of production across borders. Due to easier communication, increasing information flow, and changes in coordination costs, the various stages of the production process are no longer conducted at geographically close locations. The rising fragmentation of production implies that more and more stages of the production process are faced with international competition. In the past decades, the competitiveness of countries was mainly determined by domestic firms. In most cases, these companies competed 'sector to sector' with similar firms from other countries. The competition was usually taking place in the sphere of the price and quality of traded products.

Fragmentation of production, the process by which different production stages of final goods are conducted in different countries, has been increasing over the last several decades. It is clear that this tendency reflects the globalization process in the world economy. An illustrative example of this phenomenon, often discussed in economic literature, is the case of the German automobile industry, which is a leader in the car world market, at least in terms of traditional competition indexes (Dudenhoeffer, 2005). Since a big share of the intermediate production stages

* AGH University of Science and Technology in Cracow, Faculty of Management, Department of Applications of Mathematics in Economics, e-mails: henryk.gurgul@gmail.com (Henryk Gurgul), llach@zarz.agh.edu.pl (Łukasz Lach).
The authors would like to thank two anonymous referees for the valuable comments on an earlier version of the paper.
[1] Financial support for this paper from the National Science Centre of Poland (Research Grant no. DEC-2015/19/B/HS4/00088) is gratefully acknowledged.

of German cars is conducted in some other countries (mostly in Europe) and, therefore, a considerable part of the value of German cars is produced abroad, one can ask the following question: To what extent are German cars really German? As Marin (2010) stresses, the 'super-competitiveness' of the German economy is in large part derived from the increasing use of imported intermediates. Therefore, the high share of German car exports in the world may not result from the high competitiveness of the German economy. This example clearly shows that, in the time of increasing globalization, new measures of competitiveness are needed. Such measures should be based on the value added in production by a country and on the jobs and capital involved in global production chains. In this paper, we derive such measures empirically from world input-output (IO) tables.

Fragmentation of production is one of the most important sources of structural changes in the CEE economies in transition. This process is likely responsible for the rise (after a dramatic drop in the first half of the 90s) of the share of industrial production in gross domestic product that has been observed in CEE transition economies since the last five years of the Twentieth Century[2].

The EU membership of Poland and other CEE countries, globalization and fragmentation process speeded up the rate of growth of the ratio of investment loans to consumption loans and supported growth rate of manufacturing and construction, which has been reflected in the rise of shares of these sectors in the GDP of Poland and other CEE countries.

The fragmentation of production processes can be principally classified into two main forms. The first one is usually called a "snake" and the second – a 'spider' (Baldwin and Venables, 2013). Under this notion, the 'snake' is understood as a sequence in which intermediate goods are exported from country X to Y. Next, these goods are incorporated into intermediate goods and sent from country Y to country W. This export pattern goes on until the goods reach their final stage of production. In contrary, the 'spiders' comprehend multiple parts coming together from a number of countries. The goal is a single location for assembly of a new component or final product. In the world economy, we usually observe production processes that are complex mixtures of these two types. In this paper, we label the fragmented production processes as 'chains' irrespective of their type. This notion is widely used in the economic literature despite the fact that it rather appeals to the snake-like description of this term.

[2] For example, in Poland during the years 1996–2015, the gross domestic product increased by 105%, industrial production increased by 156%, and output of the processing industry expanded by 260%. In addition, the growth rate of exports in Poland has recently outperformed the rate of growth of GDP, and net exports became positive. At the beginning of transition, the share of employees in agriculture amounted to more than 33%; but in recent years, this index has dropped to under 13%. Moreover, the share of manufacture in the GDP exhibits more than a third, which is higher than the average share in the EU-28.

The most influential framework used in the study of globalization is the so-called **global value chain** (GVC) concept. The calculation of GVC incomes is a very interesting methodology to deal with increasing fragmentation of production, and it surely adds more useful information about competitiveness than the raw export data. However, interpretation of trends in GVC income as trends in competitiveness should be conducted carefully and with a dose of criticism. The reason is that there are very interesting patterns of specialization within different manufacturing of final goods. Results of GVC analysis may be relevant for policy makers and have important policy implications. The application of this methodology with respect to value added, labor productivity, and capital efficiency in the ten new EU members in transition from the CEE region is the main topic of our contribution. We focus on a relative context and try to answer the question of how important and efficient are the CEE transition economies when it comes to building a comparative advantage of the European Union in global value chains. To the best of our knowledge, this paper is the first study dedicated to the issue of answering the latter question by using some original modifications and extensions of methodological developments on an ex-post accounting framework in global value chains presented in the recent IO literature.

The rest of the paper is organized as follows. In Section 2, we conduct a literature review. In the third section, we formulate main research hypotheses. The methodology of studying GVC, the data sources used to measure GVC incomes, as well as indexes of jobs and capital efficiency and discussion on topics that are important for assessing the validity of the empirical results are presented in Section 4. Section 5 is the most important one, as it presents the empirical results along with the respective discussion. The last section provides concluding remarks and some suggestions for future research.

2. Literature review

Early studies on fragmentation were conducted by Fukao et al. (2003) and Ando and Kimura (2005) for Japanese firms, Hanson et al. (2005) for US firms, and Marin (2006) and (2011) for German and Austrian international companies. Macroeconomic evidence has been presented by Hummels et al. (2001) and Johnson and Noguera (2012). They found increasing vertical specialization in trade for most countries. Below, we provide a brief review of the literature that provides direct evidence of fragmentation focusing on the value chains of final products.

In the light of the literature, there are growing discrepancies between growth in gross exports and the generation of incomes and jobs for workers involved in GVC. Sinn (2006) showed that the increasing imports of intermediate goods mainly from Eastern Europe were a source of decline in the German value added per unit of exports. However, export-earned value added as a share of GDP was preserved in Germany. This was possible due to increased specialization and more than proportional growth of exports. In opinion of Sinn (2006), high and rigid wages for unskilled workers were a source of over-specialization on the skill- and capital-intensive segments of the production chain. In the framework of revealed comparative advantage analysis based on gross exports, Di Mauro and Forster (2008) claim that the specialization pattern of the countries from the euro zone was nearly the same between the 1990s and 2000s. In their opinion, the reason for that is the inability of gross export statistics to include the value added in internationally fragmented production. Koopman et al. (2012) analyzed the structure of the export sector of China and found that the latter is mainly based on imported intermediates. The contributors demonstrated that value added in this sector was much lower than suggested by the gross export values. However, the rate of growth of this sector was indeed very high.

It has been found (see e.g., Koopman et al., 2012; Ottaviano et al., 2009) that structures provided by respective columns for particular industries in the Use IO table may only provide average production structures across all firms and all products in those industries and, thus, might be quite different for exporters and non-exporters. However, according to the authors, further evidence on this issue is needed to better understand the nature of the discussed problems.

In order to take into account some issues that are typical in the time of globalization (like the increasing trade in intermediate goods), the World Trade Organization and the Institute of Developing Economies/Japan External Trade Organization have suggested a new methodology of the so-called Trade in Value-Added (TiVA) instead of the common trade in gross terms (Escaith and Inomata, 2011). The WTO and OECD have provided the first empirical results based on international input-output tables. In the economic literature, the global trade network in the framework of TiVA is called a global value chain (GVC). According to GVC, value-added exports from an origin country to a destination country is understood as the origin country's value-added induced by the destination country's final demand (excluding intermediate goods exports) to the rest of the world. The GVC methodology is an extension of the methodology presented in Trefler and Zhu (2010), Bems et al. (2011), and Johnson and Noguera (2012). These papers refer to old contributions on input-output accounting with multiple regions initiated by Isard (1951) and developed in a particular work by Miller (1966). One of the main advantages of the discussed approach is the fact that the data calculated in the framework of the GVC methodology can be used for explaining the causes of structural change. Herrendorf et al. (2013) demonstrated that structural change may be interpreted as caused by non-homothetic demands or by asymmetric productivity growth. A crucial role is played by the type of data analyzed; i.e., data on production or consumption expenditure. In order to conduct GVC analysis, reliable time series data on global input-output tables is required. Such data allows us to map value added sectoral shares (which are the ones that should enter the production functions) into consumption sectoral shares (which are the ones that should enter the utility functions). The key role here is played by the import–export relation in the data, since there are quite different implications of productivity changes in open and closed economies.

The fragmentation of the global production process does not necessarily lead to an increase of unemployment in advanced countries. Grossman and Rossi-Hansberg (2008) found that offshoring may lead to lower output prices and increased demand for the final output. Therefore, the net impact on domestic jobs might be even positive.

In the study by Kuboniwa (2015), which presents an analysis of the role of Russia in GVC, a modified version of the original World Input Output Data (WIOD) was used. The study also uses an alternative definition of value-added trade based on the contribution of Trefler and Zhu (2010). The author proves that this alternative definition is bilaterally equivalent to the traditional definition of TiVA. In the paper, one may also find (rather straightforward) proof of a theorem on the identity between the sum total of a country's value-added trade balances and gross trade balances (net "gross exports" or net exports). In other words, the author claims that the sum total of the differentials between balances in value-added and those in gross terms equals zero. In addition, it is also proven in the paper that a country's total factor content of trade is simply its net exports in conventional terminology. Using several versions of aggregated data taken from WIOD, the contributor supplies evidence supporting the theorems. A modification of the approach, with respect to Russia's trade flows and value-added for sectors related to oil (both crude and refined) and natural gas, are next examined in the empirical part of the paper[3].

Our paper is widely based on the innovative paper by Timmer et al. (2013), which concerns with the World fragmentation process. The authors introduce new indicators of competitiveness that take fragmentation into account. Their method measures the value added in each activity in the process, irrespective of its position in the production network. The contributors stress that concepts like 'global supply chains' or 'international production chains' typically refer to only the physical production stages, whereas the value chain concept used by them refers to a broader set of activities, both in the pre- and post-production phases, including research and development, software, design, branding, finance,

[3] These sectors are among the key ones of both the Russian and EU economies.

logistics, after-sales services, and system integration activities. In other words, the GVC income measures of Timmer et al. (2013) will take into account the value added in all of these stages of production. This methodology seems especially important in an IT branch where (in the case of many products) a major part of value is added (Ali-Yrkko et al., 2011; Dedrick et al., 2010).

Timmer et al. (2013) use recently compiled time series of input-output data of the world (41 countries and regions, 35 industries, time span 1995–2008). The contributors analyze global value chains (GVC) based on the methodology of Trefler and Zhu (2010), Bems et al. (2011), and Johnson and Noguera (2012). In particular, for every pair of a **final good** and a **country of use**, the authors obtain a vector with the value added generated in every possible **sector** in **the countries of origin** and a vector with the corresponding number of workers involved in every possible **sector** in the **countries of origin**. Next, using a traditional static IO approach, they invert and aggregate the respective matrices of data to obtain the total value added and number of workers in each country used to produce the world **final manufactured goods**. They call these indicators **manufactures global value chains** (denoted GVCm) income and jobs. Moreover, the authors keep track of the sectors in each country (including non-manufactures, like services and agriculture) that originate the GVCm income and jobs.

Timmer et al. (2013) focus in particular on the European region as a whole, since it has undergone a strong process of integration in the past two decades (both inside and outside the European Union). Their main findings can be summarized as follows. First of all, they proved that in recent years a strong process of international fragmentation of manufacturing production across Europe has been taking place. This process was the reason for rising differences between changes in gross exports and GVC incomes. In particular, they found that growth in manufactures GVC income during 1995–2008 is essentially lower than the growth in gross manufacturing exports for all European countries (in particular for Austria, Greece, Spain, and Eastern European countries). Moreover, the authors established strong differences in the estimated indexes of comparative advantages of the EU calculated on the basis of the new measures and gross exports. They found that European GVC income is increasing fastest in activities carried out in the production of nonelectrical machinery and transport equipment, and it is declining in activities related to the production of non-durables. In the opinion of the contributors, these findings reflect observed changes in measures of comparative advantage more precisely than the suggestion based on gross export data that implies rather stable patterns of comparative advantage.

The contributors were also surprised that this pattern for both the old and new EU members is somewhat similar to the pattern for Mexico-US integration in the 1990s (Feenstra, 1998; 2010). The authors stress that the manufactures

GVC income of a country estimates the income resulted from activities on the domestic territory related to the production of final manufacturing goods by multinational corporations.

We aim to study the production fragmentation of final products. A final product is consumed while intermediate products continue on in the production process. Total consumption includes both private and public consumption as well as investment. A global value chain of a final product is defined as the value added of all activities that are directly and indirectly required to produce it. This global value chain is identified by the country-industry where the last stage of production is performed before delivery to the final user. The final stage of production in a particular country is not equivalent to the governance of the value chain by a domestic country. For example, large IT corporations from the USA govern the production networks of CEE or India.

In general, the originality of our paper is twofold. First, we focus on a relative context and try to answer the question of how important and efficient the CEE transition economies are when it comes to building a comparative advantage of the European Union in global value chains. To the best of our knowledge, this paper is the first study dedicated to the issue of answering this type of question regarding the new EU members in transition. Second, we propose original modifications and extensions of recently presented methodological developments in ex-post accounting framework in global value chains (e.g., those presented in Timmer et al., 2013), which seems to provide a background for interesting empirical deliberations on both country-specific level as well as for the whole group of CEE economies analyzed.

3. Main research conjectures

In general, each research hypothesis examined in this paper consists of three major components. First, in each hypothesis, we refer to one of the three relative measures discussed in this study. These are the relative value added in the group of CEE economies in transition and the measures of relative productivity and capital efficiency. Second, we try to express (and verify) our suppositions on the dynamics of these indexes over the period under study. Finally, we try to extend the general statements on the whole group of examined CEE transition economies by focusing on country- and sector-specific results.

In the economic literature, it is often stressed that one of the major features of the economies of CEE in the 90s was the process of **de-industrialization**, which lead to heavy losses in the secondary sector of the economy (Kalvet and Kattel, 2006). As a consequence, the economic activity in these countries was partly redirected to other sectors (especially services). The latter was accompanied with the ongoing process of globalization, especially in the sphere of economic openness and information flows; this had a significant positive causal impact on economic growth in new EU member countries from the CEE region (Gurgul and Lach, 2014). Since we are interested in the analysis of value added in CEE transition economies in the relative context with respect to the EU in total, we should also underline that, in recent years, the level of technological progress in new EU members has not reached the average EU level as of yet[4]. This implies that these economies are still not playing a key role in high-technology sectors but rather focus on usage of their natural resources/conditions and aim at specializing in rather low-technology production; e.g., in the sector of agriculture, food production, products of wood, etc. Moreover, the energy sectors of the largest EU members from the CEE (e.g. Poland) are still heavily dependent on coal and lignite, while the richest EU counties have already implemented many programs aimed at moving toward alternative sources of energy; e.g., renewable sources or nuclear energy.

On the other hand, during the period of transition, some sectors of CEE economies have grown dramatically (especially after EU accession). A good example here would be the sector of tourism. After the collapse of the Iron Curtain, the societies of CEE transition economies gained the possibility of traveling freely abroad. At the same time, the tourist offerings of the region have been continuously expanding.

All of the above-mentioned observations may have a significant impact on the levels of sector-specific value added in new EU member countries measured as a share of total value added by all EU countries in the framework of GVC. Taking into account these remarks (along with some basic characteristics of the new EU member countries from CEE), one may expect the following hypothesis to hold true:

Hypothesis 1: *The role of selected CEE economies in transition in the creation of value added with respect to the total value added in the European Union in the GVC framework was biggest in the case of the following sectors: Agriculture, Hunting, Forestry and Fishing, Mining and Quarrying, Wood and Products of Wood and Cork, Other Non-Metallic Mineral, and Other Supporting and Auxiliary Transport Activities; Activities of Travel Agencies. When it comes to the country-specific results, one may expect that, among new EU members in transition, the highest shares in the EU total GVC value added were obtained in the largest countries, e.g. Poland and Romania*[5].

[4] For example, the share of R&D expenditure in GDP in new EU members in transition is still much lower as compared to old EU countries.

[5] In all of the research hypotheses listed in Section 3, we use the WIOD names of the respective sectors.

Beside analyzing the relative sector-specific shares in the value added in the group of the ten CEE economies in this paper, we propose new GVC-embedded measures of productivity and capital efficiency and focus on an analysis of their levels and growth rates in two sub-periods: before and after EU expansion[6]. At this place one should underline that, contrary to the values of shares in the EU total GVC value added, the growth rates of productivity are expected to be higher in smaller CEE transition economies. The latter follows from the fact that the smaller the size of the country, the larger the gain from trade. A small country can successfully export the surplus production to a large country (with large market capacity), and it can take advantage of foreign trade (comp. the theory of comparative costs). Thus, small countries are usually encouraged to specialize (or even forced to, due to a scarcity of primary products). In contrary, large countries typically do not suffer from a scarcity of primary products. However, they could have problems with respect to export of (large) excess output to smaller markets (with small market capacity). Therefore, they are not forced (or even encouraged) to specialize. Taking into account these observations (along with the history of structural change in CEE transition economies in the past 25 years), the compositions of key sectors in CEE transition economies and West European countries as well as the differences between labor markets in both groups of EU countries, one may formulate the following conjecture:

> **Hypothesis 2:** *After the two decades of transition, productivity in the ten CEE economies was still much lower as compared to the EU average for most of the sectors. However, during this period, the growth rates of the measures of productivity were, in general, positive. Moreover, they increased at a higher rate after EU accession. The highest levels and growth rates of productivity were found for smaller CEE countries*.

Before the beginning of the transition, the group of CEE countries strongly relied on fixed capital, which was an important input in manufacturing. One of the consequences of economic transition was de-industrialization in CEE along with the following shift towards new types of economic activities (especially services). The latter was also accompanied with the process of privatization. Taken altogether, one may expect that the levels of capital efficiency in CEE transition economies with respect to the EU average will be much higher that the corresponding indexes of relative productivity. Moreover, one could expect that, after EU accession, the role of capital input in the process of production and creation

[6] We focus on EU expansion which took place in 2004.

of value added in the new EU members in transition could be even smaller, due to the shift towards less capital-intensive activities. These general remarks suggest the formulation of our final hypothesis:

> **Hypothesis 3:** *After the two decades of transition, capital efficiency in the ten CEE economies in 2009 was comparable to the EU average for most of the sectors. Moreover, during this period, the growth rates of the measures of the capital efficiency were, in general, positive. However, the growth rates of these indexes dropped after EU accession. The highest levels and growth rates of indexes of capital efficiency were found for smaller CEE countries.*

The hypotheses listed above will be verified using the methodology and dataset presented in Section 4. In the next section, we will briefly present the main empirical findings of the paper.

4. Dataset and research methodology

In order to analyze the dynamics of value added shares as well as the indexes of productivity and capital efficiency in the ten CEE transition economies, we will use the global value chain (GVC) approach introduced in Timmer et al. (2013). In order to provide an ex-post accounting of the value of final demand, we will trace the value added at the various stages of production in an international input-output model. First, we will briefly introduce the accounting framework drawing on the exposition in Johnson and Noguera (2012) and Timmer et al. (2013); then, we will generalize and extend their approach to analyze the value added by specific production factors in the case of the group of new EU members in transition.

Let us now shed some light on the GVC approach. Henceforth, we assume that there are C countries, S sectors, and F production factors. Under the term **country-sector**, we shall understand one specific sector operating in one selected economy. Each of these country-sectors produces one good; thus, there are SC products. It is clear that the output in each country-sector is produced by using both domestic production factors as well as intermediate inputs (which, in turn, may be provided by domestic or foreign suppliers. As usual in an input-output framework, we assume that output in each sector is either used to satisfy final demand (either at home or abroad) or serves the role of intermediate input in production processes in other sectors. Final demand consists of consumption expenditure by households, non-profit organisations serving households, and government as well as gross fixed capital formation (investment) and changes in inventories and valuables. Following Timmer et al. (2013), we use the simplifying notation for each product, with *i* denoting the source country, *j* denoting the destination country, *s* standing for the source sector, and *r* denoting the destination sector. We assume that product market clearing takes place; thus, the quantity of a good produced in a particular country-sector must equal the quantities of this product used domestically and abroad. For every year *t*, the product market clearing condition takes the following form:

$$y_i^t(s) = \sum_j f_{ij}^t(s) + \sum_j \sum_r m_{ij}^t(s,r) \tag{1}$$

Where $y_i^t(s)$ stands for the value of output in sector *s* of country *i* in year *t*, $f_{ij}^t(s)$ denotes the value of goods shipped from this sector for final use in country *j* in year *t*, and $m_{ij}^t(s, r)$ stands for the value of goods shipped from this sector for intermediate use by sector *r* in country *j* in year *t*[7].

Using simple matrix algebra, the market clearing conditions (1) for each of the *SC* goods can be combined into a compact global input-output system. In order to obtain this compact form, let Y_t denote the $SC \times 1$ vector of production in year *t*, which is obtained by row-wise concatenation of output levels (each in the form of a $S \times 1$ vector) in each country-sector:

$$Y_t = \begin{bmatrix} Y_1^t \\ Y_2^t \\ \vdots \\ Y_C^t \end{bmatrix}, \quad Y_i^t = [y_i^t(s)]_{s=1,\dots,S}, \quad i = 1,\dots,C \tag{2}$$

Analogously, we may define the $SC \times 1$ vector of global final demand (denoted as F_t) by stacking world final demand for output from each country-sector $f_i^t(s)$. The latter takes the form of a summation of demand for a product of the sector *s* from any country, i.e.,

$$f_i^t(s) = \sum_j f_{ij}^t(s), \quad i = 1,\dots,C, \quad s = 1,\dots,S \tag{3}$$

Using this notation, we may write:

$$F_t = \begin{bmatrix} F_1^t \\ F_2^t \\ \vdots \\ F_C^t \end{bmatrix}, \quad F_i^t = [f_i^t(s)]_{s=1,\dots,S}, \quad i = 1,\dots,C \tag{4}$$

[7] Note that the use of goods can be at home (in the case of $i = j$) or abroad ($i \neq j$).

For each year t, we define a $SC \times SC$ global intermediate input coefficient matrix $A_t = [a^t_{wz}]_{w,z=1,\ldots,SC}$ using the following formula:

$$a^t_{wz} = \frac{m^t_{ij}(s,r)}{y^t_i(r)}, \text{ for } w = s + (i-1)S,\ z = r + (j-1)S \tag{5}$$

where $i, j = 1, \ldots, C,\ s, r = 1, \ldots, S$. The elements a^t_{wz} represent the output from sector s in country i used as intermediate input by sector r in country j as a share of output in the latter sector in year t (Timmer et al., 2013). Using matrix A_t, we may now answer the question of which combination of various intermediate products (both domestic and foreign) are required to produce one unit of each country-sector product. Using this definition, we can now rewrite the global market clearing conditions (1) in a compact IO-based form:

$$Y_t = A_t Y_t + F_t \tag{6}$$

or equivalently:

$$Y_t = (I - A_t)^{-1} F_t \tag{7}$$

where I is an $SC \times SC$ identity matrix.

Using the approach of Timmer et al. (2013), one may attribute the value of final demand for a specific product to value added in all country-sectors that directly and indirectly participate in the production process of the final good. Throughout this paper, we define value added in a traditional way; namely, as the difference between gross output value (at basic prices) and the cost of intermediate goods and services (at purchaser's prices). For each sector s and country c, we define $p^t_c(s)$ as the value added per unit of gross output produced in year t, and create the stacked $SC \times 1$-vector p^t_{VA} containing these (direct) value added coefficients:

$$p^t_{VA} = \begin{bmatrix} p^t_{VA}(1) \\ p^t_{VA}(2) \\ \vdots \\ p^t_{VA}(C) \end{bmatrix}, \quad \text{where} \quad p^t_{VA}(c) = \begin{bmatrix} p^t_c(1) \\ p^t_c(2) \\ \vdots \\ p^t_c(S) \end{bmatrix} \tag{8}$$

In order to take 'indirect' contributions into account, we derive the $SC \times 1$-vector of value added levels v^t_{VA} as generated to produce a final demand vector F_t. To get this vector, the gross outputs needed for production of this final demand should be multiplied by the elements of the direct value added coefficient vector p^t_{VA}. After multiplying (7) by $diag(p^t_{VA})$, one gets[8]:

[8] Henceforth, for a given vector $[x_j]_{j=1,\ldots,n}$ the symbol diag(x_j) denotes the $n \times n$ diagonal matrix with elements x_j on the diagonal.

$$v_{VA}^{t} = diag(p_{VA}^{t})(I - A_{t})^{-1}F_{t} \tag{9}$$

where:

$$v_{VA}^{t} = \begin{bmatrix} v_{VA,1}^{t} \\ v_{VA,2}^{t} \\ \vdots \\ v_{VA,C}^{t} \end{bmatrix} \tag{10}$$

Using model (9), one can now multiply matrix $diag(p_{VA}^{t})(I - A_{t})^{-1}$ with any vector of final demand levels F_{0}^{t} to find out which value added levels v_{VA}^{t} should be attributed to this particular set of final demand levels in all C countries[9].

The GVC-embedded methodology outlined so far are based on suggestions of traditional IO literature and has already been applied in empirical research (see e.g., Timmer et al., 2013). Below, we will briefly present a modification of this approach aimed at analyzing dynamics of value added, productivity, and capital efficiency in ten new EU members in transition. Our goal is to establish a GVC-embedded research framework to examine the dynamics of the discussed variables in ten CEE economies in a relative context with respect to the EU total.

For each sector s^* and year t, let us define the $SC \times 1$ global final demand vector using the following formula:

$$F_{t}(s^{*}) = \begin{bmatrix} F_{1}^{t}(s^{*}) \\ F_{2}^{t}(s^{*}) \\ \vdots \\ F_{C}^{t}(s^{*}) \end{bmatrix}, \quad F_{i}^{t}(s^{*}) = f_{i}^{t}(s^{*})e(s^{*}) \tag{11}$$

where $e(s^{*}) = [e_{q}]_{q=1,\ldots,S}$ is an $S \times 1$ vector defined as:

$$e_{q} = \begin{cases} 1, & q = s^{*} \\ 0, & otherwise \end{cases} \tag{12}$$

[9] Although the ex-post accounting framework described above does not allow us to explicitly deal with the interaction of prices and quantities as in a full-fledged Computable General Equilibrium model, it is also free of bias resulting from the need of econometric estimation of various key parameters of production and demand functions. The latter, along with characteristics of the annual IO data, makes the approach particularly well-suited for a wide range of ex-post analysis (Timmer et al., 2013).

In other words, the vector $F_t(s^*)$ contains the C values of final demand for sector s^* across all the countries. For remaining sectors j $(j \neq s^*)$, the corresponding entries in vector $F_t(s^*)$ are all equal to zero. Using formulas (8)–(10), we may now establish the distribution of value added to be attributed to the set of final demand levels contained in vector $F_t(s^*)$:

$$v_{VA}^t(s^*) = diag\left(p_{VA}^t\right)(I - A_t)^{-1} F_t(s^*) \tag{13}$$

We will now focus on the construction of GVC-embedded indexes measuring the role of selected CEE economies in transition in the creation of value added with respect to the total production in the European Union. We are also interested in analyzing the dynamics of the indexes of productivity and capital efficiency in the new EU member countries in transition. Let J_{EU27} and J_{CEE} denote the sets of indexes for EU27[10] and the ten new EU members from the CEE region[11]. For each sector s^* and year t, we may calculate the value added in GVC in the whole EU27 and in the group of ten CEE economies using equation (13). Next, we may define the **Value Added Share (VAS)** attributed in the new EU members with respect to the EU27 total using the following formula:

$$VAS_{CEE|EU27}(s^*, t) = \frac{\sum_{j \in J_{CEE}} v_{VA,j}^t(s^*)}{\sum_{j \in J_{EU27}} v_{VA,j}^t(s^*)} \times 100\% \tag{14}$$

For example, we may use formula (14) to assess the size of value added attributed to the global final demand in agriculture in the ten CEE economies in transition, with respect to the value added in this sector in the whole European Union[12].

One can generalize the decomposition of the value of final demand outlined above to analyze the value and quantities used of specific production factors. In this paper, we focus on two basic types of inputs: labor and capital. We define $p_{L,c}^t(s)$ as the direct labor input per unit of gross output produced in sector s in

[10] Due to a lack of required data on Croatia throughout this paper, we do not focus on the EU28.

[11] The ten new EU members are as follows (respective abbreviations are given in brackets): Bulgaria (BGR), Czech Republic (CZE), Estonia (EST), Hungary (HUN), Latvia (LVA), Lithuania (LTU), Poland (POL), Romania (ROM), Slovakia (SVK), and Slovenia (SVN). It should be mentioned that, during the period of 2004–2010, 12 countries joined the EU. However, Malta and Cyprus have not been taken into consideration in this study because they have never been in transition.

[12] It is clear that $VAS_{CEE|EU27}(s^*,t) \in [0,100\%]$ for all sector and all years. For example, if $VAS_{CEE|EU27}(s^*,t) = 10\%$, this implies that 10% of GVC value added in the sector s^* in year t in the whole European Union was due to the economic activity carried out in the ten new member countries from the CEE region.

country *c* in year *t*. Using an analogous formula to equation (8), we can now obtain a stacked $SC \times 1$ vector p_L^t containing these (direct) coefficients. Analogously to the case of analysis of value added, the elements of vector p_L^t do not account for labor embodied in the intermediate inputs used. However, analogous to formula (13), for each sector s^* and each year *t*, we may derive all direct and indirect labor inputs needed for the production of a specific final product:

$$v_L^t(s^*) = diag\left(p_L^t\right)(I - A_t)^{-1} F_t(s^*) \tag{15}$$

Using this formula, we may now simply calculate the **ratio of labor (labor share, denoted *LS*)** used in the global production processes in sector s^* in the group of ten CEE economies, with respect to the labor used in the whole EU27:

$$LS_{CEE|EU27}(s^*, t) = \frac{\sum_{j \in J_{CEE}} v_{L,j}^t(s^*)}{\sum_{j \in J_{EU27}} v_{L,j}^t(s^*)} \times 100\% \tag{16}$$

Using analogous formulas to (15)–(16), we may define respective indicators for the second input considered – capital:

$$v_C^t(s^*) = diag\left(p_C^t\right)(I - A_t)^{-1} F_t(s^*) \tag{17}$$

Using this formula, we may now simply calculate the **ratio of capital (capital share, denoted *CS*)** used in the global production processes in sector s^* in the group of ten CEE economies with respect to the labor used in the whole EU27:

$$CS_{CEE|EU27}(s^*, t) = \frac{\sum_{j \in J_{CEE}} v_{C,j}^t(s^*)}{\sum_{j \in J_{EU27}} v_{C,j}^t(s^*)} \times 100\% \tag{18}$$

The labor ($LS_{CEE|EU27}(s^*,t)$) and capital ($CS_{CEE|EU27}(s^*,t)$) ratios defined above may serve as a basis to define the following GVC-embedded index of workforce productivity:

$$PRODUCTIVITY_{CEE|EU27}(s^*, t) = \frac{VAS_{CEE|EU27}(s^*, t)}{LS_{CEE|EU27}(s^*, t)} \times 100\% \tag{19}$$

and the GVC-embedded index of capital efficiency:

$$CAP\_EFF_{CEE|EU27}(s^*,t)=\frac{VAS_{CEE|EU27}(s^*,t)}{CS_{CEE|EU27}(s^*,t)}\times 100\% \tag{20}$$

The interpretation of both indexes defined above is rather straightforward. If, for example, $PRODUCTIVITY_{CEE|EU27}(s^*,t)$ = 50%, this implies that, in the GVC framework, the labor used to create one unit of value added in sector s^* in year t in the ten CEE economies in transition was twice as high as the labor used in all of the EU27 countries.

To measure share in value added and calculate the indexes of productivity and capital efficiency for each of the EU27 countries, we need to track country gross output and value added by industry, the global input-output matrix, final goods shipments over time, as well as labor and capital inputs. This type of data is available from the recently released World Input-Output Database (WIOD). In this paper, we use the most-recent world IO tables published by the WIOD, which cover the period 1995–2009 and provide data on 41 regions of the world (40 major countries and the aggregate data on the rest of the world). Thus, the respective IO matrices (vectors) are 1435 × 1435 (1435 × 1) in size[13]. For each sector $s^* \in \{1,35\}$ and year $t \in \{1995,2009\}$, the vector of final output Y_t, input coefficient matrix A_t, sector-specific final demand $F_i^t(s^*)$, and vector of direct value added coefficient $\hat{p}_{VA}^t$ are derived directly from the WIOD database[14]. We used the WIOD data on hours worked to proxy the direct labor input per unit of gross output (i.e., vector $\hat{p}_L^t$) and the WIOD data on gross fixed capital formation as approximation of direct capital input per unit of gross output (vector $\hat{p}_C^t$)[15].

5. Empirical results

In order to analyze the role of the new EU members in transition in the creation of value added in the European Union within the global value chain (GVC)

[13] WIOD consists of a series of detailed and reliable databases and covers 27 EU countries and 13 other major countries in the world. For more details on the WIOD database, see Timmer (2012).

[14] If e stands for a $SC \times 1$ summation vector containing ones, one may write $(\hat{p}_{VA}^t)' = (e)'(I-A)$. This implies that, in (9), the elements of the vector of distribution of the value of final output as attributed to sectors in the value chain of any product (v_{VA}^t) add up to the elements of the vector of final demand $F_i^t(s^*)$ (for a short proof, see Timmer et al., 2013).

[15] We used exchange rates published by the World Bank to transform the original WIOD data on gross fixed capital formation expressed in national currencies into comparable values expressed in US dollars.

framework, one should first take a look at data on the size and dynamics of global final demand for all of the products examined. For this purpose, we used the WIOD input-output tables expressed in the previous years' prices (covering the period 1996–2010). Next, we used the chain rule to obtain the final demand levels in 2009 expressed in 1995 US dollars. Finally, we calculated real growth rates for all 35 sectors. Table 1 presents the respective results[16].

There are some similarities when it comes to the list of the sectors with the biggest and lowest values of final demand in EU27, EU15[17], the group of ten CEE economies in transition, and the world total. The sectors of **Construction**, **Real Estate Activities**, **Public Admin and Defense**, **Compulsory Social Security**, and **Health and Social Work** were listed among the sectors with the largest levels of final demand, while the sectors of **Wood and Products of Wood and Cork** and **Water Transport** were characterized with the smallest levels of final demand (no matter the group of economies examined). Except for the **Water Transport** sector, all remaining sectors experienced more than 60% of real growth of final demand in the group of ten CEE economies under study. Moreover, in the case of some sectors, the final demand levels in the ten CEE economies have risen by more than six times during the period of 1995–2009. More attention should be given to the sector of **Mining and Quarrying** in the case of which one could see a significant rise in global final demand (in particular, a 60% rise in the case of the ten CEE economies). At the same time, the final demand for the products of this sector has fallen in the remaining group of EU countries, especially the old EU members (EU15).

In the next stage, we used equation (14) to analyze the role of new EU members in transition (as a group of ten countries and individually for each country) in the creation of value added in the European Union within the global value chain (GVC) framework. The results of this analysis are presented in Table 2[18].

---

[16] In addition to the EU-related data in Table 1, we also provide data on the discussed trade statistics for all countries in the world (see the column named World). The latter illustrates the evolution of GVC value added from a global perspective.

[17] EU15 consist of the 15 so-called **old members** of the EU.

[18] The goal of this paper is to analyze the GVC value added in the ten CEE economies expressed as a share of GVC value added in the EU27. However, one may be interested in understanding the process of catching-up of the CEE towards the EU15 as well as the absolute and relative convergence of the CEE towards the EU15. In such a case, one would be interested in taking the EU15 as a benchmark (rather than the EU27, as the CEE economies belong to the EU27). Fortunately, the transformation of the shares of relative GVC value added in CEE economies into the shares of GVC value added in CEE economies with respect to GVC EU15 is quite simple. If, for a particular sector, one denotes the GVC value added in the respective groups of countries as CEE_VA, EU27_VA, and EU15_VA, and if one defines VAS1 = CEE_VA/EU27_VA then after a simple algebraic reformulation, one may get CEE_VA/EU15_VA = VAS1/(1 – VAS1), since EU27_VA = EU15_VA + CEE_VA.

One can list several sectors in the case of which the GVC value added in 2009 in CEE transition economies (measured as a share of the EU27 GVC total value added) was largest. These were **Agriculture**, **Hunting**, **Forestry and Fishing**, **Mining and Quarrying**, **Wood and Products of Wood and Cork**, **Other Non-Metallic Mineral**, and **Other Supporting and Auxiliary Transport Activities**, **Activities of Travel Agencies**. In other words, in the case of these sectors, the group of ten new EU members in transition had the greatest effect on the value added in the whole EU in global production processes. With just a few exceptions, the highest share of value added among the ten CEE countries was usually reported for Poland[19].

When it comes to the analysis of the dynamics of the measures of value added in the examined CEE economies, one should underline that, during the period of 1995–2009, the sectors of **Mining and Quarrying**, **Wood and Products of Wood and Cork**, and **Other Non-Metallic Mineral** experienced the highest rise in the share of the EU27 GVC value added[20]. In general, the share of the GVC value added in the sector of **Mining and Quarrying** in ten CEE economies in the EU27 GVC value added rose during the transition period. Moreover, the highest growth rate was observed especially after EU accession. This result, however, should be interpreted together with the outcomes presented in Table 1; namely, the shrinking demand for coal observed in the EU15 countries accompanied with a shift towards low-carbon energy sources evident in the richest European economies. Taken altogether, these results provide solid support for the Hypothesis 1.

In the next stage, for each sector $s^* \in \{1{,}35\}$ and each year $t \in \{1995{,}2009\}$, we calculated the indexes of productivity and capital efficiency for the group of ten CEE countries examined (using formulas (19) and (20)). The results are presented in Tables 3 and 4.

In the case of most of the sectors, the index of productivity in the ten CEE economies in 2009 was still below 50%. This implies that, in general, the CEE economies in transition were still using approximately twice as much labor to create one unit of value added as compared to the EU27 average. The highest levels of the index of productivity (around 70–80%) were obtained for the sectors of **Other Supporting and Auxiliary Transport Activities**; **Activities of Travel Agencies** and **Sale**, and **Maintenance and Repair of Motor Vehicles and Motorcycles**; **Retail Sale of Fuel**.

---

[19] Using the formula that allows changing the benchmark from EU27 to EU15 (see footnote 18), one may easily see that the ordering of sectors according to the relative GVC value added in the ten CEE economies in transition is exactly identical (no matter if the EU27 or EU15 was taken as a benchmark).

[20] In the case of the sectors of **Wood and Products of Wood and Cork** and **Other Non-Metallic Mineral**, the share in the EU27 GVC value added in 2009 was more than twice as high as in 1995.

Beside the two mentioned sectors, the largest positive changes of the index of productivity during the period of 1995–2009 were obtained for the sector of **Mining and Quarrying**. Moreover, the growth rates of these indexes increased after EU accession. When it comes to the data on individual counties, it is worth underling that the index of productivity was highest most frequently for Slovenia. In other words, these results provide support for Hypothesis 2.

Quite different results were obtained for the index of capital efficiency. In the case of most of the sectors, this index in 2009 was close to (or higher than) 100% in the case of the group of ten CEE economies in transition. In general, this implies that, in order to create one unit of value added, the CEE economies were using similar amounts of capital input as the average values of this input in the whole EU27. The highest levels of the index of capital efficiency (around 110–120%) were obtained for the sectors of **Real Estate Activities, Mining and Quarrying, Renting of M&Eq and Other Business Activities**, and **Water Transport**.

The largest positive changes of the index of capital efficiency during the period of 1995–2009 were obtained for the sectors of **Electricity, Gas and Water Supply**, **Post and Telecommunications**, **Financial Intermediation**, and **Renting of M&Eq and Other Business Activities**. In contrary to the results presented in Table 4, the growth rates of these indexes dropped after EU accession[21]. These results, in turn, provide support for Hypothesis 1.

In addition to the results presented in Tables 1–4 (which were dedicated to the whole group of ten CEE economies), we conducted a detailed country- and sector-specific individual analysis of the shares in value added and indexes of productivity and capital efficiency for all ten CEE economies. Since this analysis was carried out for individual countries, the results significantly extend the group-overall information presented in Tables 1–4. We focused on the top five sectors with respect to the shares in the GVC EU27 value added. The list of the top value-added sectors is as follows: **Agriculture**, **Hunting, Forestry and Fishing**, **Mining and Quarrying, Wood and Products of Wood and Cork, Other Non-Metallic Mineral, and Other Supporting and Auxiliary Transport Activities**, **Activities of Travel Agencies**. The respective results are presented in the Figures 1–5.

It is worth mentioning that, in the case of the sector of **Mining and Quarrying**, one could notice extremely small or even negative values of VAS in the case of some CEE transition economies. The latter implies that, in the global-market framework, some CEE countries have experienced a situation when the final demand for products of the sector **Mining and Quarrying** was equal to or even smaller than the intermediate consumption. The latter is evidence of inefficiency of this sector in the ten new EU members in transition in the global framework, as it did not make enough profit to cover the cost of doing business. In general, the detailed results presented in Figures 1–5 support Hypotheses 1–3.

[21] Except for the sector of **Electricity, Gas, and Water Supply**.

Table 1

Levels and real growth rates of final demand in the 35 sectors[a]

Sector Name	Final output in 2009 (1995 USD)				Real growth in final output in 1995–2009 (1995 prices) (in %)			
	WORLD	EU27	EU15	CEE	WORLD	EU27	EU15	CEE
Agriculture, Hunting, Forestry and Fishing	1089.0	266.4	220.8	44.9	91	62	56	107
Mining and Quarrying	188.4	11.0	8.4	2.7	209	–24	–34	60
Food, Beverages and Tobacco	2554.5	820.3	721.4	97.5	76	58	50	169
Textiles and Textile Products	581.3	158.1	138.3	19.8	35	4	–2	90
Leather, Leather and Footwear	139.1	43.8	38.8	5.0	56	11	6	89
Wood and Products of Wood and Cork	43.8	23.9	20.7	3.2	–20	32	21	226
Pulp, Paper, Paper, Printing and Publishing	348.5	155.3	143.8	11.2	46	36	31	209
Coke, Refined Petroleum and Nuclear Fuel	728.6	250.6	228.5	22.1	288	242	234	359
Chemicals and Chemical Products	754.6	335.0	312.9	21.8	107	105	98	301
Rubber and Plastics	150.5	57.3	50.5	6.6	54	41	29	411
Other Non-Metallic Mineral	66.1	37.1	31.4	5.5	15	37	23	246
Basic Metals and Fabricated Metal	334.6	133.2	118.6	14.4	43	37	27	308
Machinery, Nec	1185.5	450.3	417.1	33.2	63	67	59	365
Electrical and Optical Equipment	1381.1	337.5	292.1	45.1	76	47	31	568
Transport Equipment	1711.8	615.7	553.2	62.3	69	70	57	553
Manufacturing, Nec; Recycling	461.9	178.0	155.2	22.4	62	39	27	314
Electricity, Gas and Water Supply	788.5	350.6	310.5	39.5	85	79	67	335
Construction	5823.7	1614.4	1450.7	160.0	98	89	77	345

Table 1 cont.

Sale, Maintenance and Repair of Motor Vehicles and Motorcycles; Retail Sale of Fuel	558.8	296.9	276.4	20.1	69	97	90	282
Wholesale Trade and Commission Trade, Except of Motor Vehicles and Motorcycles	2048.2	677.2	618.5	57.4	85	89	80	308
Retail Trade, Except of Motor Vehicles and Motorcycles; Repair of Household Goods	2378.3	632.0	575.1	55.3	62	89	82	208
Hotels and Restaurants	1936.3	773.0	740.4	29.7	89	112	107	330
Inland Transport	842.8	267.3	232.2	34.9	83	69	54	355
Water Transport	113.6	42.6	41.8	0.8	173	183	193	–6
Air Transport	233.4	72.9	69.8	2.7	77	98	94	281
Other Supporting and Auxiliary Transport Activities; Activities of Travel Agencies	260.8	140.8	123.4	16.9	102	106	95	241
Post and Telecommunications	829.8	259.2	234.7	23.9	168	144	128	645
Financial Intermediation	2097.6	648.9	610.0	37.7	137	183	173	659
Real Estate Activities	4757.6	1649.1	1534.6	110.6	101	109	100	408
Renting of M&Eq and Other Business Activities	1608.1	534.6	502.6	31.6	157	134	127	359
Public Admin and Defence; Compulsory Social Security	6110.2	1454.1	1348.8	101.5	112	100	93	281
Education	2051.6	908.5	840.8	65.8	102	90	83	310
Health and Social Work	4436.7	1698.5	1623.7	72.8	125	134	129	357
Other Community, Social and Personal Services	1970.7	744.5	683.8	58.2	93	100	91	368
Private Households with Employed Persons	96.5	70.8	67.7	2.9	114	126	122	240

Source: own elaborations based on WIOD world IO tables in previous years' prices.
[a] Dark (grey) shading indicates 5 largest (smallest) values.

Table 2

Share in value added attributed in the new EU members with respect to the EU27 total[a]

Sector Name	Value added in CEE countries (share in EU27) (in %)				CEE country with highest value added (share in EU27)			
	2009	1995–2009 Change	Average annual change 2000–2004	Average annual change 2004–2009	1995	2000	2005	2009
Agriculture, Hunting, Forestry and Fishing	13.97	2.13	0.19	0.06	POL	POL	POL	POL
Mining and Quarrying	17.65	7.81	0.22	1.41	POL	POL	POL	POL
Food, Beverages and Tobacco	10.96	4.31	0.26	0.44	POL	POL	POL	POL
Textiles and Textile Products	9.57	3.65	0.24	0.31	POL	POL	POL	POL
Leather, Leather and Footwear	8.30	2.62	0.11	0.38	POL	POL	ROU	POL
Wood and Products of Wood and Cork	13.31	7.27	0.33	0.99	POL	POL	POL	POL
Pulp, Paper, Paper, Printing and Publishing	6.69	3.60	0.17	0.47	POL	POL	POL	POL
Coke, Refined Petroleum and Nuclear Fuel	10.40	4.88	0.30	0.47	POL	POL	POL	POL
Chemicals and Chemical Products	5.69	2.41	0.14	0.26	POL	POL	POL	POL
Rubber and Plastics	9.25	6.18	0.29	0.81	POL	POL	POL	POL
Other Non-Metallic Mineral	12.67	7.24	0.54	0.45	POL	POL	POL	POL
Basic Metals and Fabricated Metal	10.01	5.90	0.31	0.70	POL	POL	POL	POL
Machinery, Nec	7.35	4.30	0.25	0.44	POL	POL	POL	POL
Electrical and Optical Equipment	9.30	6.31	0.41	0.56	POL	POL	HUN	POL

Table 2 cont.

Transport Equipment	9.35	6.41	0.36	0.70	POL	POL	POL	POL
Manufacturing, Nec; Recycling	10.61	6.06	0.40	0.52	POL	POL	POL	POL
Electricity, Gas and Water Supply	11.17	6.92	0.43	0.66	POL	POL	POL	POL
Construction	8.44	4.34	0.20	0.58	POL	POL	POL	POL
Sale, Maintenance and Repair of Motor Vehicles and Motorcycles; Retail Sale of Fuel	6.34	2.94	0.17	0.31	POL	POL	POL	POL
Wholesale Trade and Commission Trade, Except of Motor Vehicles and Motorcycles	7.93	4.12	0.29	0.31	POL	POL	POL	POL
Retail Trade, Except of Motor Vehicles and Motorcycles; Repair of Household Goods	7.97	2.84	0.14	0.37	POL	POL	POL	POL
Hotels and Restaurants	3.85	1.79	0.11	0.18	POL	POL	POL	POL
Inland Transport	11.19	6.62	0.52	0.35	POL	POL	POL	POL
Water Transport	3.26	–1.81	–0.24	0.14	POL	POL	POL	POL
Air Transport	4.36	2.19	0.14	0.19	POL	POL	POL	POL
Other Supporting and Auxiliary Transport Activities; Activities of Travel Agencies	11.21	4.27	0.15	0.69	CZE	CZE	CZE	CZE
Post and Telecommunications	8.68	5.76	0.40	0.45	POL	POL	POL	POL
Financial Intermediation	5.13	2.86	0.19	0.24	POL	POL	POL	POL
Real Estate Activities	5.77	3.04	0.20	0.25	POL	POL	POL	POL
Renting of M&Eq and Other Business Activities	5.15	2.31	0.12	0.27	POL	POL	POL	POL

Table 2 cont.

Sector Name	Value added in CEE countries (share in EU27) (in %)				CEE country with highest value added (share in EU27)			
	2009	1995–2009 Change	Average annual change 2000–2004	Average annual change 2004–2009	1995	2000	2005	2009
Public Admin and Defence; Compulsory Social Security	6.54	2.87	0.19	0.24	POL	POL	POL	POL
Education	6.56	3.24	0.24	0.21	POL	POL	POL	POL
Health and Social Work	3.98	1.76	0.10	0.19	POL	POL	POL	POL
Other Community, Social and Personal Services	7.00	3.72	0.22	0.39	POL	POL	POL	POL
Private Households with Employed Persons	7.24	4.34	0.23	0.51	POL	POL	POL	ROU

Source: own elaborations based on WIOD world IO tables.
[a] Dark (grey) shading indicates 5 largest (smallest) values.

Table 3

Indexes of productivity in the new EU members with respect to the EU27 total[a]

Sector Name	Productivity in CEE countries (in %)				CEE country with highest productivity			
	2009	1995–2009 Change	Average annual change 2000–2004	Average annual change 2004–2009	1995	2000	2005	2009
Agriculture, Hunting, Forestry and Fishing	27.43	6.62	0.37	0.72	CZE	SVK	EST	SVK
Mining and Quarrying	62.36	46.97	2.69	5.02	SVN	SVN	SVN	SVK
Food, Beverages and Tobacco	30.44	11.52	0.74	1.02	SVN	SVN	SVN	SVK

Table 3 cont.

Textiles and Textile Products	23.69	5.06	0.18	0.82	SVN	SVN	SVN	SVN
Leather, Leather and Footwear	21.29	0.42	–0.18	0.55	SVN	SVN	SVN	SVN
Wood and Products of Wood and Cork	39.98	23.59	1.12	3.09	SVN	SVN	SVN	LVA
Pulp, Paper, Paper, Printing and Publishing	49.00	25.82	1.51	2.69	SVN	SVN	SVN	SVN
Coke, Refined Petroleum and Nuclear Fuel	24.83	11.47	0.64	1.28	SVN	SVN	HUN	SVK
Chemicals and Chemical Products	39.44	18.14	1.58	0.58	SVN	SVN	SVN	SVN
Rubber and Plastics	53.33	30.30	1.94	2.72	SVN	SVN	SVN	SVN
Other Non–Metallic Mineral	37.47	19.74	1.81	0.42	SVN	SVN	SVN	SVN
Basic Metals and Fabricated Metal	41.22	13.94	0.86	1.33	SVN	SVN	SVN	SVN
Machinery, Nec	35.62	22.33	1.55	1.71	SVN	SVN	SVN	SVN
Electrical and Optical Equipment	40.15	20.52	1.72	0.82	SVN	SVN	SVN	SVN
Transport Equipment	41.36	22.72	1.68	1.48	SVN	SVN	SVN	SVN
Manufacturing, Nec; Recycling	34.27	11.34	0.55	1.45	SVN	SVN	SVN	SVN
Electricity, Gas and Water Supply	32.32	18.10	1.02	1.98	SVN	SVN	SVN	SVK
Construction	39.42	15.53	1.48	0.19	SVN	SVN	SVN	SVN
Sale, Maintenance and Repair of Motor Vehicles and Motorcycles; Retail Sale of Fuel	72.58	42.89	1.56	6.83	SVN	SVN	SVN	SVN
Wholesale Trade and Commission Trade, Except of Motor Vehicles and Motorcycles	36.97	13.91	1.30	0.22	SVN	SVN	SVN	SVN
Retail Trade, Except of Motor Vehicles and Motorcycles; Repair of Household Goods	46.73	6.72	0.05	1.57	SVN	SVN	SVN	SVN

Table 3 cont.

Sector Name	Productivity in CEE countries (in %)				CEE country with highest productivity			
	2009	1995–2009 Change	Average annual change 2000–2004	Average annual change 2004–2009	1995	2000	2005	2009
Hotels and Restaurants	35.03	10.92	0.70	0.97	SVN	SVN	SVN	SVN
Inland Transport	45.01	24.81	1.84	1.60	SVN	SVN	SVN	SVK
Water Transport	31.15	10.63	0.63	1.08	SVN	SVN	SVN	SVN
Air Transport	40.54	25.00	1.65	2.13	SVN	SVN	SVN	SVN
Other Supporting and Auxiliary Transport Activities; Activities of Travel Agencies	88.22	55.61	4.21	3.37	SVN	SVN	SVN	SVN
Post and Telecommunications	54.79	33.75	2.38	2.49	SVN	SVN	SVN	SVN
Financial Intermediation	60.46	31.12	2.87	0.61	SVN	SVN	SVN	SVN
Real Estate Activities	22.56	9.25	0.94	–0.03	SVN	SVN	SVN	SVN
Renting of M&Eq and Other Business Activities	58.33	31.08	2.12	2.47	SVN	SVN	POL	SVN
Public Admin and Defence; Compulsory Social Security	37.18	10.52	0.47	1.45	SVN	SVN	SVN	SVN
Education	28.22	11.90	0.84	0.87	SVN	SVN	SVN	SVN
Health and Social Work	29.32	13.33	0.81	1.30	SVN	SVN	SVN	SVN
Other Community, Social and Personal Services	47.32	19.64	0.86	2.77	SVN	SVN	SVN	SVN
Private Households with Employed Persons	28.23	14.52	1.66	–0.51	SVN	SVN	SVN	SVN

Source: own elaborations based on WIOD world IO tables.
[a] Dark (grey) shading indicates 5 largest (smallest) values.

Table 4

Indexes of capital efficiency in the new EU members with respect to the EU27 total[a]

Sector Name	Capital efficiency in CEE countries (in %)				CEE country with highest capital efficiency			
	2009	1995–2009 Change	Average annual change 2000–2004	Average annual change 2004–2009	1995	2000	2005	2009
Agriculture, Hunting, Forestry and Fishing	98.55	–6.21	–0.39	–0.58	LVA	BGR	HUN	LTU
Mining and Quarrying	125.55	9.11	–1.21	5.29	BGR	POL	POL	POL
Food, Beverages and Tobacco	88.30	–3.18	–0.37	0.12	BGR	BGR	POL	LTU
Textiles and Textile Products	88.85	–0.73	–0.50	1.07	POL	LTU	POL	LTU
Leather, Leather and Footwear	94.01	0.28	–0.13	0.40	LVA	LTU	POL	SVK
Wood and Products of Wood and Cork	94.78	14.74	–0.24	4.29	BGR	BGR	POL	LTU
Pulp, Paper, Paper, Printing and Publishing	90.35	1.17	–1.16	3.20	BGR	BGR	POL	LTU
Coke, Refined Petroleum and Nuclear Fuel	100.98	16.99	0.41	3.22	LVA	BGR	LTU	LTU
Chemicals and Chemical Products	78.32	1.79	–0.26	1.10	BGR	BGR	POL	LTU
Rubber and Plastics	63.32	–11.43	–0.52	–1.57	BGR	BGR	LTU	LTU
Other Non–Metallic Mineral	79.89	2.58	–0.26	1.30	BGR	BGR	POL	SVK
Basic Metals and Fabricated Metal	86.08	8.03	0.01	1.97	BGR	BGR	POL	LTU
Machinery, Nec	78.43	–0.10	–0.83	2.06	BGR	BGR	POL	LTU
Electrical and Optical Equipment	84.71	2.12	0.15	0.16	BGR	BGR	POL	LTU
Transport Equipment	79.66	–5.76	–2.06	3.72	BGR	BGR	LTU	LTU
Manufacturing, Nec; Recycling	80.35	–6.60	–1.01	0.89	BGR	LTU	POL	SVK
Electricity, Gas and Water Supply	107.20	39.13	1.12	6.98	BGR	BGR	POL	HUN
Construction	72.34	–0.22	0.03	–0.14	BGR	LTU	POL	SVK

Table 4 cont.

Sector Name	Capital efficiency in CEE countries (in %)				CEE country with highest capital efficiency			
	2009	1995–2009 Change	Average annual change 2000–2004	Average annual change 2004–2009	1995	2000	2005	2009
Sale, Maintenance and Repair of Motor Vehicles and Motorcycles; Retail Sale of Fuel	91.96	2.96	–0.44	1.84	POL	LTU	POL	LTU
Wholesale Trade and Commission Trade, Except of Motor Vehicles and Motorcycles	85.14	9.43	1.16	–0.53	BGR	LTU	POL	SVK
Retail Trade, Except of Motor Vehicles and Motorcycles; Repair of Household Goods	103.51	–12.29	–1.75	1.29	POL	LTU	POL	LTU
Hotels and Restaurants	75.17	–6.77	–0.53	–0.37	LVA	LTU	POL	LTU
Inland Transport	94.54	–8.99	–0.51	–0.97	BGR	BGR	POL	LTU
Water Transport	119.38	–12.34	–2.64	3.51	SVN	BGR	POL	LTU
Air Transport	89.52	–17.89	–2.33	1.36	BGR	BGR	ROU	LTU
Other Supporting and Auxiliary Transport Activities; Activities of Travel Agencies	80.43	–28.06	–3.82	2.54	BGR	BGR	POL	LTU
Post and Telecommunications	94.99	38.63	4.01	–0.36	BGR	BGR	POL	LTU
Financial Intermediation	85.68	24.25	2.46	–0.08	LVA	EST	LTU	SVK
Real Estate Activities	128.38	–21.78	–1.56	–1.55	LVA	LVA	ROU	EST
Renting of M&Eq and Other Business Activities	118.94	30.58	2.92	0.34	LVA	HUN	POL	SVK
Public Admin and Defence; Compulsory Social Security	95.45	–14.72	–1.95	1.20	POL	BGR	CZE	HUN
Education	71.79	–10.23	–0.77	–0.63	LVA	LTU	LTU	LTU
Health and Social Work	77.46	1.77	–0.05	0.57	LVA	LTU	HUN	LTU
Other Community, Social and Personal Services	88.86	–1.09	0.45	–1.40	LVA	BGR	LTU	LTU
Private Households with Employed Persons	79.93	–1.39	0.74	–2.19	BGR	BGR	POL	LTU

Source: own elaborations based on WIOD world IO tables.
[a] Dark (grey) shading indicates 5 largest (smallest) values.

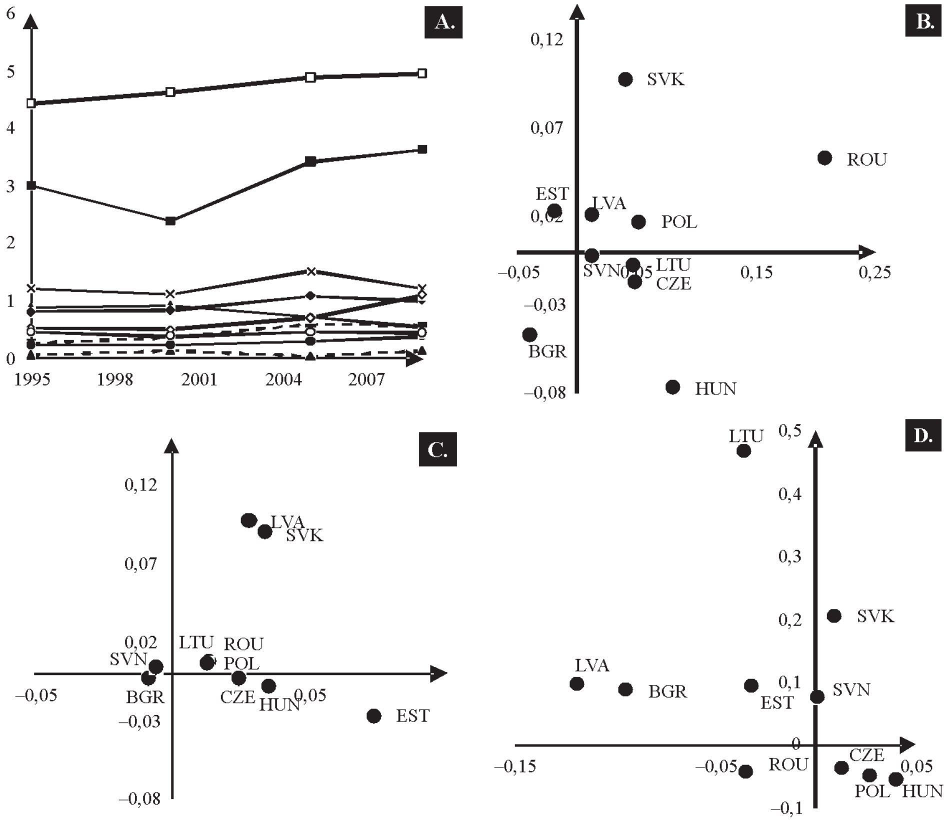

Figure 1. Share in value added and indexes of productivity and capital efficiency attributed in the sector of **Agriculture, Hunting, Forestry and Fishing** in the new EU members with respect to the EU27 total

Figure notes:

Panel **A:** Share in value added (in%) in the new EU members with respect to the EU27 total.

Legend: BGR, CZE, EST, HUN, LTU, LVA, POL, ROU, SVK, SVN

Panel **B:** Change of share in value added (in %) in the new EU members with respect to the EU27 total (horizontal axis: change in 2000–2004 period, 2004–2009: change in 2004–2009 period).

Panel **C:** Annual change of index of productivity (in %) in the new EU members with respect to the EU27 total (horizontal axis – before EU accession, vertical axis – after EU accession).

Panel **D:** Annual change of index of capital efficiency (in %) in the new EU members with respect to the EU27 total (horizontal axis – before EU accession, vertical axis – after EU accession).

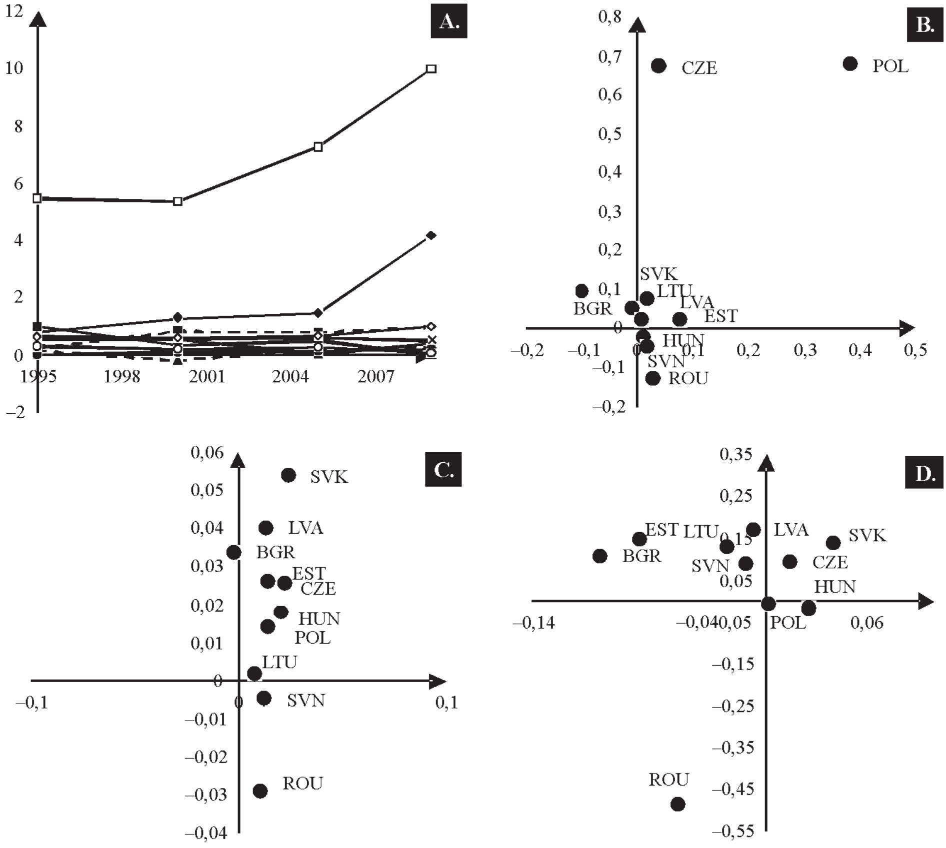

Figure 2. Share in value added and indexes of productivity and capital efficiency attributed in the sector of **Mining and Quarrying** in the new EU members with respect to the EU27 total

Figure notes:

Panel **A:** Share in value added (in%) in the new EU members with respect to the EU27 total.

Legend: BGR, CZE, EST, HUN, LTU, LVA, POL, ROU, SVK, SVN

Panel **B:** Change of share in value added (in %) in the new EU members with respect to the EU27 total (horizontal axis: change in 2000–2004 period, 2004–2009: change in 2004–2009 period).

Panel **C:** Annual change of index of productivity (in %) in the new EU members with respect to the EU27 total (horizontal axis – before EU accession, vertical axis – after EU accession).

Panel **D:** Annual change of index of capital efficiency (in %) in the new EU members with respect to the EU27 total (horizontal axis – before EU accession, vertical axis – after EU accession).

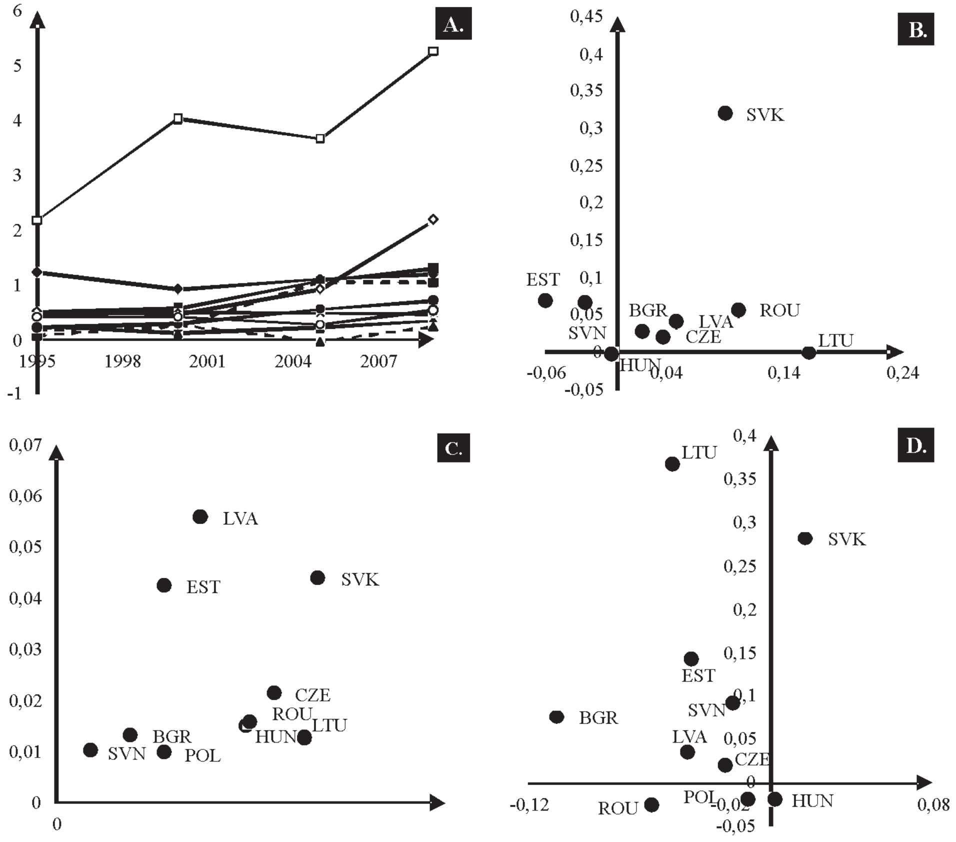

Figure 3. Share in value added and indexes of productivity and capital efficiency attributed in the sector of **Wood and Products of Wood and Cork** in the new EU members with respect to the EU27 total

Figure notes:

Panel **A:** Share in value added (in%) in the new EU members with respect to the EU27 total.

Legend: BGR, CZE, EST, HUN, LTU, LVA, POL, ROU, SVK, SVN

Panel **B:** Change of share in value added (in %) in the new EU members with respect to the EU27 total (horizontal axis: change in 2000–2004 period, 2004–2009: change in 2004–2009 period).

Panel **C:** Annual change of index of productivity (in %) in the new EU members with respect to the EU27 total (horizontal axis – before EU accession, vertical axis – after EU accession).

Panel **D:** Annual change of index of capital efficiency (in %) in the new EU members with respect to the EU27 total (horizontal axis – before EU accession, vertical axis – after EU accession).

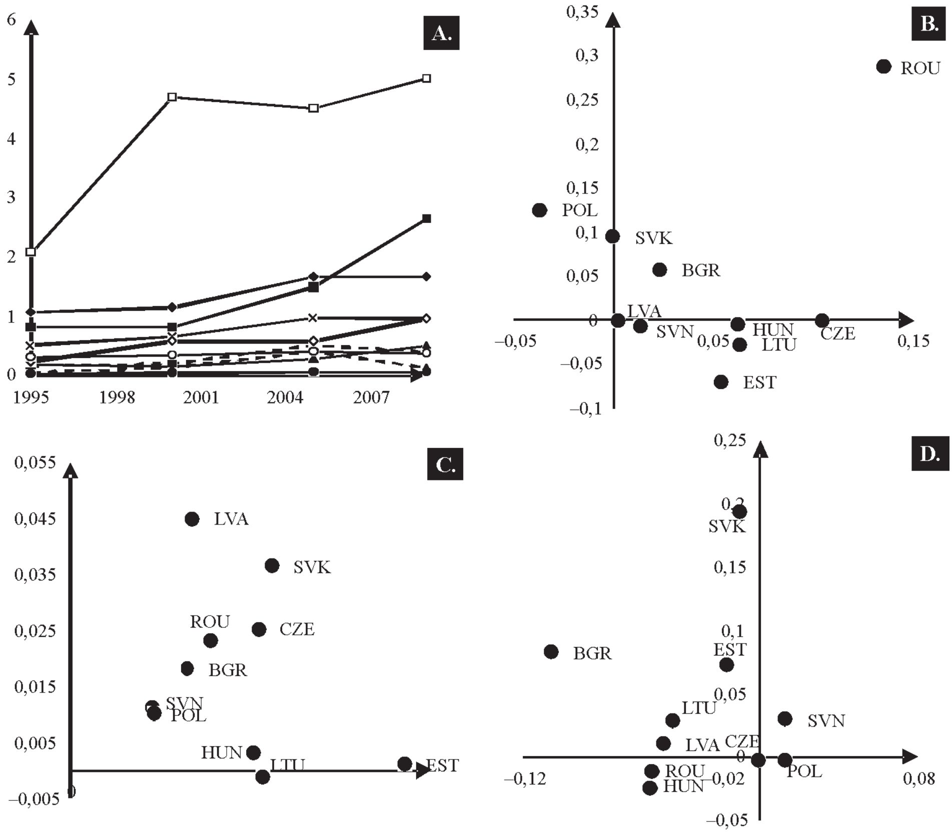

Figure 4. Share in value added and indexes of productivity and capital efficiency attributed in the sector of **Other Non-Metallic Mineral** in the new EU members with respect to the EU27 total

Figure notes:

Panel A: Share in value added (in%) in the new EU members with respect to the EU27 total.

Legend: BGR, CZE, EST, HUN, LTU, LVA, POL, ROU, SVK, SVN

Panel **B:** Change of share in value added (in %) in the new EU members with respect to the EU27 total (horizontal axis: change in 2000–2004 period, 2004–2009: change in 2004–2009 period).

Panel **C:** Annual change of index of productivity (in %) in the new EU members with respect to the EU27 total (horizontal axis – before EU accession, vertical axis – after EU accession).

Panel **D:** Annual change of index of capital efficiency (in %) in the new EU members with respect to the EU27 total (horizontal axis – before EU accession, vertical axis – after EU accession).

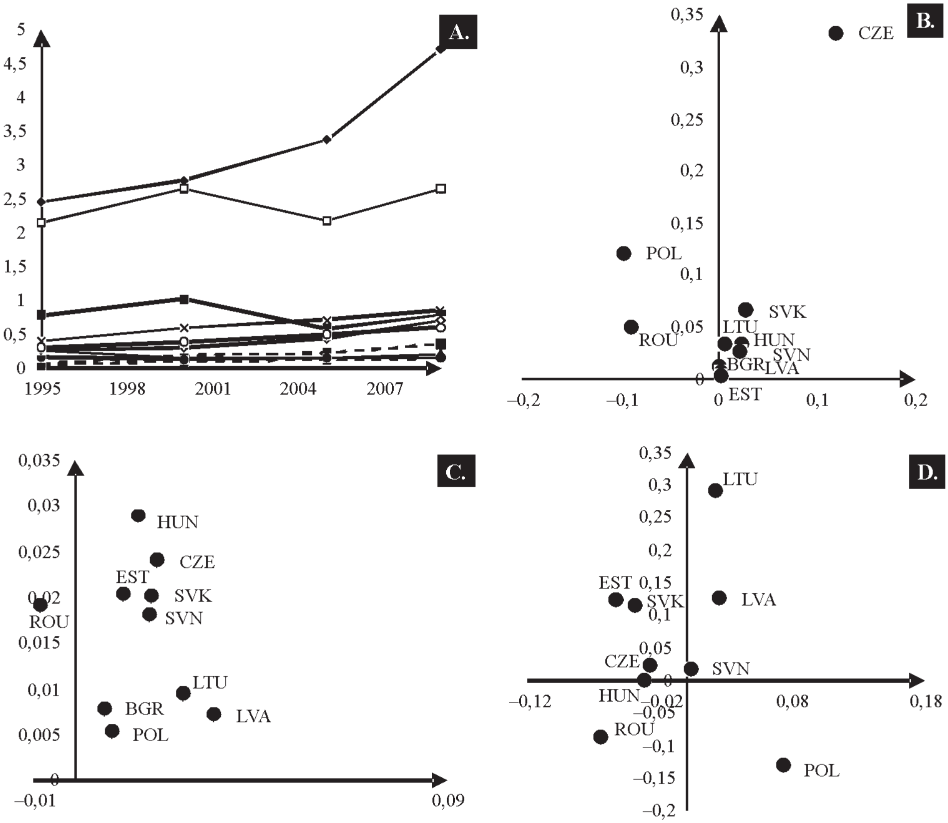

Figure 5. Share in value added and indexes of productivity and capital efficiency attributed in the sector of **Other Supporting and Auxiliary Transport Activities**, **Activities of Travel Agencies** in the new EU members with respect to the EU27 total

Figure notes:

Panel **A:** Share in value added (in%) in the new EU members with respect to the EU27 total.

Legend: BGR, CZE, EST, HUN, LTU, LVA, POL, ROU, SVK, SVN

Panel **B:** Change of share in value added (in %) in the new EU members with respect to the EU27 total (horizontal axis: change in 2000–2004 period, 2004–2009: change in 2004–2009 period).

Panel **C:** Annual change of index of productivity (in %) in the new EU members with respect to the EU27 total (horizontal axis – before EU accession, vertical axis – after EU accession).

Panel **D:** Annual change of index of capital efficiency (in %) in the new EU members with respect to the EU27 total (horizontal axis – before EU accession, vertical axis – after EU accession).

6. Concluding remarks

Due to decreasing costs of communication and coordination, it has become more profitable to apply fragmentation of the production process. The stages of production are conducted at its lowest-cost locations. Knowledge about the size and future development of splitting of international production is still not extensively and sufficiently investigated in the economic literature. Some empirical papers have been concerned with cross-border fragmentation based on foreign investment flow data of firms and their affiliates. Using a decomposition technique that has recently become feasible due to the development of the World Input-Output Database, one can trace the value added by all labor and capital that is directly and indirectly used for the production of final manufacturing goods. The production systems of manufacturing goods are susceptible to a large extent to international fragmentation. The reason is that many stages of the production process can be conducted in different countries with little differences in quality yet with an essential difference in price.

Most of the previous empirical studies were dedicated to high-end electronic products and focused at one point in time. Therefore, one may ask very important questions concerning the extent to which these findings also represent more-general patterns. Another question is this: How pervasive is the process of international production fragmentation for a large number of other products? Finally, a very important research objective is the determination of specialization patterns between high-income and emerging economies that participate in these production chains.

This paper refers mostly to the abovementioned research questions. It is one of the first studies dedicated to the importance and efficiency of the CEE transition economies in the process of building a comparative advantage of the European Union in global value chains. We proposed original modifications and extensions of the recently presented methodological developments in ex-post accounting framework in global value chains in order to obtain detailed empirical results, both for the whole analyzed group of CEE economies as well as at a country-and-sector-specific level.

The empirical results show that the role of the selected CEE economies in transition in creating value added with respect to the total value added in the European Union in the GVC framework was biggest in case of the sectors related to agriculture, mining, wood products, metal production, and travel and tourism. The highest shares in the EU total GVC value added were obtained in the largest countries (e.g., Poland). We also found that, after two decades of transition, the GVC-embedded measures of workforce productivity in the ten CEE economies in 2009 were still much lower as compared to the EU average for most of the sectors. However, we found that, during this period, the growth rates of these indexes were, in general, positive. Moreover, these indexes were increasing, especially after

EU accession. The highest levels and growth rates of GVC-embedded workforce productivity were found for the smaller CEE countries.

Different conclusions arise from the analysis of the GVC-embedded indexes of capital efficiency. After two decades of transition, the measures of capital efficiency in the ten CEE economies in 2009 were comparable to the EU average for most of the sectors. Moreover, during this period, the growth rates of these indexes were, in general, positive. However, the growth rates of these indexes were dropping after EU accession. The highest levels and growth rates of indexes of capital efficiency were, once again, found for smaller CEE countries.

Despite our efforts, it is likely that some aspects of tracing the dynamics of relative GVC-embodied value added in new EU members in transition were not captured in our study. In further research on fragmentation in CEE countries, one should test whether international fragmentation is mostly regional (i.e., it is observed only within certain groups of neighboring countries and regional trade blocs) or whether it is more global (i.e., the production process also involves countries outside the region). The results of such an analysis would have significant implications for the shape of trade policies. In the case of fragmentation within regions, regional trade agreements are sufficient to create a rise in welfare from supply chain trade. In contrary to the fragmentation within the region, the global value chains would need multiregional trade agreements.

One may also claim that, for all industries, Poland seems to be the driver of the relative GVC value added due to the size effect of this economy in the group of ten CEE economies under study. At the same time, the development of Poland has been rather different from the other CEE countries examined. However, since the absolute size matters, the role of the other countries in value added creation in CEE could be partly hidden in the presentation chosen in this paper. Therefore, an interesting direction for future research would be to analyze the data for individual countries and group the industries according to characteristics such as low/medium/high-tech, etc., in order to shed some light on the differences between the CEE countries under study. It the transition period, one could list rather liberal countries like Poland, Slovenia, and Hungary, which – from a perspective of their institutions – were much closer to the Western economies than the Baltic States. Thus, it is quite likely that differences among these countries may still exist. Moreover, Bulgaria and Romania became members of the EU in 2007; therefore, there could also be some delay in the development of fragmentation process in the case of these two countries.

Special attention should be given to the mining industry. In general, the relative GVC value added in this sector in the ten CEE economies with respect to the EU27 was rising in the transition period, mostly because of the shrinking coal production and use observed in the EU15 countries caused by the shift towards low-carbon energy sources in the richest European economies. In other words, the significant role of CEE economies in the creation of the relative GVC value added in the mining sector seems to be spurious. Thus, it seems reasonable to put more attention in future research to industries such as food, electrical and optical equipment, transport equipment and trade, or education and health & social work, which recently seem to be more important than mining for the relative GVC value added in the CEE economies in transition. This type of sector-oriented analysis could help us better understand the structural change and the importance of these industries in CEE economies and, thus, deserves considerable attention in the context of future research.

Among general directions for future research, one could also list an attempt to conduct an analogous study in a framework of dynamic input-output modeling.

References

[1] Ali-Yrkko, J., Rouvinen, P., Seppala, T. and Yla-Anttila, P. (2011) 'Who captures value in global supply chains?', *Journal of Industry, Competition and Trade*, 11(3), pp. 263–278.

[2] Ando, M., and Kimura, F. (2005) 'The Formation of International Production and Distribution Networks in East Asia', in Takatoshi, I. and Rose, A.K. (eds) International Trade in East Asia, NBER – East Asia Seminar on Economics, vol. 14, pp. 177–216.

[3] Baldwin, R., Venables, A.J. (2013) 'Spiders and Snakes: Offshoring and Agglomeration in the Global Economy', *Journal of International Economics*, vol. 90(2), pp. 245–254.

[4] Bems, R., Johnson, R.C. and Yi, K.-M. (2011) 'Vertical linkages and the collapse of global trade', American Economic Review, vol. 101(3), pp. 308–312.

[5] Dedrick, J., Kraemer, K.L. and Linden, G. (2010) 'Who profits from innovation in global value chains? A study of the iPod and notebook PCs', Industrial and Corporate Change, vol. 19(1), pp. 81–116.

[6] Di Mauro, F. and Forster, K. (2008) 'Globalisation and the competitiveness of the euro area', ECB Occasional Paper No. 97.

[7] Dudenhoeffer, F. (2005) 'Wieviel Deutschland steckt im Porsche?', Ifo Schnelldienst, vol. 58(24), pp. 3–5.

[8] Escaith, H. and Inomata, S. (2011) Trade Patterns and Global Value Chains in East Asia: From trade in Goods to Trade in Tasks, Geneva: WTO/IDE-JETRO.

[9] Feenstra, R.C. (1998) 'Integration of trade and disintegration of production in the global economy', *Journal of Economic Perspectives*, vol. 12(4), pp. 31–50.

[10] Feenstra, R.C. (2010) Offshoring in the Global Economy: Microeconomic Structure and Macroeconomic Implications, Cambridge: MIT Press.

[11] Fukao, K., Ishido, H. and Ito, K. (2003) 'Vertical Intra-Industry Trade and Foreign Direct Investment in East Asia', *Journal of the Japanese and International Economies*, vol. 17(4), pp. 468–506.
[12] Grossman, G. and Rossi-Hansberg, E. (2008) 'Trading tasks: a simple theory of offshoring', *American Economic Review*, vol. 98(5), pp. 1978–1997.
[13] Gurgul, H. and Lach, Ł., (2014) 'Globalization and economic growth: Evidence from two decades of transition in CEE', *Economic Modelling*, vol. 36, pp. 99–107.
[14] Hanson, G.H., Mataloni, R.J.and Slaughter, M.J. (2005) 'Vertical Production Networks in Multinational Firms', Review of Economics and Statistics, vol. 87(4), pp. 664–678.
[15] Herrendorf, B., Rogerson, R. and Valentinyi, A. (2013) 'Two perspectives on preferences and structural transformation', American Economic Review, vol. 103(7), pp. 2752–2789.
[16] Hummels, D., Ishii, J. and Yi, K-M. (2001) 'The Nature and Growth of Vertical Specialization in World Trade', *Journal of International Economics*, vol. 54(1), pp. 75–96.
[17] Isard, W. (1951) 'Interregional and Regional Input-Output Analysis: A Model of a Space Economy', *Review of Economics and Statistics*, vol. 33(4), pp. 318–328.
[18] Johnson, R.C. and Noguera, G. (2012) 'Accounting for Intermediates: Production Sharing and Trade in Value Added', *Journal of International Economics*, vol. 86(2), pp. 224–236.
[19] Kalvet, T. and Kattel, R. (2006) Creative Destruction Management: Meeting the Challenges of the Techno-economic Paradigm Shift, Tallinn: Praxis.
[20] Koopman, R., Wang, Z. and Wei, S.-J. (2012) 'Estimating domestic content in exports when processing trade is pervasive', *Journal of Development Economics*, vol. 99(1), pp. 178–189.
[21] Kuboniwa, M. (2015) 'Russia's Global Value Chain Using Modified World Input-Output Data', *Eastern European Economics*, vol. 53, pp. 277–308.
[22] Marin, D. (2006) 'A New International Division of Labor in Europe: Outsourcing and Offshoring to Eastern Europe', *Journal of the European Economic Association*, vol. 4, pp. 612–622.
[23] Marin, D. (2011) 'The Opening Up of Eastern Europe at 20: Jobs, Skills, and «Reverse Maquiladoras»' in Jovanovic, M.N. (ed.) International Handbook on the Economics of Integration, vol. 2, Chap. 13, Cheltenham: Edward Elgar, pp. 293–323.
[24] Miller, R.E. (1966) 'Interregional feedback effects in input-output models: some preliminary results', Papers in Regional Science, vol. 17(1), pp. 105–125.
[25] Ottaviano, G., Taglioni, D. and di Mauro, F. (2009) 'The euro and the competitiveness of European firms', *Economic Policy*, vol. 24(57), pp. 5–53.

[26] Sinn, H.-W. (2006) 'The pathological export boom and the bazaar effect: how to solve the German puzzle', *The World Economy*, vol. 29(9), pp. 1157–1175.
[27] Timmer, M.P. (2012) The World Input-Output Database (WIOD): Contents, Sources and Methods, WIOD Working Paper 10, [Online], Available: http://www.wiod.org/publications/papers/wiod10.pdf [12 May 2014].
[28] Timmer, M.P., Los, B., Stehrer, R. and de Vries, G.J. (2013) 'Fragmentation, Incomes and Jobs. An Analysis of European Competitiveness', *Economic Policy*, vol. 28(76), pp. 613–661.
[29] Trefler, D. and Zhu, S. (2010) 'The structure of factor content predictions', *Journal of International Economics*, vol. 82(2), pp. 195–207.

Managerial Economics
2016, vol. 17, no. 1, pp. 59–75
http://dx.doi.org/10.7494/manage.2016.17.1.59

Henryk Gurgul*, Artur Machno*

The impact of asynchronous trading on Epps effect. Comparative study on Warsaw Stock Exchange and Vienna Stock Exchange[1]

1. Introduction

A very important research topic is describing the way in which different asset price movements are correlated. Modern portfolio theory methods are based on observed correlations between returns at daily or larger time scales. One expects that coarse scale correlations originate from intraday movements that are strongly correlated. This implies the important question of how to obtain better estimates of such correlations by using high-frequency data. Here, one can observe an analogous paradox as for volatility estimation under microstructure noise. It is surprising that the correlation coefficient is an increasing function of the time resolution, and that correlation very quickly decays and almost vanishes at a very high frequency. The dependence of correlations between stock prices on the sampling frequency of time series involves a phenomenon called the Epps effect. The actual correlations between returns of stocks decrease as the sampling frequency of data increases. The Epps effect has received considerable attention, not only from economists but also from mathematicians and theoretical physicists. But there are only a few contributions to the subject.

The Epps effect seems to be an unexpected phenomenon at a first glance. However it is simply explainable by different factors; e.g., the asynchrony of trade times. It is clear that 1-minute returns are, in fact, returns for the last trade prices

* AGH University of Science and Technology in Krakow, Faculty of Management, Department of Applications of Mathematics in Economics; e-mails: henryk.gurgul@gmail.com; artur.machno@gmail.com

[1] Financial support for this paper comes from the National Science Center of Poland (Research Grant DEC-2012/05/B/HS4/00810) and is gratefully acknowledged by Henryk Gurgul.
Financial support for this paper comes from the Dean of Faculty of Management, AGH University (statutory activity no. 15/11.200.296) and is gratefully acknowledged by Artur Machno.

before the appropriate times. If one takes very short horizons or non-liquid assets, then some returns are calculated for the same trade. Therefore, those returns are, of course, zero. They do not give any additional contribution to correlations. It is clear that the situation where there are no trades in a time interval and, therefore, no correlation is the extreme case. However, even if the asynchrony is not in its extreme case, the Epps effect is present.

Financial transactions supply information about current prices. In addition, they assure fixing of the market value of traded stocks until the next trade takes place. From this point of view, the financial literature considers two main groups of contributions to the Epps effect. The first group of authors is concerned with the fact that price dynamics are not synchronous across stocks. This means that transactions are conducted at different times for different shares. The second one relates the Epps effect to actual lagged correlations. The main reason for this phenomenon is the temporary or permanent effect of individual trades on the price dynamics.

The main goal of this paper is to compare the Epps effect on the stock markets in Vienna and Warsaw. There are indications suggesting that the VSE and WSE may be linked. The linkage is mostly indirect because of a strong relationship with Germany. First of all, the VSE and WSE are somewhat similar in some aspects. Both stock markets are of similar size. The main indices of these markets are among the largest in Central and Eastern Europe, and they have been quoted for a similar period of time[2]. In addition, the VSE and WSE have been competing markets in recent years. On the other hand, the VSE is a developed market, while the WSE still belongs to emerging markets despite of development in recent years. The important channel of interrelations may be foreign trade. Germany is the most important trading partner for both the Austrian and Polish economies.

The computations in the empirical part of this paper are made in the R environment. All codes that have been used are available from the authors upon request. Additionally, the authors can provide the data and codes needed to reproduce all tables and plots used in this article.

The main finding of the study is that the Epps effect is surprisingly different for Austrian and Polish assets. The analysis shows that all considered Polish pairs exhibit a similar relationship in terms of correlation and the impact of the asynchrony on it. Moreover, the results for the WSE are as generally expected. In the case of Austrian pairs, we have found several aberrations. Only one analyzed Austrian pair has similar properties to all Polish pairs.

[2] ATX20 index (VSE) is quoted from January 2 1991, and the WIG20 index (WSE) is published from April 16, 1994.

To the best of our knowledge, this is the first contribution dealing simultaneously with the Epps effect on the stock markets in Vienna and Warsaw. The comparison of the Epps effect (and particularly the asynchrony trading impact in it) on stocks from the given stock markets is of the most valuable and novel part of the paper.

The content of this paper is as follows; the next section briefly reviews the most important recent contributions concerning the Epps effect, and the main conjectures are presented in the third section. The following section presents the methodology, and the fifth section is concerned with the dataset and empirical results. Finally, we summarize the major conclusions and suggest directions for future research.

2. Literature review

In his contribution in 1979, Thomas Wake Epps provides empirical evidence of a dramatic drop in correlations between stocks with a decreasing sampling time horizon. This observation has been reported for different markets.

In the opinion of most of the authors, it is not obvious which type of price formation process is the source of the Epps effect. Lundin et al. (1998) found an essential inverse link between correlation and activity. They found that growth in the quantity of assets traded reduces the Epps effect. This observation is in favor of the view that non-synchronicity is the main source of correlation drop at higher frequencies.

According to several authors (e.g., Conlon et al., 2009), a number of financial mechanisms of stock price and trading-volume formation are the source of the Epps effect. The goal of Mastromatteo et al. (2011) was not to derive a complete description of the Epps effect. The contributors stressed that they rather aimed at identifying statistical causes that can be compensated for directly, without the need to adjust parameters, model calibrations, or an optimal sampling frequency. They identified two major causes: the asynchrony of the time series and the impact of discretization by tick-size.

Toth and Kertesz (2009) assumed that there is a price of an asset at any time. However, it is observable only at the times of trade. This assumption is suggested by the observed Epps effect. In the opinion of the authors, the impact of the asynchrony is weak in relation to the impact of a static lag for which they developed a model. The authors think that the source of decline of the Epps effect with time is increased market efficiency.

Munnix et al. (2010) found that the calculated correlation consists of an actual correlation (the coefficient that would be observed if prices were quoted continuously and priced with a continuous value) and an uncorrelated part caused by asynchronous trading. The authors demonstrated that the asynchrony of trades and the decimalization of stock prices are the main sources of decline of the correlation coefficients towards smaller return intervals (i.e., the Epps effect). Munnix et al. (2010) argued that these distortions are of a purely statistical origin and depend on the properties of the time series.

The most recent contribution by Gurgul and Machno (2015) is concerned with the impact of asynchronous trading on the Epps effect for some stocks quoted on the WSE in both theoretical and empirical aspects.

To conclude, there are certainly many phenomena contributing to the Epps effect. However, one of the most important sources of the Epps effect seems to be the asynchronous trading.

To the best of the knowledge of the authors, this is the one of the first studies concerning the impact of asynchronous trading on the Epps effect in the stock exchanges in Vienna and Warsaw.

3. Description of the dataset

In this part of the paper, we briefly describe the stocks chosen for the analysis. We use four stocks for both stock exchanges under study. Each four pairs are the most-frequently-traded stocks of a particular exchange. Below, we give a short data description of datasets from both stock exchanges. Our analysis is based on the tick-by-tick data from both stock exchanges. We use the most recent tick-by-tick data available to us for the Vienna and Warsaw stock exchanges (2013 and 2014, respectively).

3.1. Warsaw Stock Exchange data

The dataset from the Warsaw Stock Exchange used in our contribution encompass every operation that took place on the WSE from 1.01.2014 to September 22.2014. We stress that the electronic system of the WSE has been changed since August 01.2013. The most important change from the point of view of our analysis is the rise of exactness of the trade time from seconds to microseconds. In this way, it overcame the problem of two or more transactions at the same time. Since this date, each two transactions are distinguishable in the time. The trade takes place on the WSE between 9:00 a.m. and 5:05 p.m. In addition, there is a break from 4:50 p.m. to 5:00 p.m. Moreover, during the last five minutes, trade on the WSE is not conducted as it is during normal hours. To avoid biases in the results caused by data, we restrict our dataset and take into account the transactions conducted not later than 4:50 p.m. in our analysis. After these restrictions, the time span of our analysis includes 470 minutes of trade each day.

The WSE is the biggest stock exchange from the countries that entered the European Union in 2004 or since. The capitalization of the WSE is about 1.2 trillion PLN. Now, 487 stocks are listed in the Warsaw Stock Exchange.

The transactions conducted on the WSE are quoted in Polish zloty (PLN). The exchange rates during the first days of February 2016 are approximately given by 1 USD = 4.04 PLN and 1 EUR = 4.41PLN. In the text, all prices are expressed in PLN for Polish stocks.

The most-frequently-traded stocks on the WSE during the period under study (January 01, 2014 – September 22, 2014) were the four stocks described below.

KGHM Polska Miedź (KGHM) is the largest mining and metallurgy company in Poland. Its main product is copper. The market value of the company is approximately 19 billion PLN. The company employs about 30,000 employees.

Powszechna Kasa Oszczędności Bank Polski (PKOBP) is the largest Polish bank. The market value of the bank is about 36 billion PLN. PKOBP employs about 26,000 employees.

Polska Grupa Energetyczna (PGE) is the largest Polish energy producer and provider. The biggest investor of PGE is the Polish government (it owns more than 60% of the shares of PGE). The value of the company is about 33 billion PLN. The company employs about 41,000 employees.

Powszechny Zakład Ubezpieczeń (PZU) is the largest Polish financial institution. Its business area is insurance, and it dominates the Polish insurance market. The value of the company is about 37 billion PLN. The company employs more than 17,000 employees.

3.2. Vienna Stock Exchange data

The dataset from the Vienna Stock Exchange used in this contribution includes every operation that took place on the VSE from January 01, 2013 to December 31, 2013. The exactness of the trade time for this dataset is 0.01 seconds. The continuous trade hours on the VSE are during the time span, from 9:00 a.m. to 12:00 p.m. and then 12:03 p.m. to 5:30 p.m. This time span consists of 507 minutes of trade each day, with a 3-minute gap in the middle.

The VSE is a relatively small mature market. The capitalization of the VSE is about 87.7 billion EUR. There are 96 listed companies, and the main equity index is the ATX20 (consisting of 20 companies).

All transactions on the VSE are quoted in Euro (EUR). In the text, we present all values in EUR for Austrian stocks.

The most-frequently-traded assets on the VSE during the period January 01, 2013 to December 31, 2013 are described below.

Erste Group Bank AG (EBS) is one of the biggest financial institutions in Central and Eastern Europe. The main subsidiaries of the group are Austrian Erste Bank der oesterreichischen Sparkassen AG, Czech Česká spořitelna, Slovakian Slovenská sporiteľňa, Hungarian Erste Bank Hungary, Croatia Erste & Steiermärkische Bank, Serbian Erste Bank a.d. Novi Sad, Romanian Banca Comercială Română, and Slovenian Banka Sparkasse. The value of the group is about 10 billion EUR, and it employs about 46,000 employees.

OMV (formerly Österreichische Mineralölverwaltung) is an international oil and gas company. It is active in many areas of the oil and gas field, like exploration, production, refining, etc. The company employs about 26,000 employees.

Andritz AG (ANDR) is a plant engineering company. It consists of Andritz Hydro, Andritz Pulp and Paper, Andritz Metals, Andritz Separation, and Andritz Feed & Biofuel. The value of the company is about 1 billion EUR, and it employs about 24,000 employees.

Voestalpine AG (VOE) is an international steel-based corporation. Among others, the company is active in railway systems and the tool steel industry. The value of the company is about 5 billion EUR, and it employs about 46,000 employees.

4. Duration

In this part of the paper, visualization of the data is presented. We have randomly chosen two days (namely, February 19 and July 7, 2014) for the descriptive analysis for the WSE. Figure 1 presents all trades of the four analyzed stocks from the WSE on the examplary day. The stocks are very liquid; therefore, it is hard to see differences in the durations. By duration, we mean the time that elapses between trades.

Analogously, we have randomly chosen two days (namely, April 5 and May 15, 2013) for the descriptive analysis for the VSE. Figure 2 presents all trades of the four analyzed stocks from the VSE on the exemplary day.

In Table 1, the numbers of trades of the stocks on the VSE and WSE during the analyzed periods are collected. We also included summary statistics for the duration on both stock exchanges. The duration distribution in the case of the VSE and WSE that are strongly linked with greater markets seems to be an interesting topic; however, the modeling of duration is not the topic of this paper.

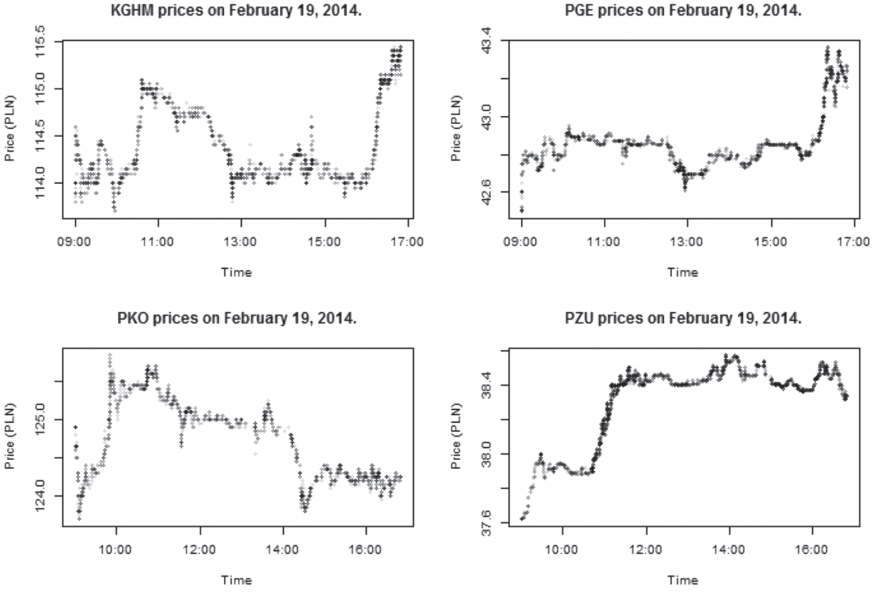

Figure 1. All transactions of KGHM, PKO, PGE, and PZU on February 19, 2014. The color of dots is grey with 90% transparency

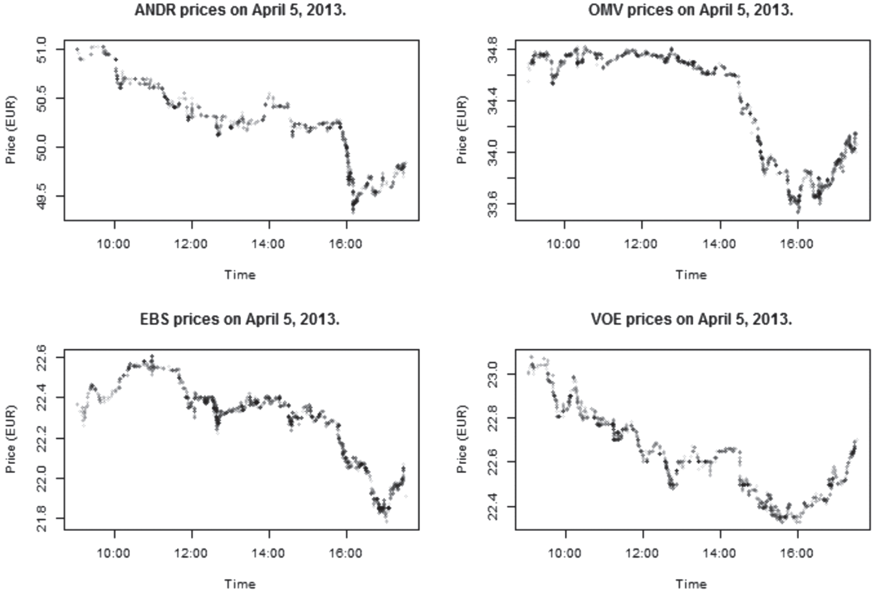

Figure 2. All transactions of ANDR, EBS, OMV, and VOE on April 5, 2014. The color of dots is grey with 90% transparency

The descriptive statistics are quite similar for the stocks under study and days on the WSE. One can see that the minimums are the same (1 microsecond). This is probably caused by the electronic system. The first quartiles take the values from the interval between 4 and 7 microseconds. This is a very small interval. Here raises the question: How is it possible that a quarter of operations take place immediately after the preceding one? Means and the number of transactions are strictly related characteristics of trade on a stock exchange. The mean is the total trade time on the day (470 minutes, 28,200 seconds) denominated by the number of transactions. The third quartiles start from 4.1 seconds, and their upper bound is 5.9 seconds. This also is not a wide interval. The wider intervals determine maximums. They vary from 14 to 26.3 minutes. Contrary to other quartiles, the median of the duration quite strongly varies for the analyzed datasets. This observation is even strongly pronounced if the analysis is conducted separately for various days. The distribution of durations is "U" shaped; this means that the durations are either very low or relatively large. There are only a few little durations near the median. This is probably the reason for the strong variation of the median across stocks and days.

The exactness of the data from the VSE due to its electronic system is lower; therefore, the minimum is zero for every stock. The other statistics show that the trading is much more intense on the WSE as compared to the VSE. What is apparent and worth further, deeper analysis is that the first quartiles for Polish stocks are more than 10,000 times lower than Austrian stocks, while means and median are about 10 times lower. It might be the case that the electronic systems work essentially differently on those stock exchanges.

Table 1

Descriptive statistics for tick data

	Trades	Min.	Q_1	Median	Mean	Q_3	Max.
KGHM	695 202	0.000001	0.000004	0.0522	7.381	4.3810	844
PKOBP	645 032	0.000001	0.000005	0.0911	7.955	4.1330	977
PGE	497 428	0.000001	0.000005	0.0954	10.31	4.9860	1160
PZU	443 598	0.000001	0.000007	0.3516	11.57	5.9110	1580
ANDR	183 768	0	0.07	3.79	41.07	43.93	538
EBS	326 368	0	0.12	3.74	23.12	25.71	506
OMV	159 464	0	0.08	4.07	47.54	37.02	759
VOE	214 768	0	0.17	6.87	35.14	39.42	846

Note: The total number of transactions for a stock during the analyzed periods is in the column "Trades." Descriptive statistics for the time between trades are presented in seconds.

The analysis of the Epps effect on both stock markets requires transformation of the tick-by-tick data into returns of different lengths. We conducted this transformation of the tick-by-tick data into a return series for the chosen intervals by the R code (which is available upon request from the authors).

In Figure 3 and the following figures, we present plots for the transformed data for companies from both stock markets. The most important difference between the tick-by-tick data and the transformed data is that the latter's length is fixed and the labels (times) are fixed. For an easy visual inspection, 1-minute data in Figures 3 and 4 and 10-minute data in Figures 5 and 6 are included. The return is computed as the difference between the last trades before a certain time divided by the prior one. The return in the first interval from the chosen time span is calculated by defining the prior as the first transaction of the day. If the case of a lack of transactions in the last interval, we set by definition the return at zero. It is clear that this problem takes place in cases when narrow intervals are defined. Visualized return series are similar to discrete data for an asset whose maximum and minimum prices on the chosen day do not much differ. The changes of prices themselves are discrete because of the tick size. The calculated returns are transformed ticks denominated by the current price. Hence, if the price is stable during the day, the time series of returns looks discrete.

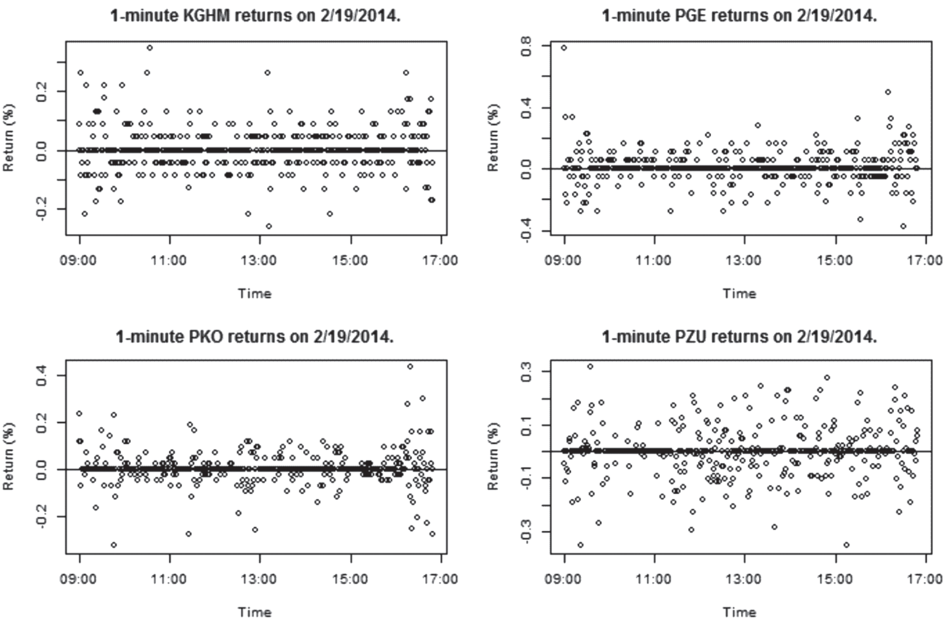

Figure 3. 1-minute percentile returns of KGHM, PKO, PGE, and PZU on February 19, 2014

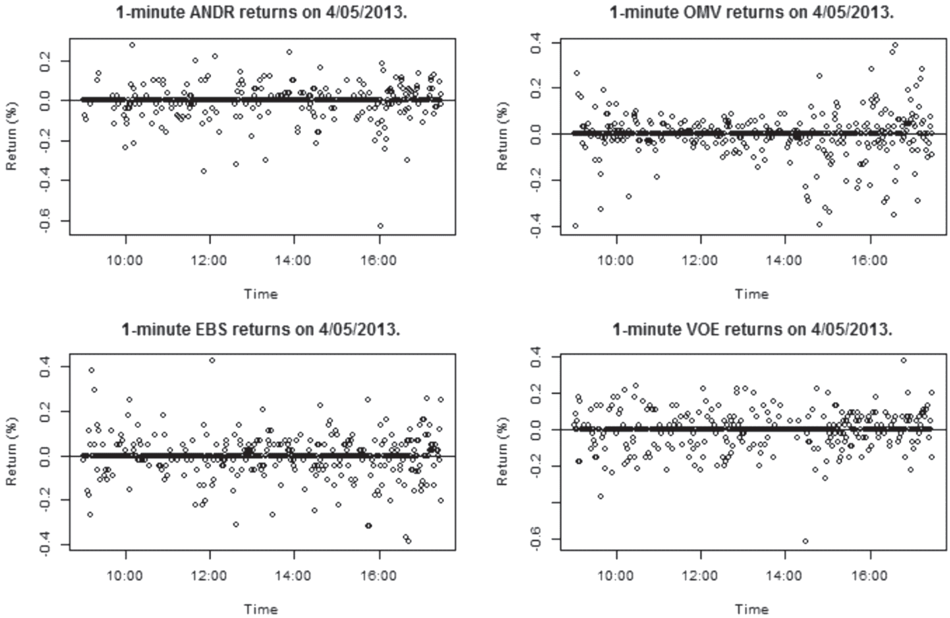

Figure 4. 1-minute percentile returns of ANDR, EBS, OMV, and VOE on April 5, 2013

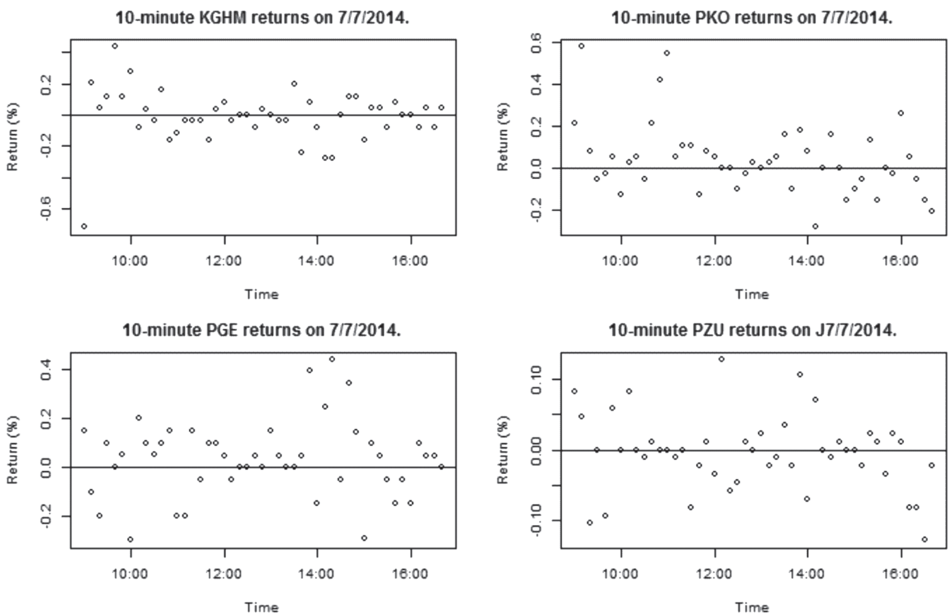

Figure 5. 10-minute percentile returns of KGHM, PKO, PGE, and PZU on July 7, 2014

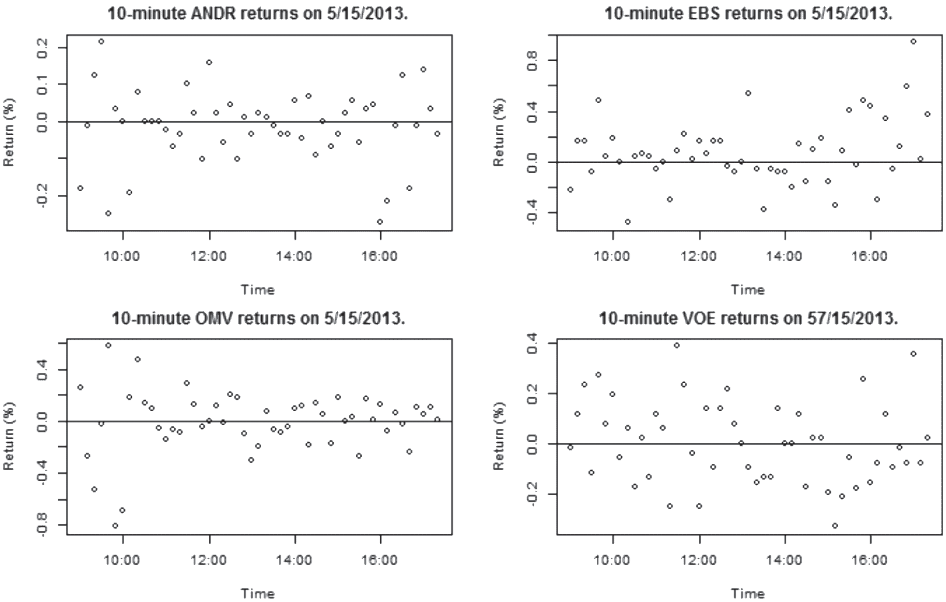

Figure 6. 10-minute percentile returns of ANDR, EBS, OMV, and VOE on May 15, 2013

It is clear that descriptive statistics of the returns may depend on the sampling frequency, stock market, stock, and particular day under study. We include descriptive statistics for 1-, 5-, and 10-minute returns of the four assets from each of the WSE and VSE. These statistics were computed independently on each day; therefore, we obtained 182 numbers in the case of the WSE and 248 numbers in the case of the VSE for each statistic and for each stock. The included boxplots illustrate the results and show variability across sampling frequencies, stocks, and days.

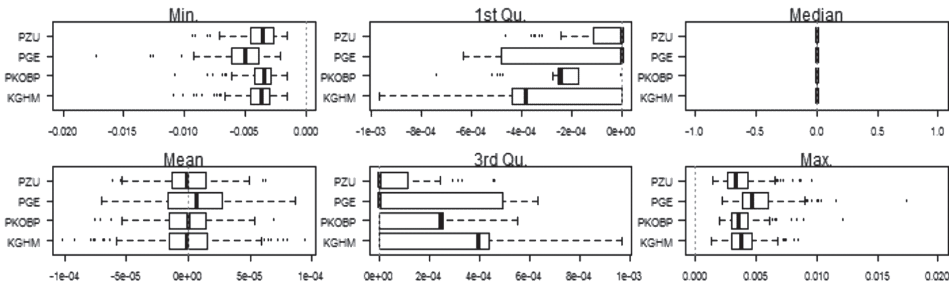

Figure 7. Descriptive statistics for 1-minute percentile returns for every trading day during 1/1/2014-9/22/2014 for WSE stocks. For minimum and maximum, the two extremely outlying observations have not been presented

Figures 7, 8, and 9 stand for six descriptive statistics (minimum, first quartile, median, mean, third quartile, and maximum) computed for the intraday returns of each Polish stock and every day. Since the dataset includes 182 trading days, each boxplot is constructed using 182 numbers. Analogously, Figures 10, 11, and 12 are prepared for Austrian stocks.

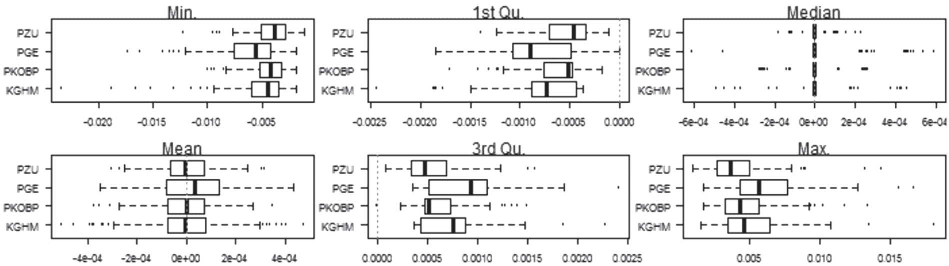

Figure 8. Descriptive statistics for 5-minute percentile returns for every trading day during 1/1/2014-9/22/2014 for WSE stocks

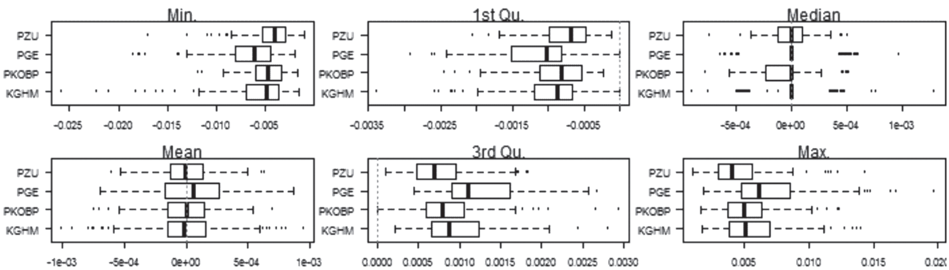

Figure 9. Descriptive statistics for 10-minute percentile returns for every trading day during 1/1/2014-9/22/2014 for WSE stocks

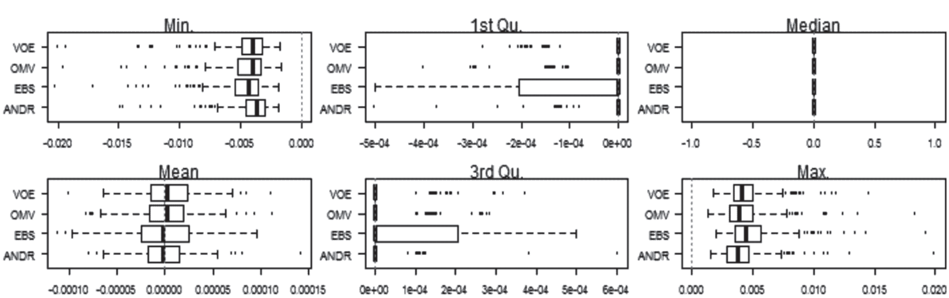

Figure 10. Descriptive statistics for 1-minute percentile returns for every trading day during 1/1/2013-12/31/2013 for VSE stocks

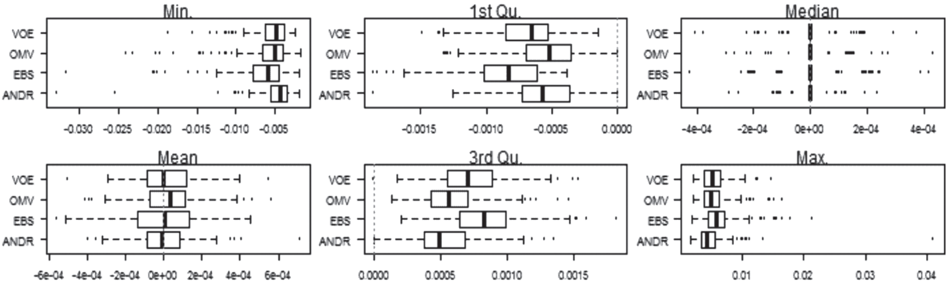

Figure 11. Descriptive statistics for 5-minute percentile returns for every trading day during 1/1/2013-12/31/2013 for VSE stocks

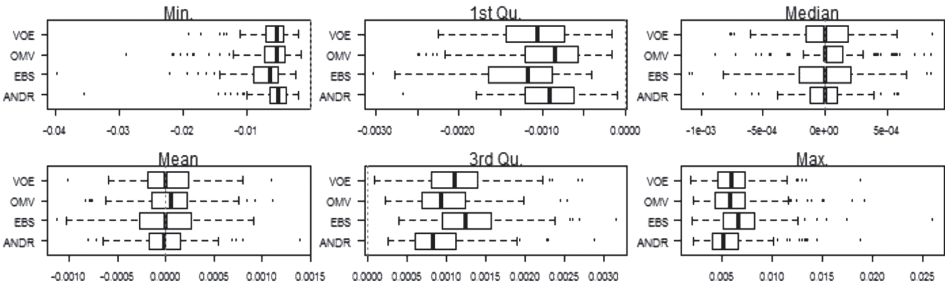

Figure 12. Descriptive statistics for 10-minute percentile returns for every trading day during 1/1/2013-12/31/2013 for VSE stocks

5. Epps effect

In this part of the paper, we briefly outline the model of compensation in the correlation estimation for the asynchronous trading. Moreover, we include the visualization of the Epps effect for companies from both markets under study. In the paper, the impact of asynchronous trading times on correlation is demonstrated. As explained in an earlier section of the paper, the series of returns from intervals of a certain length is, in fact, the series of returns from intervals between the time of the last trade before the beginning of the original interval to the time of the last trade before the original interval ends. This is the reason why the observed correlation is lower than the correlation between unobserved real-price series.

The method for compensation is based on the following theorem (which is proven in Gurgul and Machno, 2015):

Notation:

Let, for $i = 1,2$, the set $\left\{ r_{\Delta t}^{i}\left(t \right) \right\}_{t=0,\Delta t,\ldots,T}$ be the observed return's time series for asset i on time interval length Δt; let $\left\{ \tilde{r}^{i}\left(t \right) \right\}_{t=0,\Delta\tilde{t},\ldots,T}$ be the underlying true return's time series, where $\Delta\tilde{t}$ is a unit of time. Let $\tilde{g}^{i}(t)$ and $g_{\Delta t}^{i}(t)$ be the normalized returns defined in (2) and (3); $N_{\Delta t}^{i}(t)$ is the number of $\Delta\tilde{t}$ intervals in interval $[\gamma^{i}(t), \gamma_{+}^{i}(t)]$, where $\gamma^{i}(t)$ is the last trade of asset i before time t:

$$\gamma_{+}^{i}\left(t \right) := \min_{\substack{j=1,2,\ldots \\ \gamma^{i}(t+j\Delta t) \neq \gamma^{i}(t)}} \left\{ \gamma^{i}\left(t + \mathrm{j}\Delta t \right) \right\}.$$

If the following assumptions hold:

A1. Series $N_{\Delta t}^{i_1}(t)$ are independent of each other, and of series $\tilde{r}^{i_2}(t)$, for $i_1 \neq i_2$

A2. Variables $\tilde{r}^{1}(t)$ and $\tilde{r}^{2}\left(t + t_1 \right)$ are independent, for $t_1 \neq 0$

then the corrected correlation (the correlation between real unobserved series) is calculated as:

$$\operatorname{corr}_c\left(r_{\Delta t}^{1}, r_{\Delta t}^{2} \right) := \operatorname{corr}\left(\tilde{g}^{1}, \tilde{g}^{2} \right) = \frac{1}{T}\sum_{j=1}^{T} g_{\Delta t}^{1}\left(t \right) g_{\Delta t}^{2}\left(t \right) \frac{\sqrt{\left\langle N_{\Delta t}^{1} \right\rangle \left\langle N_{\Delta t}^{2} \right\rangle}}{\bar{N}\left(t_j \right)} \tag{1}$$

where $\langle\ldots\rangle$ means the average over analyzed period [0, T] and $\bar{N}(t)$ is the number of $\Delta\tilde{t}$ intervals in interval $\left[\max\left\{ \gamma^{1}\left(t \right), \gamma^{2}\left(t \right) \right\}; \min\left\{ \gamma^{1}\left(t + \Delta t \right), \gamma^{2}\left(t + \Delta t \right) \right\} \right]$.

In order to check the existence and properties of the Epps Effect on both stock markets under study, we transformed tick data into intraday data with intervals between 1 and 30 minutes with frequency by 1 second. It is possible to transform tick-by-tick data into intraday data of narrower intervals. However, this time series with very narrow intervals would not be more informative. Moreover, it would not be easy to conduct from the point of view of software capacity. Of course, the transformed time series is longer for narrower time intervals; for example, there are more 1-minute returns than 10-minute returns during the trading day.

The Epps effect takes place among all pairs of stocks listed on the WSE. Figures 13 and 14 show averaged correlations between the analyzed stocks and averaged corrected correlations for stocks from the WSE and VSE, respectively. The adjusted values corresponding to realized (not corrected) correlations are calculated as follows: for each trading day out of 182 days (in the case of the WSE) and 248 days (VSE) under study, we converted data into intraday data with the given interval. The length of the interval is marked on the x-coordinate of the point. In the next step, we computed correlations between returns of stocks belonging to our sample on each day. The average of these 182 correlations (or 248, in the case of the VSE) constitutes the y-coordinate of the particular point.

The value corresponding to the corrected correlation is obtained analogously. Instead of correlations between returns, values calculated using formula (1) are averaged.

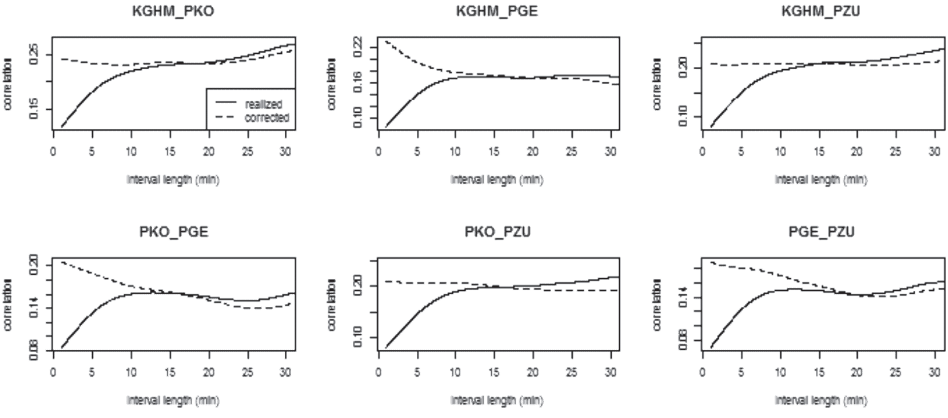

Figure 13. Realized and corrected correlations between stock returns according to data interval length for WSE stocks

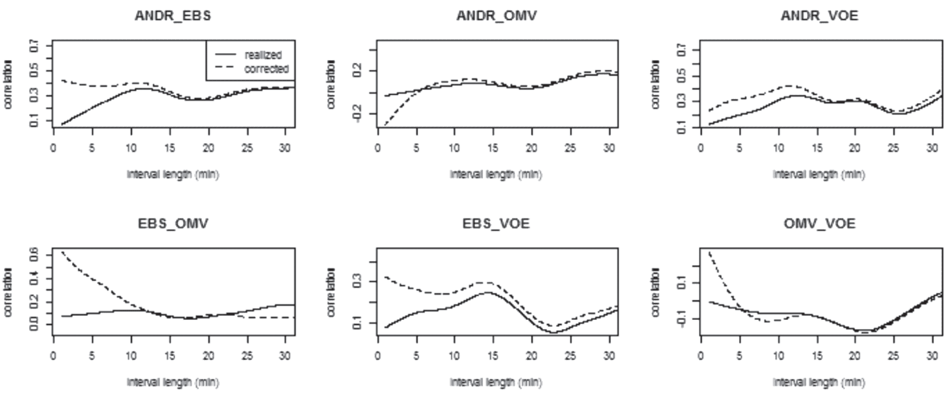

Figure 14. Realized and corrected correlations between stock returns according to data interval length for VSE stocks

Analyzing Figures 13 and 14, we make an interesting observation. In the case of the WSE, the correction clearly works; the Epps effect is not visible after correction. However, it appears that the corrected correlations might exhibit an opposite property. The corrected correlation becomes lower for wider intervals in some cases. From a theoretical point of view, the corrected correlation should be constant under the assumptions of the presented theorem. It is true that the assumptions are not fully met; however, one could expect that, even after the correction, a smaller Epps effect caused by other factors (such as lead-lag correlation or the discreteness of the data) is still present. Another surprising observation is the lower value of corrected correlation than the realized correlation for wider intervals.

The observed correlation and corrected correlation for the VSE is much more surprising, especially as compared to the corresponding results for the WSE. The correlations between Austrian stocks do not possess a typical pattern. We observe that the VOE and the OMV are basically uncorrelated on the intraday level; thus, the analysis does not make much sense. The ANDR and EBS constitute the most-correlated pair. We observe that, in the case of this pair, the patterns of the correlation and the corrected correlation are similar to pairs from the WSE.

6. Conclusions

The correlations between high-frequency financial data are strongly affected by asynchronous trading. Contrary to widespread beliefs, our study suggests that asynchronous trading may be the only essential cause of the Epps effect in the data under study on the WSE. However, the presence of the Epps effect on the VSE is not clearly pronounced. The Polish stocks under study are traded on one stock exchange (no dual listing) and are among the most-liquid assets in Warsaw. In correlations with the asynchronous effect removed, the Epps effect is no more visible.

The analysis does not show uniform properties among Austrian assets. For the most-liquid and most-correlated pair (namely, ANDR-EBS), the analysis shows similar results as for Polish stocks. However, for less-liquid and less-correlated assets, the correlation pattern itself does not exhibit a clear Epps effect. Moreover, the suggested correction seems to not work, especially for the ANDR-OMV pair.

The hypothesis that there are others significant sources of the Epps effect on the WSE seems to be false. However, on the VSE, the existence and sources of the Epps effect is not clear; therefore, additional analysis is needed.

In addition (in the case of the WSE), the relationship between the corrected correlation and sampling frequency shows an unexpected pattern. First of all, the corrected correlation is lower for wide intervals (wider than 20-minute) for all pairs from the sample. The second unexpected observation is that, in some corrected correlation patterns, evidence for the reciprocal Epps effect was found (i.e., the corrected correlation decreases as the sampling frequency of data decreases).

In the case of the VSE, the correlation does not increase when the sampling frequency decreases in some pairs and intervals (indicating a lack of the Epps effect). There is no visible pattern for the corrected correlations; for instance, the correction seems to overestimate the correlation for the OMV-VOE pair for high-frequency data (left side of the graph in Figure 14). However, for the ANDR-OMV pair, the correction seems to underestimate the correlation. Moreover, corrected correlations are not lower than observed correlations for less-frequent data, which is the case on the WSE.

The main question in the context of this analysis is the connection between the intensity of trading and the Epps effect. From a theoretical point of view, the impact of asynchronous trading on the Epps effect should be stronger for less-liquid assets. However, the conjecture that the Epps effect is stronger for less-liquid pairs (the Austrian assets are less liquid than Polish ones) is not confirmed in this study. The addressed topics need further research based on the recent data.

References

[1] Conlon, T., Ruskin, H.J. and Crane, M. (2009) 'Multiscaled Cross-Correlation Dynamics In Financial Time-Series, *Physica A* 388(5), pp. 705–714.

[2] Epps, T.W. (1979) 'Comovements in stock prices in the very short run', *Journal of the American Statistical Association*, vol. 74, pp. 291–298.

[3] Gurgul, H. and Machno, A. (2015) 'The Impact of Asynchronous Trading on Epps Effect on Warsaw Stock Exchange', *Central European Journal of Operational Research*, forthcoming, DOI: 10.1007/s10100-016-0442-y.

[4] Lundin, M., Dacorogna, M. and Muller, U.A. (1998) 'Correlation of High Frequency Financial Time Series', in Lequeux P. (ed.) Financial Markets Tick by Tick, Wilney and Sons.

[5] Mastromatteo, I., Marsili, M. and Zoi, P. (2011) 'Financial correlations at ultra-high frequency: theoretical models and empirical estimation', *The European Physical Journal B*, vol. 80, pp. 243–253.

[6] Munnix, M.C., Schafer, R. and Guhr, T. (2010) 'Compensating asynchrony effects in the calculation of financial correlations', *Physica A* 389, pp. 767–779.

[7] Toth, B. and Kertesz, J. (2009) 'The Epps effect revisited', *Quantitative Finance*, vol. 9, pp. 793–802.

Managerial Economics
2016, vol. 17, no. 1, pp. 77–92
http://dx.doi.org/10.7494/manage.2016.17.1.77

Henryk Gurgul*, Robert Syrek**

The logarithmic ACD model: The microstructure of the German and Polish stock markets[1]

1. Introduction

The timing of transactions (i.e., the quantity purchased in a period of time) is often a key economic variable; as such, it should be modeled or forecasted. The microstructure of financial markets is investigated using transaction-by-transaction data. The timing of these transactions can be very important in order to understand market participant behavior from the point of view of economic theory.

In recent years, there have been a lot of contributions dealing with the financial market microstructure. These contributions have focused both on theoretical models and empirical findings. Nowadays, most exchanges (NYSE, NASDAQ, Paris Bourse, Frankfurter Boerse) and even smaller exchanges like Vienna and Warsaw compile databases of tick-by-tick data which, depending on the exchange, give information on the trade process (time of the trade, price, volume) and the bid-ask quote process (time of quotes, bid and ask quotes, depths) or the state of the order book. Researchers can now work in new empirical and theoretical areas. However, in order to use high-frequency data, new econometric tools are necessary.

High-frequency data about investor activity typically arrives at irregular time intervals. Classic standard econometric techniques were good for the treatment of fixed time intervals. Typically, researchers aggregated market data to some fixed time intervals. The most-frequently used data in the case of consumption was monthly or yearly data. However, stock market transactions are very often

* AGH University of Science and Technology, Faculty of Management, Department of Applications of Mathematics in Economics, e-mail: henryk.gurgul@gmail.com

** Jagiellonian University in Kraków, Institute of Economics, Finance and Management, e-mail: robert.syrek@uj.edu.pl

[1] Financial support for this paper from the National Science Center of Poland (Research Grant DEC-2012/05/B/HS4/00810) is gratefully acknowledged by Henryk Gurgul.

conducted within a fraction of a second. It is plausible that, in the case of a short-time interval, one can observe many intervals with no new information, so the data becomes heteroskedastic. On the other hand, we cannot analyze the micro-structure properties of the data in the case of a long interval. The consequence of multiple transactions is the averaging of the timing and characteristics of particular transactions, meaning that the researcher can lose certain valuable information.

The problems of market microstructure are complex, because the rate of arrival of transactions may exhibit calendar or seasonal effects (dependent on the day, week, or year), which implies problems with the determination of an appropriate length-of-time interval. The frequency may depend not only on a news release but also sometimes on an unobservable factor that is to some extent deterministic but also partly stochastic. In this case, one says that there is a stochastic process underlying trading activity. The breaking point in research into the microstructure of financial markets was provided by Engle and Russel (1997) and (1998).

In our contribution, we try to compare the microstructure of selected stocks comprised from the DAX30 and WIG20 using models of the ACD type under the assumption of generalized gamma and Burr distributions of durations.

The remaining part of our paper is scheduled as follows: Section 2 contains a short overview of existing models in the literature and the methods used in studies of microstructure. In Section 3, the methodology is presented; and in Section 4, the results of our computations are shown and discussed in detail. Finally, Section 5 presents the conclusions of the paper.

In the next section, we conduct a brief literature review on the topic.

2. Literature overview

Early contributors on market microstructure (for example Hafner, 1996; Eddelbutel and McCurdy, 1998; and Guillaume et al., 1997) concentrated on high-frequency data modeled by the so-called *"fixed interval"* econometric models. These models include the stochastic volatility- and GARCH-type models. The common feature of this kind of modeling is that the data is regularly sampled at a very high frequency. However, one main drawback of these models is that they do not take into account the irregular spacing of data. The contributors tried to conduct an alternative to regular sampling, using time transformation techniques. The aim was to transform irregularly spaced data into fixed-interval data.

Engle and Russel (1997) and (1998) treated arrival times as random variables that follow a stochastic point process. Arrival time depends on (financial) random variables such as volume, bid ask spread, or price. The authors derived a new model for dependent-point processes. The set of parameters typical of the stochastic process consisted of past events in order to reflect the transaction process. The most-important application of the model was to the measurement and forecasting of the intensity of transaction arrivals. The contributors parameterized the conditional intensity as a function of the time between past events. In addition, they allowed some natural extensions; e.g., taking into account the characteristics of past transactions. Also, outside influence cannot be excluded. The authors, assuming the dependence of the conditional intensity on past durations, called their model the *Autoregressive Conditional Duration* (ACD) model. The ACD model is the counterpart of the GARCH model.

This model for the durations between two successive market events (such as the buying or selling of a security) takes into account a clustering effect in the durations. In the model, short (long) durations tend to be followed by short (long) durations. This effect resembles that which is found in the volatility of many financial time series. The contributors applied the model to the modeling of the foreign exchange market and the IBM stock. In a subsequent paper, Engle (2000) linked the ACD-duration model with a GARCH model for returns. This combination allowed the modeling of irregularly timed data. The following extensions or counterparts were developed by a number of contributors.

Ghysels and Jasiak (1998) suggested the stochastic volatility duration (SVD) model. Gramming and Maurer (2000) derived an ACD model based on the Burr distribution (an alternative to the Weibull distribution). Jasiak (1998) extended ACD to the fractionally integrated ACD model. This model enables the modelling of long-range dependence in the durations.

Bauwens and Veredas (2004) defined a class of models for the analysis of durations, known as stochastic conditional duration (SCD) models. These models are based on the assumption that the durations are generated by a dynamic stochastic latent variable. The model yields a wide range of shapes hazard functions. The estimation of the parameters is conducted by quasi-maximum likelihood and by using the Kalman filter. The model is applied to trade, price, and volume durations of stocks traded at the NYSE. The authors also investigate the relation between price durations, spread, trade intensity, and trading volume.

The most-recent research tries to reflect the new realities on financial markets implied by the introduction of new technology and high-frequency trading, especially after the crisis of 2008.

Based on order-level data from 2008, Hasbrouck and Saar (2013) found that some traders on the NASDAQ could respond to events such as changes in the limit order book in 2–3 milliseconds. Ye et al. (2013) proved that trading could be conducted at even faster speeds. They suggested that high-frequency traders' need for speed depends on the particular strategies that they follow.

According to Jones (2012), Brogaard et al. (2012), and Carrion (2013), most market participants and researchers are convinced that HFT market trading enhances market quality by reducing spreads and raising informational efficiency. Some contributions (e.g., Kirilenko et al., 2011; Easley et al., 2011; 2012a, and Madhavan, 2013) express concerns that HFT market trading can induce market instability. However, Brogaard et al. (2011), O'Hara (2011), and O'Hara et al. (2013) stress that the bulk of liquidity provision in many markets is provided by high-frequency traders.

Angel et al. (2011) stress that, when retail orders do go to the NYSE, they often benefit from liquidity provided by DMMs (designated marker makers) and SLPs (strategic liquidity providers), many of which are actually high-frequency trading firms. Another observation is that the trading costs of retail traders have been falling over the past 30 years. In addition, this decline has sped up in recent years. Using data from the Toronto Stock Exchange, Malinova and Park (2013) show empirically that retail trading costs have fallen because of the presence of HFTs. Hendershott et al. (2011) provide evidence that algorithmic trading particularly improved market quality with respect to improved liquidity and the enhanced informativeness of quotes. Boehmer et al. (2015) supported this prediction on the basis of data from 39 markets.

HFTs are looking for the fastest way to trade. Technological innovations are key factors in making a decision to trade or not. For exchanges and markets, provision of these innovations is a key factor in their competitiveness (and survival). The profitability of the new fast technology of trading is discussed in Brogaard et al. (2014), Cespa and Vives (2013), Pagnotta and Philippon (2011), and Biais et al. (2015). According to Laughlin et al. (2014), the speeding up of communication to about three milliseconds between Chicago and New York markets increased costs significantly. Haldane (2011) stresses the role of the speed of high-frequency trading in these words: "Adverse selection today has taken on a different shape. In a high speed, co-located world, being informed means seeing and acting on market prices sooner than competitors. Today, it pays to be faster than the average bear, not smarter. To be uninformed is to be slow".

Hasbrouck and Saar (2009) underline that technology allows orders to be submitted (and cancelled) instantaneously. The optimum strategies use this option in order to apply complex trading strategies.

One of the most-important questions is whether one can actually link "buy" and "sell" trades with upcoming information (Easley et al., 2012a, 2012b, 2013). In their opinion, the active side of the trade is oriented more to the spread than the actual content of the released information.

O'Hara (2015) stresses that a fundamental change in how traders trade and how markets operate can be observed in recent years. In her opinion, the high-frequency algorithms operate across the market and use the power of technology to forecast price movements of securities. The forecasts take into account the behavior of correlated assets. Thus, the starting point of empirical analyses should be to assess the predictive power of market variables, both within and across markets. The main focus in the future should be oriented towards understanding the changing nature of the market, including understanding the changing nature of market data.

More complete surveys of HFT topics may be found in reviews by Biais and Wooley (2011), Angel et al. (2011), Jones (2012), and Goldstein et al. (2014).

3. Methodology

We shall consider the dynamic parametrization of the conditional mean function (Engle and Russell, 1998):

$$\psi_i := \psi_i(\theta) = E[x_i \mid \mathcal{F}_i; \theta]$$

where $\mathcal{F}_i$ denotes the information set up to observation t_{i-1} (beginning of i-th duration x_i between two events occurred at times t_{i-1} and t_{i-1}) and θ is the vector of parameters.

It is assumed that standardized durations:

$$\varepsilon_i = \frac{x_i}{\psi_i}$$

are independent and identically distributed random variables with $E[\varepsilon_i] = 1$. Variation in autoregressive conditional duration models arise from different choices of functional form for the conditional mean function and choices of distribution of standardized durations.

The most basic specification assumes linear parametrization of the conditional mean function (Engle and Russell, 1998):

$$\psi_i = \omega + \sum_{j=1}^{P} \alpha_j x_{i-j} + \sum_{j=1}^{Q} \beta_j \psi_{i-j}$$

where $\omega > 0$, $\alpha_j \geq 0$, $\beta_j \geq 0$ for all j and $\sum_{j=1}^{P} \alpha_j + \sum_{j=1}^{Q} \beta_j < 1$. The first three constraints ensure that conditional durations are positive, whereas the last inequality ensures the existence of an unconditional mean of duration. Bauwens and Giot (2000) propose two extensions of the linear ACD model. Models called logarithmic ACD are of the forms:

$$\ln\psi_i = \omega + \sum_{j=1}^{P}\alpha_j \ln\varepsilon_{i-j} + \sum_{j=1}^{Q}\beta_j\psi_{i-j}$$

and

$$\ln\psi_i = \omega + \sum_{j=1}^{P}\alpha_j\varepsilon_{i-j} + \sum_{j=1}^{Q}\beta_j\psi_{i-j}$$

We refer to the different specifications as $LACD_1$ and $LACD_2$ respectively. There are no sign restrictions on parameters to ensure the positivity of conditional duration.

For each cited specification, researchers have to choose a distribution for standardized durations. In their seminal paper, Engle and Russel (1998) study exponential and Weibull distributions (it is worth mentioning that the former is used in quasi maximum likelihood estimation).

In this paper, we try to fit generalized gamma and Burr distributions. The density of the generalized gamma distribution (Lunde, 2000) is given as:

$$f(\varepsilon) = \frac{\gamma\varepsilon^{\kappa\gamma-1}}{\theta^{\kappa\gamma}\Gamma(\kappa)}\exp\left\{-\left(\frac{\varepsilon}{\theta}\right)^{\gamma}\right\}$$

with $\theta = \dfrac{\Gamma(\kappa)}{\Gamma\left(\kappa + \frac{1}{\gamma}\right)}$ and $\kappa, \gamma > 0$. The generalized gamma distribution includes the Weibull distribution for $\kappa = \gamma = 1$ and the exponential distribution if $\kappa = 1$.

Gramming and Maurer (2000) examine the Burr distribution, whose density is as follows:

$$f(\varepsilon) = \frac{\theta\kappa\varepsilon^{\kappa-1}}{\left(1+\sigma^2\theta\varepsilon\kappa\right)^{\frac{1}{\sigma^2}+1}}$$

where $\theta = \sigma^{2\left(1+\frac{1}{\kappa}\right)}\dfrac{\Gamma\left(\frac{1}{\sigma^2}+1\right)}{\Gamma\left(\frac{1}{\kappa}+1\right)\Gamma\left(\frac{1}{\sigma^2}-\frac{1}{\kappa}\right)}$ with $\sigma > 0$ and κ such that $-\kappa < 1 < \dfrac{\kappa}{\sigma^2}$.

We get the Weibull distribution if $\sigma \to 0$ and exponential distribution if, additionally, $\kappa = 1$.

A concept that is often used in duration analysis is the hazard function. Assuming that duration X is a continuous random variable, the hazard function is defined as (see, for example, Bauwens and Giot, 2001):

$$h(x) = \lim_{dx \to 0} \frac{P\left[x \le X < x + dx \mid X \ge x\right]}{dx}$$

In the formula above, the numerator is the probability that the event occurs in the interval $[x, x + dx)$, given that it has not occurred before, while the denominator is the width of the interval. The fraction represents the rate of event occurrence per second. Taking the limit as dx goes to 0 we obtain an instantaneous rate of occurrence. This leads to an alternative definition of hazard function of the form:

$$h(x) = \frac{f(x)}{S(x)} = \frac{-dS(x)}{dx}$$

Where $f(x)$ and $S(x)$ are the density and survival functions of random variable X, respectively. It can be shown that the hazard functions of an exponential distribution is flat, while the hazard function of Weibull can be either flat or monotone (increasing and decreasing), depending on the values of the parameters. In the case of Burr and the generalized gamma, the hazard function can take many different shapes (including non-monotone cases), depending on parameter values.

4. Empirical results

We consider tick-by-tick transactions of DAX30 and WIG20 companies. The first dataset contains the prices of 25 companies from 2013.03.22 to 2013.05.17 (37 trading days). The second dataset contains observations from 2013.09.02 to 2013.10.18 (35 trading days) for 12 actively traded Polish companies. To compute price durations, one needs to set up a price threshold. Taking into account the large number of time series, it is not possible to choose one universal price threshold. In addition, the tick sizes of companies are very different (based on upper and lower price bands). To avoid incomparable either short or long duration series we set a price threshold based on the tick size of each company. For most DAX30 companies we set the price threshold as 0.1 euro (which equals a 20 tick size) and different values for Polish companies (this is the result of huge differences in the number of trades and scale of prices). We remove overnight durations and durations corresponding to events recorded outside regular opening hours (9:00 to 17:30 for German and 9:30 to 16:50 for Polish companies).

In Table 1 we present the main descriptive statistics of plain-price durations along with the result of Ljung-Box test with 15 lags.

Table 1

Descriptive statistics of raw-price durations (number of observations (*N*), mean, standard deviation, minimum, quantiles, maximum, and Ljung-Box test statistic)

DAX30					
Statistic	**Min**	**0.25q**	**Median**	**0.75q**	**Max**
N	1626	2772	3071	3570	5802
Mean	198.72	320.89	372.62	410.55	695.58
S.D.	267.2	440.4	537.16	613.92	955.03
Min	1	1	1	1	1
0.25q	36	63	69	83	147
Median	106.5	175.5	196	216.5	381.5
0.75q	252	394	474.25	518	873.75
Max	2778	5194	6332	8514	22975
LB(15)	326.31	672.72	776.62	1111.87	2565.88
WIG20					
Statistic	**Min**	**0.25q**	**Median**	**0.75q**	**Max**
N	1544	1975.75	2316	3382	4470
Mean	218.75	291.68	418.28	491.43	623.31
S.D.	426.03	627.67	894.34	1116.63	1513.92
min	1	1	1	1	1
0.25q	16	17.5	23	35.625	57
Median	65	80	102.5	133.75	226
0.75q	229	280.38	364.63	440.31	674.25
Max	6537	9434.75	12068.5	16952.5	21037
LB(15)	72.23	193.86	293.73	613.96	1507.94

The computation results confirm the stylized facts observed in the duration data. There is an overdispersion and autocorrelation in all series under study. On average, Polish companies exhibit weaker serial correlation yet the highest overdispersion. In both cases, the series of price durations exhibit a diurnal pattern (which may be different for each day of the week). In a way similar to Bauwens and Giot (2000), we take into account the time of day and the day of the week. We use cubic splines with nodes set every 60 minutes with two additional nodes: 10 minutes after the market opens and 10 minutes before the end of the session.

The table 2 presents the descriptive statistics of diurnally adjusted price durations (plain durations divided by seasonal component).

Table 2

Descriptive statistics of adjusted price durations (number of observations, mean, standard deviation, minimum, quantiles, maximum, and Ljung-Box test statistic)

DAX30					
Statistic	**Min**	**0.25q**	**Median**	**0.75q**	**Max**
N	1626	2772	3071	3749,25	5802
Mean	1.017	1.027	1.036	1.042	1.064
S.D.	1.096	1.186	1.214	1.248	1.634
min	0.001	0.001	0.002	0.002	0.003
0.25q	0.194	0.241	0.258	0.272	0.299
Median	0.521	0.622	0.636	0.653	0.7
0.75q	1.193	1.347	1.364	1.375	1.422
Max	8.817	11.74	13.689	15.131	25.356
LB(15)	257.79	410.408	458.694	667.238	3730.499
WIG20					
Statistic	**Min**	**0.25q**	**Median**	**0.75q**	**Max**
N	1544	1975.75	2316	3382	4470
Mean	1.01	1.022	1.034	1.042	1.08
S.D.	1.745	1.928	2.005	2.121	2.42
min	0.001	0.001	0.001	0.002	0.003
0.25q	0.051	0.065	0.087	0.11	0.14
Median	0.253	0.306	0.35	0.408	0.459
0.75q	0.983	1.104	1.152	1.215	1.292
Max	19.666	28.698	32.101	36.076	70.971
LB(15)	107.421	112.962	357.678	511.005	1224.436

With this procedure, the mean of adjusted durations is close to 1. It can be seen that seasonal adjustment reduces overdispersion and autocorrelation.

We estimate (by the maximum likelihood estimation method) models being combinations of the parametrization of conditional mean functions and distributions (with a total number of 12 different models, the models are restricted with lag order of $P = Q = 1$)[2]. We select a model that best fits in several ways. First, we restrict our attention to models that "remove" autocorrelation (we apply the Ljung-Box test to residuals and their squares). Denoting by $f_i(x_i \mid \mathcal{F}_i)$ the sequence of one-step-ahead density forecasts, we calculate the probability integral transform PIT (Diebold et. al, 1998):

$$z_i = \int_{-\infty}^{x_i} f(u)\,du$$

and apply the Anderson-Darling and Cramer von Mises GOF test. Finally, we choose a model associated with the smallest BIC and "significant" parameters (details of estimation results are available from the authors upon request). In Tables 3 and 4, we present the estimation results and the testing for "best" autoregressive conditional duration models for all companies under study.

Restriction $P = Q = 1$ in all specifications is sufficient to describe clustering in price duration series. It follows from the values of the Ljung-Box test statistics for residuals (with lowest p-values of 0.09 [Thyssen] and 0.14 [JSW]), and the results of this test applied to squared residuals (lowest p-values: 0.06 [Allianz] and 0.73 [KGHM]). In most cases, $LACD_1$ models fit the best. The $LACD_1$ model implies that the concave news impact curve (relationship between ε_{i-1} and x_i) is asymmetric and that the difference in the impact of innovations with $\varepsilon_i < 1$ is larger than with $\varepsilon_i > 1$ (Hautsch, 2003). In one case (BASF), the outcome is the opposite. Sum $\alpha + \beta$ for linear parametrization (four cases) equals at least 0.93 and confirms the clustering of durations (this implies a slowly decreasing autocorrelation function). For logarithmic parametrization, clustering increases with parameter β. Comparing results (using quartiles for parameters β), we can conclude that these values are a little higher for Polish companies.

Given the results of Anderson-Darling and Cramer von Mises testing, we conclude that the assumed distributions are correct, with a strong rejection of both exponential and Weibull distribution.

In about two-thirds of the cases, the Burr distribution fits better than the generalized gamma distribution.

Regarding the parameters of these distributions, in only two cases (Linde and Kernel) is the hazard function monotone (decreasing starting at ∞). In all of the remaining cases, the hazard function has an inverted-U shape. Analyzing parameters of generalized gamma distributions, we find one case that corresponds to a U-shaped hazard function (PGNiG), starting at ∞, and tending to ∞ as ε tends to ∞ (Bauwens and Giot, 2001). In all of the remaining cases, the hazard function has an inverted-U shape.

[2] We use R environment and package ACDm for all computations (https://cran.r-project.org/web/packages/ACDm/index.html)

Table 3

Estimation and testing results for DAX30 companies

Company	Model	Dist.	LB(15)	LB2(15)	A-D	CvM	ω	α	β	κ	μ
Adidas	LACD1	Burr	12.66	7.09	0.60	0.59	0.09	0.16	0.92	1.13	0.29
Allianz	LACD1	Burr	16.87	24.12	0.91	0.95	0.09	0.16	0.86	1.07	0.19
BASF	LACD2	GG	19.88	20.32	0.12	0.17	–0.16	0.16	0.94	1.48	0.75
BAYER	LACD1	Burr	16.68	10.55	0.89	0.86	0.11	0.20	0.89	1.15	0.27
Beiersdorf	LACD1	GG	15.98	7.79	0.11	0.13	0.07	0.12	0.91	1.28	0.81
BMW	LACD1	GG	6.11	2.94	0.80	0.84	0.09	0.17	0.85	2.54	0.59
Continental	ACD	Burr	18.54	4.75	0.92	0.92	0.07	0.17	0.76	1.05	0.14
Daimler	LACD1	Burr	11.80	4.65	0.86	0.79	0.12	0.23	0.77	1.13	0.22
Deutsche Bank	LACD1	Burr	13.23	16.19	0.50	0.59	0.09	0.22	0.84	1.23	0.23
Deutsche Börse	LACD1	GG	11.22	7.63	0.08	0.10	0.08	0.14	0.87	1.48	0.74
Deutsche Post	LACD1	Burr	7.66	1.43	0.80	0.86	0.10	0.17	0.85	1.08	0.21
Deutsche Telekom	LACD1	Burr	13.21	9.36	0.74	0.70	0.10	0.19	0.90	1.13	0.24
Deutsche_Lufthansa	LACD1	GG	17.58	2.79	0.17	0.20	0.10	0.18	0.87	2.20	0.61
EON	LACD1	GG	19.74	6.45	0.50	0.43	0.11	0.19	0.95	3.55	0.46
Fresenius SE	LACD1	GG	21.51	16.56	0.19	0.20	0.08	0.13	0.88	1.23	0.84
Fresenius Medical Care	LACD1	Burr	19.07	5.29	0.68	0.63	0.09	0.15	0.93	1.06	0.21
Heidelberg Cement	LACD1	GG	22.56	13.59	0.15	0.20	0.06	0.12	0.90	1.41	0.79
Henkel	LACD1	Burr	13.57	5.22	0.76	0.83	0.10	0.18	0.85	1.13	0.25
Infineon	LACD1	Burr	15.31	7.03	0.79	0.76	0.13	0.23	0.80	1.09	0.25
K&S	ACD	Burr	15.90	12.03	0.77	0.73	0.07	0.20	0.73	1.07	0.19
Lanxess	LACD1	Burr	10.11	14.50	0.76	0.75	0.10	0.18	0.86	1.09	0.25

Table 3 cont.

Company	Model	Dist.	LB(15)	LB2(15)	A-D	CvM	ω	α	β	κ	μ
Linde	LACD1	Burr	20.27	10.50	0.87	0.97	0.09	0.13	0.91	0.97	0.16
SAP	LACD1	Burr	16.82	4.31	0.96	0.92	0.13	0.24	0.89	1.13	0.28
Siemens	LACD1	Burr	14.49	5.42	0.22	0.22	0.11	0.19	0.91	1.11	0.25
Thyssen	LACD1	Burr	22.60	3.76	0.45	0.40	0.08	0.14	0.92	1.10	0.26

LB(15) denotes the value of the Ljung-Box test statistics applied to residuals, while LB2(15) that applied to squared residuals, A-D and CvM are p-values in GOF testing, parameter μ refers to γ for generalized gamma distribution, and σ^2 for Burr distribution.

Table 4

Estimation and testing results for WIG20 companies

Company	Model	dist.	LB(15)	LB2(15)	A-D	CvM	ω	α	β	κ	μ
PKOBP	LACD1	GG	20.77	1.23	0.62	0.51	–0.17	0.16	0.92	3.42	0.36
PZU	LACD1	GG	4.89	0.49	0.11	0.10	0.35	0.29	0.85	15.02	0.15
KGHM	LACD1	GG	20.94	11.33	0.22	0.32	0.15	0.19	0.91	3.54	0.39
PEKAO	LACD1	GG	20.61	4.44	0.47	0.50	0.20	0.21	0.91	3.29	0.36
PKNORLEN	LACD1	GG	12.66	0.41	0.12	0.08	0.15	0.14	0.93	3.81	0.32
PGE	LACD1	GG	17.88	0.67	0.10	0.10	0.28	0.27	0.84	5.63	0.26
PGNIG	ACD	GG	9.13	2.41	0.10	0.10	0.06	0.28	0.70	2.24	0.38
TPSA	LACD1	GG	15.43	4.72	0.09	0.12	0.23	0.20	0.89	4.39	0.28
ASSECOPOL	LACD1	GG	11.06	2.83	0.18	0.18	0.10	0.11	0.95	1.45	0.56
JSW	ACD	GG	20.98	5.76	0.14	0.13	0.03	0.12	0.86	2.87	0.41
KERNEL	LACD1	Burr	8.50	1.46	0.16	0.20	0.10	0.10	0.94	0.77	0.12
LOTOS	LACD1	GG	12.73	1.76	0.19	0.22	0.12	0.12	0.97	2.71	0.39

LB(15) denotes the value of the Ljung-Box test statistics applied to residuals, while LB2(15) that applied to squared residuals, A-D and CvM are p-values in GOF testing, parameter μ refers to γ for generalized gamma distribution, and σ^2 for Burr distribution.

5. Conclusions

To conclude, logarithmic ACD models are useful tools for describing transaction processes on the Frankfurt and Warsaw Stock Exchanges.

The conducted empirical analysis of raw price duration for selected companies listed on the DAX30 and WIG20 shows that, on average, Polish companies exhibit weaker serial correlation yet higher overdispersion than German companies. Both statistics depend on quantiles. For higher quantiles they become larger. This dependence is weaker in the case of adjusted price durations with respect to overdispersion, and is more visible in autocorrelation. In addition, the dependence of autocorrelation on the quantiles of raw data is much more pronounced in Polish than in German price durations.

The fitted ACD model for price durations for almost all companies listed on both stock markets under study is LACD1. While for most German companies on the DAX30, the Burr distribution fits better than generalized gamma distribution, the latter distribution fits well in the case of Polish blue chips. Analyzing series by hazard function, we note a similarity of hazard functions for companies from both markets, the functions in general displaying a U-shaped pattern.

References

[1] Angel, J., Harris, L. and Spatt, C. (2011) 'Trading the 21st century', *Quarterly Journal of Finance*, vol. 1 (1), pp. 1–53.

[2] Bauwens, L. and Giot, P. (2000) 'The logarithmic ACD model: an application to the bid-ask quote process of three NYSE stocks', *Annales d'Economie et de Statistique*, vol. 60, pp. 117–149.

[3] Bauwens, L. and Giot, P. (2001) Econometric Modelling of Stock Market Intraday Activity, Dordrecht: Springer Science, Business Media.

[4] Bauwens, L. and Veredas, D. (2004) 'The stochastic conditional duration model: a latent variable model for the analysis of financial durations', *Journal of Econometrics*, vol. 119, pp. 381–412.

[5] Biais, B. and Wooley, P. (2011) 'High Frequency Trading', Working Paper, Toulouse University, IDEI.

[6] Biais, B., Foucault, T. and Moinas, S. (2015) 'Equilibrium high-frequency trading', *Journal of Financial Economics*, vol. 116(2), pp. 292–313.

[7] Boehmer E., Fong K.Y.L. and Wu J. (2015) 'International evidence on algorithmic trading,' AFA 2013 San Diego Meetings Paper.

[8] Brogaard, J., Hendershott, T. and Riordan, R. (2011), 'High Frequency Trading and Price Discovery', Working Paper, [Online], Available: http://ssrn.com/abstract=1928510.

[9] Brogaard, J., Garriott, C. and Pomranets, A. (2012) 'Is more high-frequency trading better? ' Working Paper.

[10] Brogaard, J., Hendershott, T.J. and Riordan, R. (2014) ‘High-frequency trading and price discovery’, *Review of Financial Studies*, vol. 27(8), pp. 2267–2306.
[11] Carrion, A., (2013) ‘Very fast money: high-frequency trading on NASDAQ’, *Journal of Financial Markets*, vol. 16(4), pp. 680–711.
[12] Cespa, G. and Vives, X. (2013) ‘The Welfare Impact of High Frequency’, Working Paper.
[13] Diebold, F.X., Gunther, T.A. and Tay, A.S. (1998) ‘Evaluating density forecasts with applications to financial risk management’, *International Economic Review*, vol. 39, pp. 863–883.
[14] Easley, D., Lopez de Prado, M. and O’Hara, M. (2011) ‘The Microstructure of the «Flash Crash»’: Flow Toxicity, Liquidity Crashes, and the Probability of Informed Trading’, *Journal of Portfolio Management*, vol. 37, pp. 118–128.
[15] Easley, D., Lopez de Prado, M. and O’Hara, M. (2012a) ‘Flow Toxicity and Liquidity in a High Frequency World’, *Review of Financial Studies*, vol. 25, pp. 1457–1493.
[16] Easley, D., Lopez de Prado, M. and O’Hara, M. (2012b) ‘Bulk Classification of Trading Activity’, *SSRN Electronic Journal*, vol. 03, DOI: 10.2139/ssrn.1989555.
[17] Easley, D., Lopez de Prado, M. and O’Hara, M. (2013) ‘Optimal Execution Horizon’, *Mathematical Finance*, vol. 25(3), pp. 640–672.
[18] Eddelbutel, D., McCurdy, T.H. (1998) ‘The Impact of News on Foreign Exchange Rates: Evidence from very High Frequency Data’, Manuscript, Rotman School of Management and Institute for Policy Analysis, University of Toronto.
[19] Engle, R. (2000) ‘The Econometrics of Ultra-High Frequency Data’, *Econometrica*, vol. 68, pp. 1–22.
[20] Engle, R. and Russell, J. (1997) ‘Forecasting the Frequency of Changes in Quoted Foreign Exchange Prices with the Autoregressive Conditional Duration Model’, *Journal of Empirical Finance*, vol. 4, pp. 187–212.
[21] Engle, R. and Russell, J. (1998) ‘Autoregressive Conditional Duration. A New Model for Irregularly Spaced Transaction Data’, *Econometrica*, vol. 66, pp. 1127–1162.
[22] Ghysels, E. and Jasiak, J. (1998) ‘GARCH for Irregularly Spaced Financial Data: The ACD-GARCH model’, *Studies in Nonlinear Dynamics and Econometrics*, vol. 2, pp. 133–149.
[23] Goldstein, M., Kumar, P. and Graves, F.C. (2014) ‘Computerized and High Frequency Trading’, *The Financial Review*, vol. 49 (2), pp. 173–433.
[24] Gramming, J. and Maurer, K.O. (2000) ‘Non-Monotonic Hazard Functions and the Autoregressive Conditional Duration Model’, *The Econometrics Journal*, vol. 3, pp. 16–38.
[25] Guillaume, D.M., Dacorogna, M.M., Dave, R.R., Muller, U.A., Olsen, R.B. and Pictet, O.V. (1997) ‘From the Bird’s Eye to the Microscope: A Survey of New Stylized Facts of the Intra-Daily Foreign Exchange Markets’, *Finance and Stochastics*, vol. 1, pp. 95–129.

[26] Hafner, C. (1996) 'Estimating High Frequency Foreign Exchange Rate Volatility with Nonparametric ARCH Models', *Journal of Statistical Planning and Inference*, vol. 68, pp. 247–269.
[27] Haldane, A. (2011) 'The Race to Zero', Bank of England Speeches, Speech given to the International Economic Association Sixteenth World Congress, July 2011.
[28] Hasbrouck, J. (2013) 'High Frequency Quoting: Short-term Volatility in Bids and Offers', Working paper, [Online], Available: http://ssrn.com/abstract=2237499.
[29] Hasbrouck, J. and Saar, G. (2009) 'Technology and Liquidity Provision: The Blurring of Traditional Definitions', *Journal of Financial Markets*, vol. 12, pp. 143–172.
[30] Hasbrouck, J. and Saar, G. (2013) 'Low-Latency Trading', *Journal of Financial Markets*, vol. 16 (4), pp. 646–679.
[31] Hautsch, N. (2004) Modelling Irregular Spaced Financial Data – Theory and Practice of Dynamic Duration and Intensity Models, Berlin Heidelberg: Springer-Verlag.
[32] Hendershott, T., Jones, C. and Menkveld, A. J. (2011) 'Does Algorithmic Trading Increase Liquidity?', *Journal of Finance*, vol. 66, pp. 1–33.
[33] Jasiak, J. (1998) 'Persistence in Intertrade Durations', *Finance*, vol. 19, pp. 166–195.
[34] Jones, C. (2012) 'What do we know about high frequency trading?', Columbia University Working Paper
[35] Kirilenko, A.A., Kyle, A.S., Samadi, M. and Tuzun, T. (2011) 'The flash crash: the impact of high-frequency trading on an electronic market', Working Paper, CFTC and University of Maryland.
[36] Laughlin, G., Aquirre, A. and Grundfest, J. (2014) 'Information Transmission between Financial Markets in Chicago and New York', *Financial Review*, vol. 49, pp. 283–312.
[37] Lunde, A. (2000) 'A Generalized Gamma Autoregressive Conditional Duration Model', Discussion Paper, Aarlborg University.
[38] Madhavan, A. (2013) 'Exchange-traded funds, market structure, and the flash crash', *Financial Analysts Journal*, vol. 68, pp. 20–35.
[39] Malinova, K. and Park, A. (2013) 'Do retail traders benefit from improvements in liquidity?', Working Paper.
[40] O'Hara, M. (2015) 'High Frequency Market Microstructure', *Journal of Financial Economics*, vol. 116, pp. 257–270.
[41] O'Hara, M. (2011) 'What is a quote?', *Journal of Trading*, vol. 5(2), pp. 10–16.
[42] O'Hara, M., Saar G., and Zhang, Z. (2014) 'Relative Tick Size and the Trading Environment', *SSRN Electronic Journal*, vol. 1, DOI: 10.2139/ssrn.2463360

[43] Pagnotta, E. and Philippon, T. (2011) ‘Competing on Speed’, NBER Working Papers 17652, National Bureau of Economic Research.
[44] Ye, M., Yao, C. and Jiading, G. (2013) ‘The Externalities of High-Frequency Trading’, Working Paper 2013.

Managerial Economics
2016, vol. 17, no. 1, pp. 93–121
http://dx.doi.org/10.7494/manage.2016.17.1.93

Magdalena Kludacz-Alessandri*

Non-financial dimensions of measurement and assessment in the performance model for hospitals

1. Introduction

Performance measurement in the public healthcare system has become a more and more popular research problem throughout Europe and the world. In particular, these studies represent the link between the achievement of its major objectives and the available resources. The professional literature concerning performance measurement in the public healthcare system contains various opinions. There are many opinions that are critical of the idea of implementing a performance-measurement system in the public sector – this idea has been assessed as not credible or, at best, difficult to implement. On the other side, there is the idea of Jones et al. (2000) that performance measurements in the public healthcare sector could be defined and measured through a system of indicators where efficiency measures are especially important.

According to Shaw (2003), performance-measurement systems should be defined in a published national or regional plan that clarifies the values and participation of various stakeholders. Many European countries have developed frameworks for performance measurements at the national level. The concept of performance measurement within public hospitals and public-sector healthcare entities has already been proposed and explained by many authors for several countries, including: Tawfik-Shukor et al. (2007) for Holland; Guisset et al. (2002) for Belgium; Le Pogam et al. (2009) for the United Kingdom; Berg et al. (2005) for France, Sweden, and Denmark; Ştefănescu et al. (2011) for Romania; and Davis et al. (2013) and Gauld et al. (2011) for New Zealand.

* The College of Economics and Social Sciences in Plock, Warsaw University of Technology, e-mail: m.kludacz@pw.plock.pl

Non-financial factors play a valuable role in the assessment of the performance of hospitals. According to Goddard et al. (1999), some non-financial factors enrich the overall assessment of hospital performance by adding information that goes beyond the factors that are amenable to quantification within the formal system. Non-financial performance assessment can focus on various measures, ranging from the quality of food provided in hospitals to the volume of complaints received in a year. For instance, the indicators relating to waiting times are often connected with national and local targets regarding the effectiveness of the healthcare system. The main function of non-financial indicators in performance assessment is to act as a safety net in order to identify poor quality in hospitals. They could also be used to identify the best practice in order to achieve the aims regarding promoting good performance at the national level.

There are a lot of papers that explore the impact of non-financial factors on the assessment of hospital performance. For instance, according to Byrd (2013), Nelson et al. (1992), and Oszustowicz (1992), non-financial factors such as occupancy and quality are relevant for risk assessment. They have the potential to impact on revenues, expenses, and the ability of a hospital to repay its debt. As shown in a KPMG report (2008), increased risk is reflected in the higher cost of capital for hospitals with poorer quality-of-care ratings. Patients are expected to use the information about quality to choose providers who give better care, which, in turn, will impact the revenues of competing hospitals.

An interesting example showing how non-financial factors can affect the assessment of hospitals on the national level was described by Besley et al. (2009) and Bevan and Hood (2006). In 2001, the star ranking system for annual assessment of hospital performance was introduced in England. This system was based on non-financial indicators which measured targets relating to waiting times, cleanliness, and treatment-specific data.The base for this system were non-financial indicators that measured targets related to waiting times, cleanliness, and treatment-specific data. The number of stars awarded to hospitals was a reference point for many providers. The star ratings were published in the national and local media, and poor performance could end up being a reason for dismissal of hospital management teams. On the other hand, very good performance was a reason for increasing hospital's autonomy from the central government.

In reality, most of the performance-measurement systems for healthcare focus only on quality outcomes. There are not many systems that can connect the measurement of quality performance with financial measurements. In fact, according to Kludacz (2012), a properly designed system for performance measurement would encourage activities that could improve not only the quality of medical services but also the financial situation as well as the utilization of resources within hospitals and the whole healthcare system.

One of the most interesting initiatives undertaken in Poland to develop and implement a system for measuring financing in (as well as the non-financing performance of) Polish healthcare was the research project conducted within the grant of the Polish Minister of Science and Higher Education, entitled 'Conception of hospital reporting for an integrated system of performance assessment' carried out by a team from the Chair of Cost Accounting at the University of Szczecin represented by Hass-Symotiuk et al. (2010). The aim of the project as a whole was to identify and define sets of medical and economic information relevant for the purposes of an integrated performance measurement and assessment system designed for public hospitals, with the use of selected groups of indicators. This goal was realized within a framework of eight research stages:

1) defining the concepts, purpose, and components of the performance measurement and assessment system, presentation of the assumptions, and the stages of its construction;
2) analysis and evaluation of information sets generated by hospitals, arising from legal regulations and the needs of different healthcare-system stakeholders;
3) examining the information needs of the founding bodies of public hospitals (county, marshals' offices, and medical universities) and analysis of indicators that could be used to assess the effects of hospital activities;
4) examining the information needs of the National Health Fund regarding contracting and settlement of agreements concluded with hospitals;
5) analysis of the role and information needs of the marshals' governor and the Ministry of Health;
6) development of a model for measurement and evaluation of the performance of public hospitals, desirable from the point of view of different stakeholders;
7) verification of the proposed concept of a performance model using Data Development Analysis (DEA);
8) developing an information-reporting standard for the public hospitals that will allow the use of the performance model to assess the achievements of hospitals by the executives of hospitals, founding bodies, and the Ministry of Health.

The aim of this article is to present the results of the sixth stage of this project – a framework of a performance model with the key non-financial dimensions to measure hospital performance and a set of valid and reliable indicators related to these dimensions that could be used for supporting hospitals in assessing their results. In this Polish research project, it was assumed that a comprehensive performance model should be useful, not only for hospitals but also for other entities operating in the Polish healthcare system.

The general framework of the performance model was developed through extensive review of the literature on hospital performance projects, empirical studies conducted in hospitals, founding bodies, provincial branches of the National Health Fund and the Ministry of Health, and through a series of workshops that gather experts representing managers and chief accountants working in hospitals. The first selection of dimensions and indicators was based on an analysis of the different models of hospital performance measurement that were applied in various hospitals and countries. The general framework for the model (especially for the dimension and indicator selection) was built on strong empirical material. A survey was carried out in 36 Polish hospitals, in 35 funding bodies, and in 4 provincial branches of the National Health Fund. The problems regarding the framework of the model were also discussed and analyzed with the experts. Finally, the following outcomes were achieved:

– design of the structure of a performance model for measuring hospital performance in various dimensions and in three levels of the Polish healthcare management system;
– identification of four key dimensions for assessing hospital performance: patients, internal processes, development, and finance;
– analysis of nearly 100 performance indicators in order to prepare a core and a tailored set of performance indicators with an operational definition;
– elaboration of descriptive sheets for core indicators to support hospitals in interpreting their results.

Summarizing, one of the main achievements of the research project was to develop a general theoretical model for performance measurement in Polish hospitals as well as the whole healthcare system. This model could support the activities of:

– Managers of hospitals, in the area of evaluating hospital performance.
– The National Health Fund, in the area of contracting medical services.
– Regional Governor's and marshal's offices, in the area of performance of the healthcare system at the regional level.
– The Ministry of Health, in the area of performance of the healthcare system at the central level.

The basis for developing a model for the measurement card of performance improvements in the Polish healthcare system was the balanced scorecard approach (BSC), which was developed by Kaplan et al., (2005) in the early 1990s. This is one of the best-known performance assessment frameworks developed from the organization's strategies and includes indicators related to four perspectives: finances, customers, internal processes, and learning and growth. As shown in Amado et al.

(2012), the main strength of the BSC is its way of integrating different indicators to make links between the different dimensions of performance in a single system.

However, it would be difficult to use the BSC approach in a healthcare organization without any modifications. Grigoroudis et al. (2012) claim that healthcare entities are non-profit, socially oriented organizations, so the financial dimension of the framework should be changed and considered as a constraint rather than as an objective. The remaining dimensions should also have a different emphasis.

2. Construction of a conceptual framework for the performance-measurement model

Lachmann et al. (2015) provided evidence that an effective system for hospital performance measurement depends on the regulatory environment, type of ownership, and internal actors in the hospital. In this study, it was assumed that a comprehensive performance model would be useful not only for hospitals but also for other entities operating in the Polish healthcare system.

A hospital is characterized by a large and diverse group of stakeholders that includes the patients and their families, founding bodies and hospital owners, financial supporters of the healthcare services (National Health Fund), public administration responsible for health policies, Ministry of Health, physicians in cooperation with the hospital staff and their trade unions, suppliers of medicines, service providers (of transport, utilities, energy), financial institutions (lenders, insurers), and local communities. All of these stakeholders have a great impact on the activities of the hospital; therefore, to reflect their needs, the developed model takes into account three levels of a healthcare management system, composed of:

- the central level, represented by the Ministry of Health – at this level, the functions related to the strategic management (monitoring of public health, long-term planning, and the development of national standards) are realized;
- the regional level, represented by the regional governor, the marshal's office, and the regional offices of the National Health Fund – at this level, functions related to the strategic and operational management of mid-level healthcare systems (implementation of national plans in the region, the study of the health needs of the population in the region, and the coordination of the work of local institutions) are realized;
- the local level, represented by hospitals and their funding bodies – at this level, the functions related to the operational management of local hospitals and their funding bodies (e.g., the realization of the health needs of their patients) are realized.

The research project was based on the following assumptions:

- The performance model would be useful not only for hospitals and their founding bodies but also for the provincial branches of the National Health Fund and the Ministry of Health.
- The implementation of this model would require proper identification of the purposes in the most-important areas of hospital activity and the information needs of the stakeholders.
- The number of indicators should be limited so that the structure of the model is based on 23–25 core indicators (depending on the users); but, the users of the card (e.g., individual hospitals) can choose additional complementary indicators.
- The adopted indicators should allow for the evaluation of the degree of completion of the objectives and tasks by hospitals and their stakeholders.
- The implementation of the model would require the development of information standards for public hospitals, which would enable comparisons of data for performance assessment.

The construction of the performance-measurement model for healthcare systems was realized in a few stages. One of the most-important steps was to determine the dimensions of performance measurement that were relevant to the specific operations in hospitals and in the other stakeholders. The next step was to determine a universal set of targets for individual stakeholders to be implemented at three levels of the health management system (micro, meso, and macro). The last step was to choose indicators for each dimension and performance-measurement level. From the beginning, it was assumed that the performance-measurement system would show the relationship between the objectives and indicators in various dimensions of the model.

A very important task in the project was to determine the performance-measurement dimensions for evaluating achievements in hospitals. Especially important was to answer the question: How many, and what kind of dimensions should be included in the performance model to evaluate the most-important areas of hospital activity? The selection of the dimensions was based on an analysis of the literature describing different models of hospital-performance measurements that have been applied in various countries at the central level and in single hospitals.

A very interesting model for measuring the performance of the overall national healthcare system was developed in 2003 by the World Health Organization's Regional Office for Europe. The aim of the WHO project (PATH) was to: develop and disseminate a flexible and comprehensive tool for quality improvements in hospitals; to support them in assessing their performance; to question their own results; and to translate the results into actions for improvement. This project was carried out in 20 European countries. The final conceptual model presented by Veillard et al. (2005), consisted of six dimensions that were identified for assessing a hospital's performance: clinical effectiveness, safety, patient centeredness, production efficiency, staff orientation, and responsive governance. Groene et al. (2008) reviewed other indicator projects for hospital-performance assessments and compared them to the WHO-PATH model. The problem is that most of the models concentrate mainly on the quality of the care and don't include areas such as finance and technology.

Generally, a national healthcare system can be examined in terms of various aspects, such as human resources, facilities, health information systems, the technology used in direct patient care, financing, governance, and health policies. According to Phua et al., (2014), the problem is to focus on the areas in a national healthcare system that don't meet specific standards or minimal requirements and that should be improved through public policy interventions.

The final framework of the Polish model consists of four key dimensions for assessing healthcare performance: patients, internal processes, development, and finance. The model also includes a fifth area of assessment that reflects the needs of various stakeholders representing the three levels of the healthcare-management system. This area is called 'perspective of assessment' and consists of three elements:

– macro perspective – at the central level, represented by the Ministry of Health;
– meso perspective – at the regional level, represented by the governor, the marshal's office, and the regional offices of the National Health Fund;
– micro perspective – at the local level, represented by hospitals and their funding bodies

It should also be noted that high hospital achievements should be coupled with the needs and demands of the society and integrated with the whole healthcare system, including the promotion of health. The structure of the developed model is presented in Figure 1.

The essence of the performance model is to simultaneously present and analyze the hospital's achievements in four dimensions of evaluation: patients, internal processes, development, and finance. All of these dimensions are important from three perspectives: microeconomic, mesoeconomic, and macroeconomic. They have been taken into account during the development of a set of indicators for various kinds of stakeholders: hospital managers and hospital owners, the governor, the National Health Fund, and the Ministry of Health. For example, the health of the patient (the 'patient' dimension) is important for the hospital (to provide adequate medical care), for the Ministry of Health (to ensure the appropriate availability of healthcare resources throughout the country), and for the National Health Fund (to control the quality and availability of medical services).

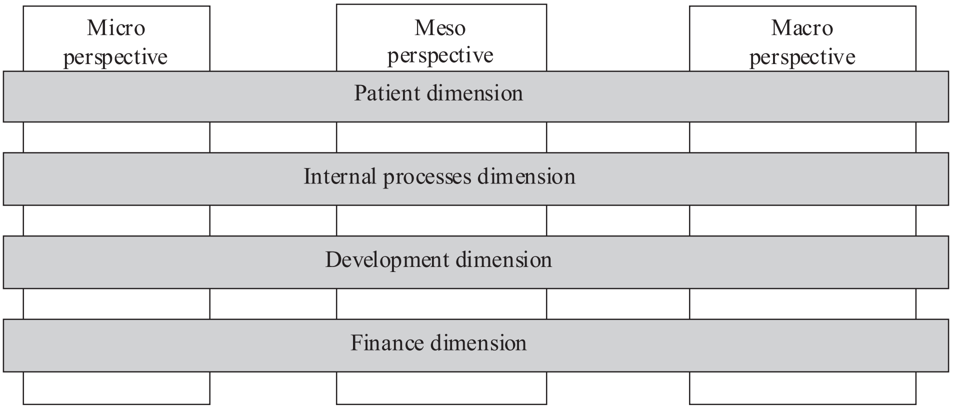

Figure 1. The conceptual performance model for the Polish healthcare system

3. Determining the goals for each dimension of the performance-measurement model

The multidimensional assessment enables a systematic and parallel concentration on the key areas of hospital activities by different stakeholders. It prevents the optimization of one area of activity at the expense of the others. It is worth noting that the objectives and indicators for each dimension can communicate directly with the other dimensions, and they can influence each other. For example, the appropriate organization of hospital wards (an 'internal processes' dimension) and the training activities of the medical staff (a 'development' dimension) affect the quality and efficiency of patient care and, therefore, the level of patient satisfaction (a 'patient' dimension). Finally, it also affects the costs of the medical service and the hospital financial performance (a 'finance' dimension). The dimensions of hospital performance are described in Table 1.

It is worth noting that the key dimension in the performance model is the patient dimension, which brings together the key issues related to the quality of the medical services and of the final product. The quality of the medical services can be measured as the attentiveness of care providers to patient needs and expectations (e.g., the quality of the medical equipment and the quality of the 'hotel' functions of a hospital). Santiago (1999) underlines that the quality of the product can affect the final result of treatment – whether the health of patient is improved and whether he is able to avoid death, disability, discomfort, and further disease.

Table 1

The dimensions of the hospital performance-measurement model

Dimension	Description
Patient	This dimension of performance wherein a hospital places patients at the center of care and pays particular attention to the needs of the patients, their expectations, their autonomy, their access to a hospital support network, communication, confidentiality, trust, dignity, their choice of a provider, and their desire for prompt, timely care [WHO, 2000a]. **From the micro perspective** – the dimension wherein the hospital properly takes care of high-quality medical services for all patients and achieves the desired results (appropriate care, treatment outcomes, clinical effectiveness). **From the macro/meso perspective** – the responsibility to address social needs, ensure continuous and coordinated care, health promotion, and ensure healthcare for all citizens of the country/region regardless of their race, physical, cultural, social, demographic, or economic characteristics.
Internal processes	This dimension is concerned with the operational aspects of medical activities that ensure better use of the resources as the key processes of the hospital and healthcare system. It indicates the most-important factors that influence the efficiency of the hospital, its position on the medical market, and its ability to generate value for the patients. **From the micro perspective** – this dimension focuses on the optimal use of resources in the key internal processes that hospitals should realize in order to create a foundation for long-term development. The selection of targets and measures in this dimension requires an analysis with a consideration of hospital resources. **From the macro/meso perspective** – this dimension should take into account the state of all material resources available in the national/regional healthcare system (medical equipment, drugs, medicinal preparations) that are necessary for the proper performance of medical services.
Development	This dimension attempts to identify the factors, skills, and tools necessary to stimulate the development of the hospital and to create its value. An important task is to look for new methods of improving medical activities and analyzing the key areas that are important for development. **From the micro perspective** – this dimension examines the aspects of medical activities that are crucial for the development of the hospital. The important areas in this dimension are: information systems within the organization and medical technologies as well as the employees and their potential, skills, experience, and satisfaction. An especially essential area in the analysis of this dimension is the degree to which hospital staff are appropriately qualified to deliver the required patient care, have opportunities for continued learning and training, work in good conditions, and are satisfied with their work (Klazinga et al., 2001).

Table 1 cont.

Dimension	Description
Develop-ment	**From the macro/meso perspective** – it is important to take care of the appropriate availability of medical staff in the national/regional healthcare system and predict the demand for new professionals in the future. This dimension also focuses on the national/regional development of new technologies and information infrastructures supporting innovation as well as initiating research and development work in healthcare.
Finance	The financial dimension of the model mirrors the hospital's mission statement regarding financial safety. This dimension refers to a hospital's financial health, efficiency, profitability, and determines the possible ways to reach financial success. **From the micro perspective** – it indicates the most rational management of the hospital's funds. Financial-dimension measures are usually associated with profitability. Typical measures would include operating profits, return on assets, and profit margins. In a public hospital (where the main objective is to balance the budget and to settle liabilities), the performance assessment should focus on such problems as how to balance costs with revenues and how to maintain financial liquidity. **From the macro/meso perspective** – it will be important to answer the following questions: What are the national/regional costs of the healthcare system? Are the medical entities operating in the national/regional healthcare system financially stable? How should the optimal financing of medical services be ensured?

According to Shaller (2007), patients are the most-important group of customers in the healthcare system, and they should be at the center of the performance-assessment system. The patient should be the main focus of the strategies in each hospital. From the patient dimension, it is important to determine who is a patient in a hospital and which medical services are the most expected and desirable from the point of view of the patients. From the macroeconomic perspective, it is also important to determine the level of satisfaction of the patient and to improve the availability of medical services in different regions of the country. The active participation of patients and their full involvement in the treatment process (as well as their acceptance of treatment methods) are necessary to achieve the desired results. This acceptance is dependent on the quality of the medical services and patient trust towards the doctors.

An important stage of the project was determining the key questions for each dimension to meet the information demands of all stakeholders. Such questions are often identified in other performance-measurement models and help to choose the indicators for measuring hospital performance, (as shown by Griffith et al., 2000; and Niven, 2011). The dimensions (along with strategic questions) are presented in Figure 2.

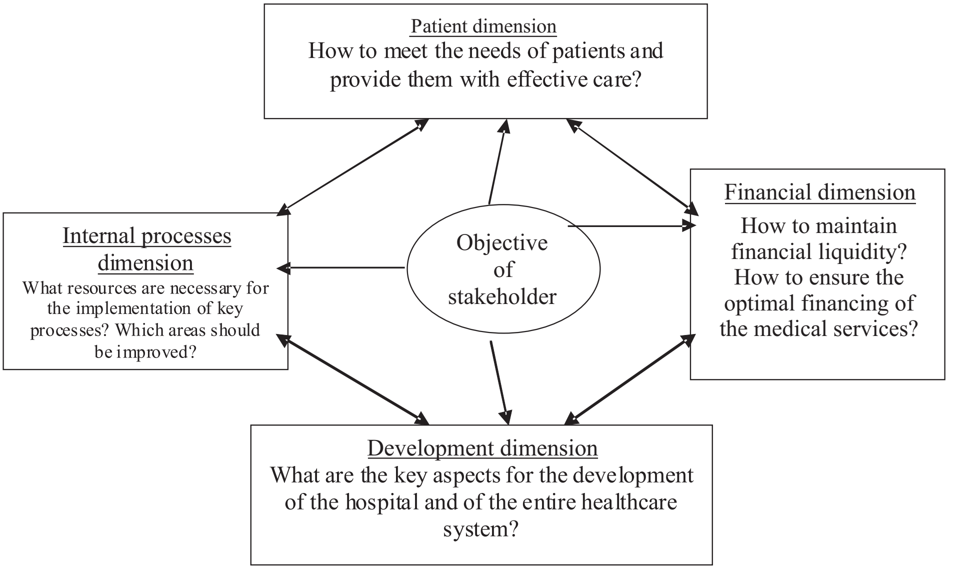

Figure 2. The dimensions of a performance model with strategic questions

All of these dimensions were analyzed independently. This multi-dimensional approach to measuring performance was based on the assumption that satisfactory achievements in the healthcare system should rely on:

- highly qualified medical staff who can use current medical knowledge,
- available medical technology and resources,
- professional medical care and high quality medical services,
- optimal health outcomes.

Assessing the performance of a healthcare system begins with defining its goals. The model takes into account the universal goals that a hospital and the individual stakeholders should realize in all of the defined dimensions of performance. The determination of these goals was conducted through the analysis of tasks carried out by the various stakeholders representing the three levels of the healthcare system.

According to the World Health Organization (2000b), the main goal of a health system is the delivery of effective, preventative, and curative health services to the full population, equitably and efficiently, while protecting individuals from catastrophic healthcare costs. Kruk et al. (2008) adds that the state is, therefore, responsible for the needs and demands of the population and is obligated to ensure the availability, accessibility, acceptability, and quality of health services.

The universal objectives determined for a healthcare system, together with the key clinical and economic goals of hospitals and other stakeholders, were a significant reference point for constructing a set of indicators for the performance model. The objectives determined in the three non-financial dimensions of the model are presented in Table 2.

Table 2

The goals of hospitals and other stakeholders for the non-financial dimensions of the performance-measurement model

Dimension	Perspective	Goals
Patient	micro	– improving the satisfaction of patients and their families, and taking care of their safety – improving the quality of medical services and treatment conditions – optimizing treatment results – improving the clinical effectiveness of the medical services
	meso/ macro	– developing health policies – improving the health of the society in the region/country – ensuring high-quality medical services in the region/country – promoting health and health programs in the region/country
Internal processes	micro	– optimizing treatment processes in hospitals – improving the availability of medical services in hospitals – improving employee productivity in hospitals – optimizing material resources in hospitals and improving the efficiency of resource usage
	meso/ macro	– improving the availability of medical entities in the region/ country – improving the availability of medical staff in the region/ country – improving the availability of material resources (medical equipment, hospital beds, drugs) in the region/country – increasing access to treatment and reducing waiting times for health services in the region/country

Table 2 cont.

Develop-ment	micro	– improving the satisfaction and qualifications of the medical staff in the hospital – developing modern medical technology in the hospital – improving the information systems in the hospital – implementing new medical procedures and their standardization – improving the management systems in hospitals
	meso/ macro	– improving the training systems for medical staff in the country/region – developing the hospital infrastructure in the region/country – distributing information about the healthcare system – improving the healthcare management system – developing the DRG (Diagnostic Related Groups) system in the country

It is worth noting that the goals of hospitals vary by their ownership type. For instance, Villa et al. (2013) underlines that for-profit hospitals have to maximize profits, whereas private non-profit hospitals must balance multiple objectives such as profits, quality of care, quantity of care, and charitable care. The goals presented in the micro perspective of the model were selected only for public hospitals.

4. Selection of indicators for the performance-measurement model

The next step in the project was to define the indicators for each dimension and perspective of the model. The idea was to select such indicators that could measure whether the goals specified for all dimensions and all three levels of the health management system were realized.

Performance indicators are seen as a promising answer to the demands for increased transparency, accountability, and quality within healthcare (Berg et al., 2005). The literature abounds with definitions of the indicators that can be used to evaluate many aspects of a hospital's achievements. However, most of the performance indicators concern the dimensions connected with the quality of medical services; e.g., patient-safety, clinical effectiveness, safety, and appropriateness of the care (as well as its equitable delivery). On the other hand, we required indicators that could also measure the efficiency and resource use of medical facilities.

Hundreds of different indicators were analyzed in order to establish the set of indicators for the model. Also, each key indicator was checked for its appropriateness to the activities of public hospitals. As a result of this work, a group was selected consisting of dozens of measures, taking into account empirical research conducted in hospitals, founding bodies, provincial branches of the National Health Fund, and the Ministry of Health.

In the course of the study, it was necessary to decide how many indicators should be chosen for different dimensions and perspectives of the model. In the literature, there is a wide range in the number of indicators recommended for use in various systems of evaluations – ranging from 13 to 44. The upper limit of this range seems to be significantly above the levels recommended in the literature and definitely exceeds the ability of managers to focus on such numbers. The problem with the amount of indicators also encompasses both the resources and the costs involved in the process of measurement, data collection, and the analysis and interpretation of the indicators.

Using opinions obtained during the survey, the authors of the project choose 30 indicators. The indicators that were the highest rated by the respondents of the survey were further evaluated by experts working in hospitals (hospital managers, chief accountants), and the final selection of indicators for the model was completed using the following selection criteria:

- Ability for use and implementation – it was important to answer the following questions: Can the indicators be accepted by their potential users? Are the indicators easy to calculate?
- Importance and validity – it was important to answer the following questions: Do the indicators reflect aspects that are relevant for their users and important for the current healthcare system? Do the experts and users agree that the measurement of the indicator is necessary? Do the indicators measure achievements in the appropriate dimension? Are the indicators related to the objectives in the various dimensions of measurement?
- Availability – it was important to answer the following questions: Is the measurement data readily available? Is the measurement data available at a reasonable price?
- Credibility, reliability and usefulness – it was important to answer the following questions: Does the information obtained as a result of measurement reflect the achievements of the hospital? Is it possible to control it? Can hospitals rely on these indicators in case of problems? Is the measurement data reliable? Can the indicator be used for benchmarking? Is the indicator related to other indicators that measure the achievements of hospitals? Does the indicator measure the trait for analysis?

As a result of the research, the indicators were divided into two groups:

– A set of core indicators, including a limited group of 23–25 standard indicators in each perspective.
– A set of 70 complementary indicators, including non-standard indicators used only in specific situations and adjusted to the individual needs of hospitals.

The set of core indicators selected for non-financial dimensions and all perspectives of the model is presented in Table 3, while the set of complementary indicators for a microeconomic perspective is presented in Table 4. Similar sets of complementary indicators can be developed for the meso and macro perspectives.

Table 3

The set of core indicators for non-financial dimensions of the performance model

Dimension	Indicators	
	micro perspecive	**meso/macro perspective**
Patients	– average wait time for medical services (in days) in the hospital – patient satisfaction in the hospital – number of patient complaints – rate of nosocomial infections in the hospital – mortality rate in the hospital – average length of stay for curative care in the hospital	– Healthy Life Years in the region/country – life expectancy in the region/country – self-perceived health in the region – psychological well-being (Energy and Vitality Index) in the region/country – average length of stay for curative care in the region/country – average wait time for medical services (in days) in the region/country
Internal processes	– number of physicians and nurses per patient in the hospital – rate of labor productivity – surgical theatre usage – curative care bed occupation rate in the hospital – medical technology per patient in the hospital – number of available beds in the hospital	– number of physicians and nurses per capita in the region/country – medical technology per capita in the region/country – curative care bed occupancy rate in the region/country – operation theatres per capita in the region/country – number of available beds per capita in the region/country – number of hospital wards in the region/country

Table 3 cont.

Dimension	Indicators	
	micro perspecive	**meso/macro perspective**
Development	– place of the hospital in the national ranking for quality – training costs per employee in the hospital – value of the new investments in the hospital per year – level of the employee satisfaction in the hospital – value of new medical technologies in the hospital per year – degree of computerization in the hospital	– value of the new investments per capita in the region/country – value of new medical technologies per capita in the region/country – number of accredited hospitals in the region/country – level of development of a national/regional "e-health" program – level of development of a DRG (Diagnostic Related Groups) system in the region/country – number of young specialists per capita in the region/country

Table 4

The set of selected tailored indicators of non-financial dimensions for the micro perspective

Dimension	Indicators
Patients	– number of statements of claim against the hospital – average length of stay in the various wards of the hospital – percentage of patients with individual treatment plans – average length of treatment for various patients (DRG cases) – number of Acute Oncology Services (AOS) – number of patients discharged without a health improvement – readmission rate – number of emergency admissions – blood culture contamination rate – mortality rate in the hospital wards – agreement between the initial and final diagnoses
Internal processes	– number of midwives per patient in the hospital – number of hospital wards by specialty – hospital beds by type of care; e.g., psychiatric, long-term-care beds, other beds – physicians by medical speciality per capita; e.g., oncology, radiology, cardiology – laboratory utilization rate – average age of the medical equipment – number of outpatient departments – number of theatre rooms – curative care bed occupation rate in various hospital wards

Table 4 cont.

Development	– number of hours of training per employee per year – number of Clinical Practice Guidelines – number of implemented innovations – rate of implementation of accreditation standards – number of key management areas covered by the information standards – number of employees who have upgraded their skills – staff turnover – value of modern medical equipment purchased in the last two years divided by the total value of the equipment – degree of computerization in the medical part of the hospital

Both sets of indicators were selected based on scientific evidence after the surveys were carried out in hospitals, branches of the NHF, and the funding bodies. The results of discussions with experts were also taken into account. The performance model deliberately avoids presenting operational definitions of indicators. Hospitals and stakeholders can use core and tailored indicators from a micro and meso perspective and adapt them to their needs, but these indicators have to be operationally defined by them. Tailored indicators from a macro perspective should reflect national specific priorities and can be used for comparisons of hospitals.

There is much discussion in the literature about the problems of measuring some indicators regarding the quality of health care. One such broadly discussed indicator is patient satisfaction. According to Al-Abri and Al-Balushi (2014), there is no consensus among the literature on how to define the concept of patient satisfaction in healthcare. For instance, Nelson et al. (1992) defined patient satisfaction as a degree of congruency between patient expectations of ideal care and their perceptions of real care received. On the other hand, Mohan and Kumar (2011) pointed out that patient satisfaction is related to emotions, feelings, and [individual] perception of the healthcare services.

Another problematic indicator is labor productivity. Productivity is usually defined as a measure of the effectiveness and efficiency of an organization in generating output with the available resources. Labor productivity is a measure of economic output per unit of labor input. According to Brill (2015), hospital labor productivity is notoriously difficult to assess because of problems with measuring a hospital's output.

The labor input means the quantity of utilized labor and is not complicated to measure. Wulong and Morin (2014) claim that labor input should be measured by the total worked hours of doctors, nurses, and administrative staff. Also, according to OECD (2001), the most-appropriate measure for the volume of labor input is the number of hours actually worked. However, labor input should also include such factors as sick leave and holiday leave.

The challenge is to measure the actual hospital output, because hospitals provide different types of services to different types of patients. In the past, the volume of output was measured by the labor costs for medical and administrative staff, but this did not allow for measurement of labor productivity at the national or regional level. According to Chansky and Garner (2015), current labor-productivity indicators usually rely on volume-based measurement of hospital services, which are separated into outpatient visits and inpatient courses of treatment. So they developed and analyzed two concepts to measure a physical count of the services provided – one based on the number of procedures and the second based on the number of complete administered treatments adjusted for disease severity. Chansky et al. (2013) proposed the measurement of inpatient services on the basis of patient discharge data. Each discharge should be assigned to an appropriate Diagnosis-Related Group (DRG) that corresponds to the primary diagnosis as well as to the procedures and services used during treatment.

An accurate method to measure patient satisfaction is the quantitative approach. According to Jose et al. (2006), the most-popular assessment tool for conducting patient satisfaction studies are standardized questionnaires. Veillard et al. underline that there are three broad approaches to patient surveys. They can measure patient experience regarding received care, patient satisfaction, or the gap between patient expectations and perceived experience. These three approaches should be complementary.

Freeman (2002) emphasized the importance of distinguishing the use of internal and external indicators. It is worth noting that the performance model contains both kinds of indicators. External indicators can be used by governments, regional governors, and funding bodies to assess the availability of medical services, the health of the population, and the quality of care by healthcare providers, and they can be compared on a macro level. Internal indicators, on the other hand, can be used by healthcare providers to monitor and improve the outcomes of their care processes.

For each indicator in the core set of indicators, a descriptive sheet has been developed, which contains the following information:

- name and description of indicator,
- calculation formula,
- frequency of measurement,
- links to the dimension and goals of measurement,
- person responsible for measurement,
- indicator target value and interpretation guidance,

– information about data collection needed to calculate index,
– supplementary indicators to measure similar areas.

The model for the descriptive sheet is presented in Table 5.

Table 5

Model of the descriptive sheet

Elements of descriptive sheet	**Characteristics of indicator**
The name of indicator	
Dimension of measurement	
Calculation formula	
Description of indicator	
The frequency of measurement	
Indicator target value	
Sources of data necessary for measurement	
Supplementary indicators	
The person responsible for measurement	

Descriptive sheets were prepared for core indicators, with a detailed description of each indicator, formulas for the calculation, reference values, rules for frequency of measurement and analysis, as well as a recommendation for the appropriate usage. The results should relate to external references, together with internal comparisons over a period of time, and give guidance on interpretation. I tend to agree with the WHO (2007), that the descriptive sheets for indicators are not able to answer all questions regarding the definition and other information about the indicators, especially their reference values.

For a portion of the indicators, the reference values set as parameters were based on evidence from the literature, but such a situation was observed mainly for indicators from the finance dimension. For most of the indicators from non-financial dimensions, the reference values weren't standardized at the hospital, regional, or national levels, and further work will be required to find an optimal reference value for hospitals. In the line 'indicator target value', there was often a suggestion that this value should be based on the historical results obtained by the hospital.

The reference value can be established as average hospital points, or it can be described as percentages. In some cases, references are based on exclusion criteria. Each year, the reference values should be reviewed in conjunction with any new information from the collected data, and additional modifications should be documented.

According to Gaev (2010), reference values for the indicators may be established using internal or external benchmarking or by regulation. At the hospital level, the reference value should be established by comparing the performance of the hospital against itself. For example, if a hospital achieved an average monthly curative care bed occupation rate of 72 percent last year, it may choose to use this as the reference value this year. Also, according to Stausberg et al. (2011), a reference value should be established based on a direction for better results.

Sometimes, it will be enough to connect a reference value with hospital management evaluation guidelines; for instance, establish if higher or lower indicator values indicate better performance. It is not always possible, however, to present one set of reference values for all kinds of hospitals. According to Ferretti and Zangrandi (2013), it is clear that the performance of hospital models is diverse due to the specific historical data and specifics of the organization.

For meso and macro perspectives, reference values for the indicators should be established on the basis of an external benchmark; i.e., by comparing the performance of one hospital with its peer group. In such cases, the reference values should be calculated from the merged data of all hospitals from the group. As a result of such benchmarking, a reference-values database for all types of indicators could be published and used in performance systems at the national and regional levels.

According to Shaw (2003), the reference values and objectives of hospital performance-measurement systems should be made explicit and agreed with by stakeholders. We should remember that performance-measurement systems should aim to manage and improve hospital performance rather than to generate unreliable rankings and comparisons.

Esders (2008) stressed that it is good to have such an already-established set of reference indicators that can be used by all kinds of different stakeholders with differing information needs, but it is not easy to realize this goal. Even if the selected set of indicators is stable, the descriptions of indicators are not static and are, therefore, still open to further development.

It is worth noting that the achievements of hospitals measured by the designed set of indicators should be periodically monitored. In such case, it is possible to use the card for monitoring the achievements presented in Table 6.

The performance model for monitoring the achievements of the hospital assumes a possibility for comparing the selected indicators on a quarterly basis. Such a comparison seems to be the most optimal for most of the indicators. However, in the case of large fluctuations in the monitored indicators, it is possible to compare them more often.

Table 6

Card for monitoring achievements of the hospital

Selected indicators	Quarters			
	I	II	III	IV
Patient dimension				
Indicator 1				
Indicator 2				
Internal processes dimension				
Indicator 1				
Indicator 2				
Development dimension				
Indicator 1				
Indicator 2				
Financial dimension				
Indicator 1				
Indicator 2				

Implementation of the performance model involves ensuring access to the right data as well as the political and cultural issues. Hospitals need educated people who will understand the goal and use of the performance model. Kurien and Qureshi (2011) claim that the challenge in implementing any performance-measurement system is a cultural shift in many organizations.

It is not easy to effectively implement the performance model. Bourne et al. (2003) suggest that implementation of any performance-measurement framework is fraught with complexity at varying levels and; therefore, implementation issues are critical to its success. Thakkar et al. (2009) mention four critical factors that impact the successful implementation of performance models at the hospital level, being strategy, leadership, culture, and capability. Each of these elements is interconnected with the others. Charan et al. (2008) suggest that critical factors such as these are not the only ones of relevance. According to them, successful implementation depends on; e.g., an effective information system, employee commitment, partnership with stakeholders, appropriate performance indicators, overcoming mistrust, funds for implementation, commitment by top management, awareness about performance model, and consistency with strategic goals.

The success of implementing the performance system lies especially in choosing the right indicators. In the opinion of Moreira (2008), the main disadvantage of the approaches based on balance scorecards is their partial nature, which, according to the indicators selected, may lead to contradictory conclusions. One way to solve this problem is to aggregate several partial indicators into one efficiency index. Some authors recommend the use of such aggregate measures to improve assessments from various stakeholders (e.g., funding bodies, regional and government agencies) at the national and regional levels. For instance, according to Shwartz et al. (2008) and Staiger et al. (2009), such aggregate measures can eliminate the limitations associated with individual indicators and are often used to summarize hospital performance and enhance their accountability. The approaches that are based on the aggregated efficiency measures usually involve two steps: the estimation of an efficiency frontier, and the calculation of each unit's deviation from that benchmark.

Most research that has a goal to measure efficiency is based on either parametric or non-parametric methods. Madl et al. (2008) underline that the main difference between the parametric and the non-parametric methodology is the approach to constructing an efficiency frontier that provides a benchmark by which the efficiency performance can be assessed.

The most-commonly used parametric method is Stochastic Frontier Analysis (SFA). This is a statistical technique that allows one to estimate the deviations of performance scores from the efficient frontier. In this method, a usual function with constant parameters is specified a priori, which is one of the main drawbacks of this approach. The efficiency is measured using the residuals from the estimated equation where the error term is divided into two components: inefficiency, and a statistical residual. According to Kontodimopoulos et al. (2011), the basic concept of SFA models is that the deviations are not entirely due to inefficiency, since they acknowledge that random effects outside the control of the units may affect output. Usage of this parametric method is mostly found in the econometric literature.

While the parametric techniques require the ex-ante definition of the functional form of the efficiency frontier, the non-parametric approach is primarily data-driven, because it constructs an efficiency frontier using input/output data for the whole sample. A very popular non-parametric methodology that uses mathematical-programming techniques is DEA (Data Envelopment Analysis). This allows for the measurement of efficiency without requiring the specification of a functional form. However, the existence of exogenous factors is not taken into account in the analysis.

According to Hollingsworth (2008), 75% of the papers regarding the problems of efficiency measurement in healthcare use the DEA methodology. For instance, Fragkiadakis et al. (2013) used the DEA methodology, allowing for the identification of input and output variables that should be considered, as well as their aggregation into a multi-dimensional efficiency analysis context. The analysis was based on a comprehensive set of variables related to the volume and type of services provided by Greek hospitals, their size, personnel, and costs structure. A similar methodology was used by Steinmann et al. (2004), who extended the analysis to comparisons with other EU countries.

The value developed by the authors of the integrated system of performance assessment based on the performance model is reflected in the possibility of using it not only in individual hospitals but also at the regional and central levels of the healthcare system. In such a case, it can be used to compare hospital achievements and prepare their rankings. The method that can best be used to conduct comparative analysis at the meso and macro levels using the indicators from different areas of the performance model was DEA. This method was chosen in the final stage of the research project for verification of the model Hass-Symotiuk et al. (2010).

5. Conclusions

According to Lowe et al. (2011), a financial emphasis still persists among many users of performance-measurement tools based on the BSC. Their findings indicate that an emphasis on short-term financial outcomes is a problem in many organizations.

The recently developed Polish performance model exceeds the scope of the financial dimension and enriches it with three non-financial dimensions: patients, internal processes, and development. This allows experts to concentrate not only on the tasks of the hospital but also on the objectives of other stakeholders. This model, therefore, puts an emphasis not only on financial indicators but especially on non-financial indicators. The growing importance of non-financial indicators stems from the following weakness of financial indicators:

- Financial indicators are determined on the basis of data from the past, so they are not able to warn about changes in the market of medical services.
- They are not sufficient for determining the financial results in the future. It is not possible to use them to assess the healthcare system of the future.
- Financial indicators mainly measure the effectiveness of material resource management (hospital facilities, medical equipment, information systems, medicine and medical supplies, cash, etc.) that have less and less importance in the economic value of the hospital.

The performance model developed at the level of a hospital can be the basis for the development of cards for individual hospital wards, because the indicators developed for each dimension may also be made more detailed. In addition, each department can supplement the card with indicators that are considered essential for its development. At the ward level, there can be more freedom and flexibility. The performance model also provides a clear and comprehensive interpretation, which helps one understand the results and suggest areas for improvement. The implementation of this model at the microeconomic level may be of fundamental importance for:

– the application of modern hospital management methods,
– the construction of hospital incentive schemes,
– the proper allocation of resources.

The performance model presented in the article reflects the universal non-financial goals of hospitals and key entities operating in the healthcare system. Taking into account the needs of all stakeholders makes this card not only an instrument for hospital management but also a tool for health policies. Therefore, this model can be a practical tool for monitoring as well as for internal quality improvements throughout the healthcare system. The implementation of this model at the macroeconomic level can be used by various stakeholders to:

– collect data about the achievements of Polish hospitals,
– compare the achievements of hospitals,
– initiate activities to improve the quality of medical services,
– determine the rankings of hospitals,
– improve the availability of healthcare services.

The use of a model at the macro and meso levels can also provide an opportunity to make comparisons between hospitals according to various criteria (regions, sizes) to better understand the differences between hospitals and provide opportunities for improvements in individual hospitals.

In implementing the performance-measurement model, it is very important to use tools that could enable the monitoring and benchmarking of the efficiency of hospitals using aggregated indicators from various dimensions and facilitate the implementation of the best practice guidelines and policies.

References

[1] Agyei-Mensah, B.K. (2012) 'The Impact of Contingent Factors on Performance Measures in the Rural Banks of Ashanti Region of Ghana', *International Journal of Academic Research in Business and Social Sciences August*, vol. 2, No. 8, pp. 78–110.

[2] Al-Abri, R. and Al-Balushi, A. (2014) 'Patient satisfaction survey as a tool towards quality improvement', *Oman Medical Journal*, vol. 29(1), pp. 3–7.

[3] Amado, C.A., Santos, S.P. and Marques, P.M. (2012) 'Integrating the Data Envelopment Analysis and the Balanced Scorecard approaches for enhanced performance assessment', *Omega*, vol. 40, No. 3, pp. 390–403.

[4] Bevan, G. and Hood, C. (2006) 'What's Measured is What Matters: Targets and Gaming in the English Public Health Care System', *Public Administration*, vol. 84, pp. 517–538.

[5] Berg, M., Meijerink, Y., Gras, M., Goossensen, A., Schellekens, W., Haeck, J., Kallewaard, M. and Kingma, H. (2005) 'Feasibility first: developing public performance indicators on patient safety and clinical effectiveness for Dutch hospitals', *Health Policy*, vol. 75, No. 1, pp. 59–73.

[6] Besley, T., Bevan, G. and Burchardi. K. (2009) 'Naming and Shaming: The Impacts of Different Regimes on Hospital Waiting Times in England and Wales', LSE Health and Social Care Discussion Paper, London: London School of Economics and Political Science.

[7] Bourne, M., Neely, A., John Mills, J., and Platts, K. (2003) 'Implementing performance measurement systems: a literature review', *International Journal of Business Performance Management*, vol. 5, No. 1, pp. 1–24.

[8] Brill A., (2015) Hospital (un)productivity, [Online], Available: http://www.modernhealthcare.com/article/20151210/NEWS/151219999 [March 2016].

[9] Byrd, J.R. Jr. (2013), 'Impact of quality performance on financial risk and cost of capital in hospitals', Ph.D., The University of Alabama at Birmingham.

[10] Chansky, B. and Garner, C.A. (2015) 'New Measure of Labor Productivity for Private Community Hospitals: 1993–2012', *Monthly Labor Review*, vol. 138, No 1.

[11] Chansky, B., Garner, C. and Raichoudhary, R. (2013) 'Measuring output and productivity in private hospitals', Chapter in forthcoming NBER book *Measuring and Modeling Health Care Costs*, University of Chicago Press.

[12] Charan, P., Shankar, R. and Baisya, R.K. (2008) 'Analysis of interactions among the variables of supply chain performance measurement system implementation', *Business Process Management Journal*, vol. 14, No. 4, pp. 512–529.

[13] Davis, P., Milne, B., Parker, K., Hider, P., Lay-Yee, R., Cumming, J. and Graham, P. (2013), 'Efficiency, effectiveness, equity (E 3). Evaluating hospital performance in three dimensions', *Health Policy*, vol. 112, No. 1, pp. 19–27.

[14] Esders, M., (2008), 'Conceptualising the Assessment of Eco-Innovation Performance. A Theory Based Framework for Deriving Eco-Innovation Key Performance Indicators and Drivers (EI-KPIs)', [Online], Available: https://www.leuphana.de/fileadmin/user_upload/Forschungseinrichtungen/csm/files/Arbeitsberichte_etc/61-0_download.pdf [April 2016]

[15] Ferretti, M. and Zangrandi, A. (2013) 'Management, Organization and Performance Measurement in Italian Healthcare System: The Experience of Sicily Region', *Emergency*, vol. 2 No. 1, pp. 24–43.

[16] Fragkiadakis, G., Doumpos, M., Zopounidis, C. and Germain, C. (2013) 'Evaluation of the Efficiency of Greek Hospitals: A Non-Parametric Framework and Managerial Implications', Working Paper, [Online], Available: http://www.fel.tuc.gr/Working%20papers/2013_02.pdf [April 2016]

[17] Freeman, T. (2002), 'Using performance indicators to improve health care quality in the public sector: a review of the literature', *Health Services Management Research*, vol. 15, No. 2, pp. 126–137.

[18] Gaev, J. (2010) 'Successful measures: benchmarking clinical engineering performance', *Health Facilities Management*, vol. 23(2), pp. 27–30.

[19] Gauld, R., Al-Wahaibi S., Chisholm, J., Crabbe, R., Kwon, B., Oh, T., Palepu, R., Rawcliffe, N. and Sohn, S. (2011) 'Scorecards for health system performance assessment: the New Zealand example', *Health Policy*, vol. 103, No. 2, pp. 200–208.

[20] Goddard, M., Russell M., Smith Peter C. (1999) 'Assessing the performance of NHS Hospital Trusts: the role of «hard» and «soft» information', *Health Policy*, vol. 48, No. 2, pp. 119–134

[21] Griffith, J.R. and King, J.G. (2000) 'Championship management for healthcare organizations', *Journal of Healthcare Management*, vol. 45, pp. 17–31.

[22] Grigoroudis, E., Orfanoudaki, E. and Zopounidis, C. (2012) 'Strategic performance measurement in a healthcare organisation: A multiple criteria approach based on balanced scorecard', *Omega*, vol. 40, No. 1, pp. 104–119.

[23] Groene, O., Skau, J.K. and Frølich, A. (2008) 'An international review of projects on hospital performance assessment', *International Journal for Quality in Health Care*, vol. 20, No. 3, pp. 162–171.

[24] Guisset, A.L., Sicotte, C., Leclercq, P. and d'Hoore, W. (2002), 'Définition de la performance hospitalière: une enquête auprès des divers acteurs stratégiques au sein des hôpitaux', *Sciences sociales et santé*, vol. 20, No. 2, pp. 65–104.

[25] Hass-Symotiuk M. (ed.) (2010) 'Koncepcja sprawozdawczości szpitali na potrzeby zintegrowanego systemu oceny dokonań', Szczecin: Wydawnictwo Naukowe Uniwersytetu Szczecińskiego.

[26] Hollingsworth, B. (2008) 'The measurement of efficiency and productivity of health care delivery', *Health Economics*, vol. 17, pp. 1107–1128.

[27] Iftikhar, A., Allah N., Shadiullah, K., Habibullah, K., Muhammad, A.R. and Muhammad, H.K. (2011) 'Predictors of patient satisfaction', *Gomal Journal of Medical Sciences*, vol. 9, No. 2, pp. 183–188.

[28] Jones R. and Pendlebury M. (2000), Public Sector Accounting, 5th Ed. London: Financial Times/Prentice Hall.

[29] José, M.Q., Nerea, G., Amaia, B., Felipe, A., Antonio, E., Cristóbal, E., Emilio, S. and Andrew, T. (2006) 'Predictors of patient satisfaction with hospital health care', *Health Services Research*, vol. 6, No. 102, [Online], Available: http://www.biomedcentral.com/1472-6963/6/102 [March 2016].

[30] Kaplan, R.S. and Norton, D.P. (2005) 'The balanced scorecard: measures that drive performance', *Harvard Business Review*, vol. 83, No. 7, pp. 172–180.

[31] Klazinga, N., Stronks, K., Delnoij, D. and Verhoeff, A. (2001) 'Indicators without a cause. Reflections on the development and use of indicators in health care from a public health perspective', *International Journal for Quality in Health Care*, vol. 13, No. 6, pp. 433–438.
[32] Kludacz, M. (2010) 'Model karty dokonań na potrzeby zintegrowanego systemu oceny działalności szpitali publiczych', in Koncepcja sprawozdawczości szpitali na potrzeby zintegrowanego systemu oceny dokonań, Szczecin: Wydawnictwo Naukowe Uniwersytetu Szczecińskiego.
[33] Kludacz, M. (2012) 'Financial dimension of measurement and assessment in the model of performance card for the hospitals', Proceedings in ARSA-Advanced Research in Scientific Areas, 1, pp. 613–661.
[34] Kontodimopoulos, N., Papathanasiou, N.D., Flokou, A., Tountas, Y. and Niakas, D. (2011), 'The impact of non-discretionary factors on DEA and SFA technical efficiency differences', *Journal of Medical Systems*, vol. 35, No. 5, pp. 981–989.
[35] KPMG (2008) 'Briefing for audit committee members of not-for-profit healthcare systems and hospitals: Industry Insights', [Online], Available: http://www.doc88.com/p-659204461078.html [March 2016].
[36] Kruk, M.E. and Freedman, L.P. (2008) 'Assessing health system performance in developing countries: a review of the literature', *Health Policy*, vol. 85, No. 3, pp. 263–276.
[37] Kurien, G.P. and Qureshi M.N. (2011) 'Study of performance measurement practices in supply chain management', *International Journal of Business, Management and Social Sciences*, vol. 2.4, pp. 19–34.
[38] Lachmann, M., Trapp, R. and Wenger, F. (2015) 'Performance Measurement and Compensation Practices in Hospitals: An Empirical Analysis in Consideration of Ownership Types', *European Accounting Review*, [Online], Available: http://www.tandfonline.com/doi/abs/10.1080/09638180.2014.994541?journalCode=rear20 [March 2016].
[39] Le Pogam, M.A., Luangsay-Catelin, C. and Notebaert, J.F. (2009) 'La performance hospitalière: à la recherche d'un modèle multidimensionnel cohérent', *Management & Avenir*, vol. 25, No. 5, pp. 116–134.
[40] Lowe, D.J., Carmona-Moreno, S. and Reckers, P.M. (2011), 'The influence of strategy map communications and individual differences on multidimensional performance evaluations', *Accounting and Business Research*, vol. 41, No. 4, pp. 375–391.
[41] Mandl, U., Dierx, A. and Ilzkovitz, F. (2008) 'The effectiveness and efficiency of public spending' No. 301. Directorate General Economic and Monetary Affairs (DG ECFIN), European Commission.
[42] Mohan, D.R., Kumar, K.S. (2011) 'A study on the satisfaction of patients with reference to hospital services', *ZENITH. International Journal of Business Economics and Managemnt Research*, vol. 1 no.3, pp. 15–25.

[43] Moreira, S. (2008) 'Efficiency analysis of public hospitals transformed into public corporations: an application of data envelopment analysis', *Economic Bulletin*, vol. 14, No. 1, [Online], Available: https://www.bportugal.pt [April 2016].
[44] Nelson, E.C., Rust, R.T., Zahorik, A. and Rose, R.L. (1992), 'Do patient perceptions of quality relate to hospital financial performance?', *Marketing Health Services*, vol. 12, No. 4.
[45] Niven, P.R. (2011) Balanced scorecard: Step-by-step for government and nonprofit agencies, John Wiley & Sons: Hoboken, New Jersey.
[46] OECD (2001) OECD Productivity manual: A guide to the measurement of industry-level and aggregate productivity growth, [Online], Available: http://www.oecd.org/subject/growth/an_ec_gr.htm [April 2016]
[47] Oszustowicz, R.J., (1992) 'Quality of care emerges as a determinant of creditworthiness, *Healthcare Financial Management*, vol. 46, No. 3, pp. 46–58.
[48] Phua, K.-L. (2014) 'Measuring Health System Performance: A Critique, with Proposals for Improving on the WHO. Building Blocks Framework', *Open Public Health Journal*, vol. 7, pp. 12–15.
[49] Santiago, J.M. (1999), 'Use of the balanced scorecard to improve the quality of behavioral health care', *Psychiatric Services*, vol. 50, No. 12, pp. 1571–1576.
[50] Shaller, D. (2007) Patient-centered care: what does it take?, Commonwealth Fund, [Online], Available: http://www.commonwealthfund.org/ [April 2016]
[51] Shwartz, M., Ren, J., Peköz, E.A., Wang, X. and Cohen, A.B. (2008) 'Restuccia JD. Estimating a composite measure of hospital quality from the Hospital Compare database: differences when using a Bayesian hierarchical latent variable model versus denominator-based weights', *Medical Care*, vol. 46, pp. 778–785.
[52] Shaw, C. (2003) How can hospital performance be measured and monitored?, World Health Organization Regional Office for Europe, Copenhagen, WHO Regional Office for Europe, Health Evidence Network report, [Online], Available: http://www.euro.who.int/document/e82975.pdf. [April 2016]
[53] Staiger, D.O., Dimick, J.B., Baser, O., Fan, Z. and Birkmeyer, J.D. (2009), 'Empirically derived composite measures of surgical performance', *Medical Care*, vol. 47, pp. 226–233.
[54] Ştefănescu, A., Ţurlea, E. and Calu, D. A. (2011) 'Performance measurement in Romanian public hospitals', *African Journal of Business Management*, vol. 5, No. 2, pp. 552–576.
[55] Steinmann, L., Dittrich, G., Karmann, A., and Zweifel, P. (2004), 'Measuring and comparing the (in) efficiency of German and Swiss hospitals', *European Journal of Health Economics*, vol. 5(3), pp. 216–226.
[56] Stausberg, J., Halim, A. and Färber, R. (2011)'Concordance and robustness of quality indicator sets for hospitals: an analysis of routine data', *BMC Health Services Research*, vol. 11, [Online], Available: http://bmchealthservres.biomedcentral.com/articles/10.1186/1472-6963-11-106 [April 2016].

[57] Tawfik-Shukor, A.R., Klazinga, N.S., and Arah, O. (2007) 'Comparing health system performance assessment and management approaches in the Netherlands and Ontario, Canada', *BMC Health Services Research*, vol. 7, No. 1, pp. 25.

[58] Thakkar, J., Kanda, A. and Deshmukh, S.G., (2009) 'Supply chain performance measurement framework for small and medium scale enterprises', *Benchmarking: An International Journal*, vol. 16, No. 5, pp. 702–723.

[59] World Health Organization (2000a) WHOTERM Quantum Satis. A Quick Reference Compendium of Selected Key Terms Used in the World Health Report.

[60] World Health Organization (2000b) The World health report 2000 – Health systems: improving performance. Geneva, Switzerland.

[61] World Health Organization (2007) Performance Assessment Tool for Quality Improvement in Hospitals, Indicator Descriptions, WHO Regional Office for Europe

[62] Wulong, G. and Morin, S. (2014) 'Experimental Measures of Output and Productivity in the Canadian Hospital Sector, 2002 to 2010', Statistics Canada, Economic Analysis Division.

[63] Veillard, J., Champagne, F., Klazinga, N., Kazandjian, V., Arah, O.A. and Guisset, A.L. (2005), 'A performance assessment framework for hospitals: the WHO regional office for Europe PATH project', *International Journal for Quality in Health Care*, vol. 17, No. 6, pp. 487–496.

[64] Villa, S. and Kane, N. (2013), 'Assessing the impact of privatizing public hospitals in three American states: implications for universal health coverage', *Value in Health*, vol. 16, No. 1, pp. S24–S33.

[65] Veillard, J., Guisset, A.L. and Garcia-Barbero M., (2004) 'Selection of indicators for hospital performance measurement: a report on the 3rd and the 4th workshop', World Health Organization, Barcelona, Spain.

Managerial Economics
2016, vol. 17, no. 1, pp. 123–148
http://dx.doi.org/10.7494/manage.2016.17.1.123

Milena Suliga*

The reaction of investors to analyst recommendations of stocks listed on the WIG20 index

1. Introduction

This paper studies the impact of analyst recommendations of markings from the most-fluent stocks quoted on the Warsaw Stock Exchange. The main goal of the research is to check which of the recommendations significantly influence stock prices and what factors affect price reaction (besides the level of recommendation).

Among other information, analyst recommendations set a very interesting group. Assuming that recommendations are formulated on the grounds of publicly available information, they should not set unexpected events. Nonetheless, a variety of brokerage firms and houses spend a lot of money and effort on working out detailed recommendations of stocks. This can suggest that professional analysts employing advanced analytical methods in constructing their assessments usually have greater knowledge than ordinary investors, and their recommendations can be valuable to others. An additional question that could be researched is whether or not the integrity of analysts who have business connections with specific issuers may intentionally spread manipulated information. Moreover, after compiling (but before publishing) a recommendation, analysts can be considered as insiders who possess some privileged information. The presence of excess returns directly before the publication can be interpreted as a sign of potential illegal insider trading. All of the aspects mentioned above make analyst recommendations an interesting subject of study.

The impact of analyst recommendations of stock prices has been researched by many authors, who have also endeavored to locate factors that influence the strength, direction, and duration of such an impact. Among others, Elton et al.

* AGH University of Science and Technology, Faculty of Management, Department of Applications of Mathematics in Economics, e-mail: msuliga@zarz.agh.edu.pl

(1986), Stickel (1995), Womack (1996), Lin and McNichols (1998), Barber et al. (2001), Yazici and Muradoglu (2002), Jegadeesh et al. (2004), Jegadeesh and Kim (2006), Gurgul and Majdosz (2004), Glezakos (2007), Loh and Stulz (2009), Jegadeesh and Kim (2010), Murg et al. (2014), and Murg and Zeitlberger (2014) have studied this subject based on daily data. Kim et al. (1997), Busse and Green (2002), and Green (2006) focus on intraday returns, checking the occurrence of instantaneous market reactions to analyst recommendations. The results received by the authors vary depending on the researched market, the period under study, and the research approach.

The aforementioned studies suggest that analyst recommendations significantly influence the price-formation process on developed markets, and the direction of price changes is consistent with the information contained in the news. BUY recommendations are followed by positive abnormal returns, while SELL recommendations lead to negative abnormal returns observed on the day of the publication or the following day (Stickel, 1995; Womack, 1996; Jegadeesh and Kim, 2006; Murg and Zeitlberger, 2014). Beyond an immediate reaction to the recommendation, some researchers find a drift in prices that continues in the subsequent months after the event (Womack, 1996; Jegadeesh and Kim, 2006). Reactions of small and emerging markets are not so unequivocal. Glezakos (2007) and Gurgul and Majdosz (2004) do not find any visible reaction to recommendations on the Athens Stock Exchange and Warsaw Stock Exchange, respectively. Jegadeesh and Kim (2006) confirm a dependence on the strength of the reaction and the size of the market, convincing us that a developed market reacts stronger, while Murg and Zeitlberger (2014) show the opposite (via examples of the Austrian and German stock markets): abnormal returns on the smaller market are higher.

Jegadeesh et al. (2004) suggest that the research should not only involve the level of recommendation but also the change from its previous level, as it has more robust explanatory power than the level alone. The impact of recommendation is especially strong when there are the most-extensive changes, such as an upgrade to BUY from SELL or a downgrade to SELL from BUY (Stickel, 1995; Murg et al., 2014; Murg and Zeitlberger, 2014). Some authors consider additional factors that influence investor reaction; one of which the size of the company. Results obtained by Womack (1996), Barber et al. (2001), Murg and Zeitlberger (2014), and Murg et al. (2014) suggest that the strongest reaction to recommendations is observed in the case of small- and medium-sized companies.

On the U.S. equity market, Womack (1996) analyzes long-run returns for BUY and SELL recommendation changes stratified by the size of the firm (measured by market capitalization) and graphically shows the differences among companies of different sizes. Also on the U.S. equity market, Barber et al. (2001) group firms into portfolios according to their consensus analyst recommendations and focus on the profitability of investment strategies that involve this consensus. As a part of the research, they partition portfolios according to the size of the firm. They find that the difference between returns for the most highly rated and least-favorably recommended stocks is the most expressive in the group of small- and medium-sized firms. Studying companies from the Austrian and German equity markets, Murg and Zeitlberger (2014) construct a linear regression model with the absolute abnormal return as the independent variable and the market capitalization (MC) of the firm as an explanatory variable representing the size of the firm. All MC coefficients are negative, suggesting that the reaction of smaller companies is stronger. A similar analysis is performed on the Austrian market by Murg et al. (2014), who consider a model with the same endogenous variable but with more exogenous variables; one of which being MC. They also find that analyst recommendations have a stronger impact on smaller firms. As mentioned by Barber et al. (2001), this relationship between company size and investor reaction is reasoned. The smaller is the firm, the less informed the public is about it (and as a result, the more unexpected the recommendation).

Another factor that can potentially influence investor reaction is the reputation of the brokerage house issuing the recommendation. Stickel (1995) and Jegadeesh and Kim (2006) verify that the reaction is stronger when the recommendation is prepared by analysts with better reputations and from larger brokerage houses for stocks from U.S. market and G7 countries, respectively. To that end, the authors implement different regression models with categorical variables. Some other authors conclude that leading analysts avoid giving a SELL recommendation to maintain good relations with clients. Moreover, they recommend BUY too often when they have some business connections with a specific issuer (Womack, 1996; Lin and McNichols, 1998; Barber et al., 2005). This leads to an asymmetric reaction of the market to different types of recommendations. Lin and McNichols (1998) researched that, on the US market, a HOLD recommendation leads to a negative reaction of the market when it is suggested by leading analysts, because investors suspect that SELL should be warranted. Barber et al. (2005) prove that the implementation of NASD Rule 2711 in 2002 (which obligates analysts to display the percentage of particular types of recommendation [BUY, HOLD, or SELL]) effectively reduced the number of BUYs (which had previously been too many).

From the Efficient Market Hypothesis point of view, most of the studies demonstrate that the market is not semi-strong form efficient in the researched countries (not only because there is a post-recommendation drift in prices). On some markets, significant abnormal returns are observed directly before the publication of a recommendation, which suggests some information leakage and the presence of preferred customers (Yazici and Muradoglu, 2002; Gurgul and Majdosz, 2004; Murg and Zeitlberger, 2014).

The impact of analyst recommendations on stocks listed on the Warsaw Stock Exchange is researched inter alia by Gurgul and Majdosz (2004), Mielcarz et al. (2007), Podgórski and Mielcarz (2008), and Buzała (2012). All of the aforementioned authors make an inference based on event-study methodology but assuming different models. Conclusions derived from their studies are not consistent.

Gurgul and Majdosz (2004) consider the sample of 139 BUY, HOLD, and SELL recommendations for stocks listed on the WIG20 Index published from 1996 to 2003. They do not observe any significant abnormal returns after publications regardless of the type of recommendation, but significantly negative average excess returns appear directly before the publications of SELL recommendations. This can be interpreted as some leakage of information. Mielcarz et al. (2007) analyze only positive recommendations (BUY and ACCUMULATE) from the 1st of January 2005 to the 31st of December 2006. Their sample consists of 246 events. In contrast to Gurgul and Majdosz (2004), they determine two recommendation issue dates: the day when the recommendation was provided to commercial customers and the day when the recommendation became public on the Internet. Their results are opposite to those obtained by Gurgul and Majdosz (2004) and show that BUY recommendations create significantly positive abnormal returns on both assumed days of the event. According to this research, the second type of positive recommendation (ACCUMULATE) has no impact on stock prices. Podgórski and Mielcarz (2008) study only neutral and negative recommendations (for a change); cumulatively, 319 events. They analyze the same period as Mielcarz et al. (2007) and define issue dates likewise. They conclude that negative events show an immediate reflection in stock prices, as significantly negative abnormal returns are observed both on the day of delivering the recommendation to commercial customers and the day of the official publication on the Internet. They do not find any influence of neutral recommendations on the price-formation process.

Another study in the area of recommendation impact on stock prices on the Warsaw Stock Exchange is the research of Buzała (2012). 1185 recommendation dates gathered by the author cover the period from January 2010 to December 2012. Recommendations concerning stocks listed on the WIG20 Index are systematized in a five-point scale: BUY, ACCUMULATE, HOLD, REDUCE, and SELL. Issue dates are defined in two ways, like in Mielcarz et al. (2007). In addition to the research of price reaction to the five aforementioned types of recommendations, Buzała (2012) checks if changes in the level of recommendation are informative. Thereupon, he distinguishes among event upgrades and downgrades. In the group of upgrades, he additionally specifies upgrades to BUY, and in the group of downgrades, he itemizes downgrades to SELL. His findings indicate that both a first issue of an extreme recommendation (BUY or SELL) to a limited group of customers as well as its official publication influence stock prices. For other types of recommendations, there are no visible price reactions. Significant price changes also appear in groups of upgrades and downgrades. The strongest reaction can be observed in groups of upgrades to BUY and downgrades to SELL. In the case of recommendations that are initially delivered to a narrow audience, the second public issue is less significant. Post-event drift appears in the cases of the most-extreme recommendations (BUY and SELL) as well as in the situation of level changes. Significant abnormal returns consistent with the direction of the recommendation appear before the day of the event, even in the case of the first issue. This is an argument for some leakage of information that leads to a rejection of the semi-strong effectiveness of the Polish equity market.

In this paper, the effect of analyst recommendations on security prices is explored. Recommendations that are researched involve the most-fluent companies quoted on the Warsaw Stock Exchange from January 2012 to September 2015. The first part of the study concerns price reaction to positive, neutral, and negative recommendations. Subsequently, the impact of changes in recommendation levels on investor reaction is tested. The tests are conducted with the use of the event-study methodology. Finally, additional factors that could affect price reactions are studied.

The study extends foregoing research from the WSE relating to this subject. First, it covers the period that has not been the object of such a study as of yet, as the last known research on this issue was conducted by Buzała (2012) and covers the period of 2010–2011. (this period is analyzed in additional research to check if the employed method gives results similar to those acquired by Buzała). Second, it examines not only investor reaction to a given level of recommendation but also to level changes. To the best of this author's knowledge, only the research of Buzała attempted to investigate this issue until now. However, this research is more detailed than that made by Buzała, as it regards various possible level changes, not only upgrades and downgrades. Moreover, all of the aforementioned authors employing event-study methodology use parametrical test statistics. The validity of the parametrical test is addicted to the fulfillment of strict assumptions that very often go unfulfilled. In this research, the problem is resolved by subsidiary use of a non-parametrical rank test, which requires less-restrictive assumptions. Lastly, as distinct from previous studies, additional factors that could influence investor reaction are researched. Next to changes in recommendation levels, the size of the company and reputation of the brokerage house are analyzed as potential explanatory variables. These factors are pointed out as influential by research from other countries. Conterminously with Murg and Zeitlberger (2014) and Murg et al. (2014), a linear regression model is employed, but it differs significantly from those used by the aforementioned authors. First and foremost (as the information is not immediately reflected in security prices), a proper independent variable is chosen. Furthermore, the model contains different dependent variables; particularly, the size of the firm is represented by the natural logarithm of company shares in the WIG20 Index.

The structure of the paper is as follows: Section 2 describes the data and methodology applied in the research; empirical results are demonstrated and discussed in Section 3; and Section 4 concludes the paper. The appendix contains results from the period of 2010–2011 as a comparison to the research conducted by Buzała (2012).

2. Data and methodology

The dataset contains 576 analyst recommendations for 28 companies listed in the WIG20 Index from the 1st of January 2012 to the 1st of September 2015. Recommendations are taken from the database available at www.bankier.pl, which contains information about the date of the publication, level of the recommendation, actual price, target price and price at date of publication, change-in-price potential, and name of the issuing institution.

Recommendations can be expressed in one of two scales. An absolute recommendation is prepared by analysts who estimate a target price and then compare it with the current market value. Then, on the basis of underestimation or overestimation, they formulate a recommendation such as BUY, ACCUMULATE, HOLD (or NEUTRAL), REDUCE, or SELL. When the target price is compared with the current valuation of companies from a specified segment or comparative group, the recommendation is then relative. Possible recommendations in this group are in line with the market, below the market, above the market, or similar. From the initial set of 1470 recommendations, 1245 absolute recommendations were selected. The number was then reduced to 576 by removing the events that were too close in terms of the event-study methodology. Recommendations qualified for the final sample were categorized into three major groups: positive (BUY or ACCUMULATE), neutral (HOLD), and negative (REDUCE or SELL). Table 1 contains the number of recommendations in each of these groups.

To test if analyst recommendations have an impact on security prices, the event-study methodology was applied. Daily log-returns of the stocks were employed, computed from closing prices available on www.gpwinfostrefa.pl. Since the data does not contain information about whether or not a particular recommendation was previously delivered to privileged customers, the date of the event (designated by $t = 0$) is defined as the day of the publication named in the database. For each of the recommendations, pre-event and event windows are defined. Following others

(Gurgul and Majdosz, 2004; Murg, Zeitlberger, 2014), we define a pre-event window that covers 30 days (from $t = -35$ to $t = -6$). The event window contains 11 days around the date of the event. It starts five days before the publication ($t = -5$) to capture potential price changes that could indicate information leakage and end five days after it ($t = 5$) to evaluate the speed of price reaction.

Table 1

The number of absolute analyst recommendations included in the study. All recommendations relate to companies listed in the WIG20 Index from the 1st of January 2012 to the 1st of September 2015

Year	Total number of recommendations	Positive		Neutral	Negative	
		BUY	ACCUMULATE	HOLD	REDUCE	SELL
2012	135	40	9	54	9	23
2013	145	33	5	71	7	29
2014	170	66	5	66	6	27
2015	126	24	9	61	3	29
SUM	576	163	28	252	25	108

For each day in the pre-event and event windows, abnormal returns are calculated as the difference between the actual return and its expected value:

$$AR_{i,t} = R_{i,t} - E\left(R_{i,t}\right)$$

$R_{i,t}$ is a logarithmic rate of return of *i*-th company on day t. Expected returns are calculated with the classical market model from the estimation window:

$$R_{i,t} = \alpha + \beta R_{m,t} + \varepsilon_{i,t}$$

where $R_{m,t}$ is a logarithmic rate of market return represented by WIG20 return and $\varepsilon_{i,t}$ is an error on a given day. Parametric tests drawn on the model assume a normal distribution of residuals, lack of the autocorrelation, and homoskedasticity. Cumulatively, 24% of the data does not satisfy at least one of these assumptions.

Despite the fact that assumptions are not achieved in all cases, estimators received from the least-squares method are consistent. Nonetheless their effectiveness may be not satisfactory (Gurgul, 2006), and the applied parametrical test may not have assumed t-distribution. Previous research shows that applying more-complicated models does not necessarily improve results of the classical market model (see Brown and Warner, 1980; Murg et al., 2014). To support the results of the parametric test, a non-parametric rank test (which does not require the assumption of abnormal return normality) is employed. For each event, abnormal returns are divided by the standard deviation from the pre-event window, and thus, are standardized:

$$SAR_{it} = AR_{it} \ / \ S\left(AR_i\right)$$

where

$$S\left(AR_i\right) = \sqrt{\frac{1}{29}\sum_{t=-35}^{-6} AR_{it}^2}$$

Figure 1 shows the cross-sectional variance of standardized abnormal returns separately in the three defined groups of events. One can notice that there is a significant increase of volatility in the event window. This phenomenon is frequently observed (Corrado, 2011) and demands cross-sectional variance adjustment. Following Corrado and Zivney (1992), adjusted standardized abnormal returns are computed; namely:

$$SAR'_{it} = \begin{cases} SAR_{it} & t = -35,\ldots,-6 \\ SAR_{it} \ / \ S\left(SAR_t\right) & t = -5,\ldots,5 \end{cases}$$

where

$$S\left(SAR_t\right) = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}\left(SAR_{it} - \overline{SAR_t}\right)^2}$$

and N is the number of stocks in the sample.

Cross-sectionally adjusted *t*-test statistic has the form as in Corrado and Zivney (1992):

$$T_a\left(t_0\right) = \frac{1}{\sqrt{N}}\sum_{i=1}^{N} SAR'_{it_0}$$

while the non-parametric rank test proposed by Corrado (2011) is defined as:

$$T_{CZ}\left(t_0\right) = \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\frac{\left(\text{rank}\left(SAR'_{it_0}\right) - \frac{n+1}{2}\right)}{\sqrt{n\left(n+1\right)/12}}$$

n is the length of the pre-event window, while $\text{rank}\left(SAR'_{it_0}\right)$ is the rank of the adjusted standardized abnormal return of day t_0 in the group of 31 standardized abnormal returns: 30 from the pre-event window, and the remaining 1 from day t_0 that is SAR'_{it_0}. The T_{CZ} statistic is asymptotically normally distributed.

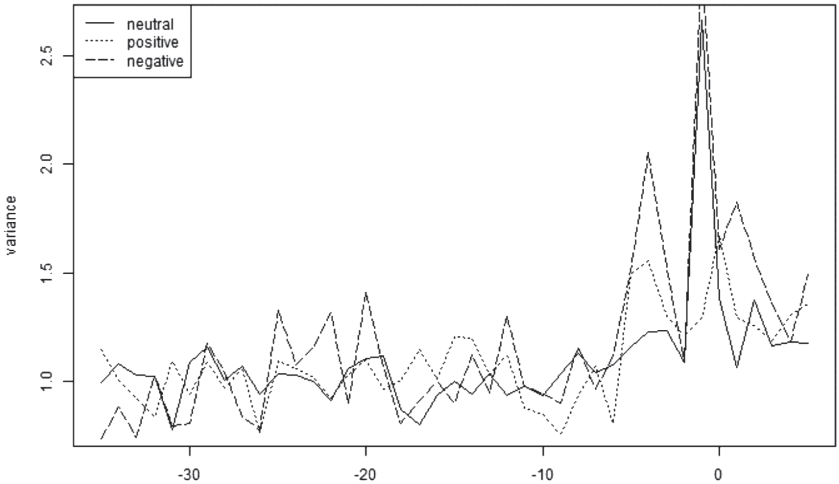

Figure 1. Cross-sectional variance of standardized abnormal returns in the pre-event and event window, separately in three groups of recommendations: positive, neutral, and negative

3. Empirical results

3.1. Results from analysis of the standard levels of recommendations

In the first step, the recommendations were divided into three clusters: positive, neutral, and negative events (as detailed in Table 1). In each of the groups separately, the impact of recommendations on security prices was tested with the parametric *t*-test and nonparametric rank test. Mean-adjusted standardized abnormal returns $\left(\overline{SAR'}\right)$ and *p*-values for both test statistics in the event window are presented in Table 2. The number of events in each group is also detailed in the table.

The strongest reaction of prices can be observed in the case of positive recommendations. Both test statistics are significantly positive at the 5% level on the day of publication and at the 1% level the following day. Two days after the issue of recommendation, prices are still influenced by the event (as suggested by the significance of the statistics at the 5% level). Adjusted standardized abnormal returns remain positive to five days after the event, and the non-parametric test statistic is significant also on the fourth and fifth days after the recommendation. This suggests that the market reacts strongly to positive analyst recommendations. Price adjustment is not immediate, but it lasts a few days. The parametrical test indicates a statistically significant abnormal return five days before the event; but as the sign of the return is negative, this is probably triggered by some confounding events.

In the group of neutral events, abnormal returns remain negative from three days before the recommendation to one day after it. However, test statistics around the event day are insignificant. This can be a signal to a tentative interpretation that hold recommendations have no impact on security prices. Only four days before the event and five days after it is the T_{CZ} statistic significantly positive; but this seems to have no direct connection to the publication of the recommendation.

Negative recommendations are marked with negative abnormal returns that continue from one day before the recommendation to five days after it. Nonetheless, only one of the two test statistics (namely, the parametric *t*-test statistic) is significantly negative at the 5% level on the day of recommendation. The non-parametric T_{CZ} statistic has the lowest value of the post-event window, but it is not significant.

Table 2

Reaction of daily returns to particular types of analyst recommendations (positive, neutral, and negative) issued between the 1st of January 2012 and the 1st of September 2015

t	**Positive recommendation (Buy/Accumulate)**			**Neutral recommendation (Hold)**			**Negative recommendation (Sell/Reduce)**		
	$\overline{SAR'}$	***p*-value**		$\overline{SAR'}$	***p*-value**		$\overline{SAR'}$	***p*-value**	
		T_a	T_{CZ}		T_a	T_{CZ}		T_a	T_{CZ}
	191 events			**252 events**			**133 events**		
–5	–0.176	**0.021**	0.104	–0.007	0.396	0.115	–0.072	0.282	0.389
–4	0.068	0.257	0.082	0.060	0.252	**0.043**	0.017	0.390	0.085
–3	0.039	0.344	0.175	–0.004	0.398	0.228	–0.071	0.285	0.367
–2	–0.056	0.295	0.360	–0.006	0.397	0.382	0.109	0.181	0.054
–1	0.068	0.256	0.099	–0.088	0.150	0.385	–0.062	0.307	0.247
0	0.291	**0.013**	**0.000**	–0.009	0.394	0.489	–0.238	**0.010**	0.112
1	0.210	**0.006**	**0.000**	–0.037	0.336	0.423	–0.153	0.085	0.159
2	0.151	**0.045**	**0.015**	0.017	0.384	0.096	–0.105	0.190	0.124
3	0.011	0.394	0.083	–0.060	0.255	0.368	–0.020	0.388	0.416
4	0.095	0.170	**0.030**	0.042	0.319	0.109	–0.029	0.377	0.494
5	0.076	0.229	**0.026**	0.045	0.307	**0.046**	–0.037	0.363	0.490

The results are contrary to those obtained by Gurgul and Majdosz (2004) but are consistent with the research of Mielcarz et. al (2007) and Podgórski and Mielcarz (2008). The market seems to react more strongly to positive events than to negative ones. Neutral recommendations seem to have no impact on stock prices.

3.2. Results from the analysis of changes in recommendation level

Research conducted by other authors suggests that an analysis of changes in the level of recommendation is much more informative than an analysis of a pure recommendation level (e.g., Stickel, 1995; Womack, 1996; Jegadeesh et al., 2004; Murg and Zeitlberger, 2014). To check this assumption, recommendations are divided into nine clusters (as detailed in Table 3). In 526 out of 576 events, it was possible to assess the previous level of recommendation and define the change. Only changes from one of the three defined groups (positive, negative, neutral) to another are studied (e.g., from positive to negative). Changes inside the group (such as the change from ACCUMULATE to BUY or from SELL to REDUCE) are not researched. The number of recommendations in each of these groups is presented in Table 3. The most numerous is the group of neutral recommendations that were previously also neutral (121 events – 23%). They are followed by positive events with no change (106 events – 20.2%). The least numerous are the most-extensive changes: from a positive to negative recommendation (19 events – 3.6%) and from a negative to positive one (13 events – 2.5%).

Table 3

Recommendation changes. Only changes from one of three defined levels to another one are studied. Changes within the specified group (e.g., the change from SELL to REDUCE) are lost

		Previous recommendation			**Sum**
		positive	neutral	negative	
Current recommendation	positive	106	50	13	169
	neutral	66	121	47	234
	negative	19	45	59	123
	Sum	191	216	119	526

Results from the analysis of recommendation changes are presented in Table 4. The table contains mean-adjusted standardized abnormal returns $\left(\overline{SAR'}\right)$ and *p*-values for both test statistics in the event window for nine groups of events.

It can be noticed that positive recommendations are marked with a positive $\left(\overline{SAR'}\right)$ from the day of the event to two days after it, regardless of the level of the previous recommendation. However, the strongest reaction is observed in the case of recommendations that were also previously positive. The T_{CZ} statistic is significant at the 5% level on the day of the event as well as three other days in the post-event window. The parametric *t*-statistic is significantly positive only two days after the recommendation. This indicates that new information is not immediately incorporated into the price but needs a few days to process. In the group of positive events that were previously neutral, only the nonparametric statistic is significantly positive at the 5% level on the day of the recommendation as well as the following day. Nevertheless, nonparametric test results are more reliable (as pointed out in Section 2). Changes from negative to positive do not show an immediate reaction to the information, as test statistics are insignificant on the days around the event. Only the T_{CZ} statistic is significant at the 5% level four days after the recommendation. However, the size of the group (13 events) is too small to make conclusions on the basis of asymptotic distribution of this statistic. Results from the first three groups of recommendations can be summarized as follows: the better the previous recommendation, the stronger positive reaction to the following one that is positive. This can be a signal that investors do not trust analysts if they change the recommendation upward to a positive one. However, if an analyst issues a second positive recommendation, this is interpreted as a credible sign of good company position.

The impact of the neutral event on security prices varies depending on the level of the previous recommendation. A neutral recommendation preceded by a positive one is interpreted as a strongly negative signal. In this group of events, both test statistics are significantly negative at the 1% level on the day of the recommendation. Nonetheless, returns reveal counter movements in the post-event window, as they are positive on three days. The nonparametric test statistic is significantly positive at the 5% level on the second and fourth days after the event. It seems that investors overreact to this kind of change in a recommendation level. Stickel (1995) and Murg and Zeitlberger (2014) also find overreaction to some changes in the level of recommendation on other stock markets. A second neutral recommendation does not influence security prices. Test statistics are insignificant in the whole event window. As might be expected, a neutral recommendation occurring after negative one is interpreted as positive change. $\overline{SAR'}$ are positive on the day of the event as well as the two following days. The significance of test statistics appears on the first and second days after the recommendation, which indicates that the information is gradually incorporated into the prices. The analysis of neutral recommendations conducted in the previous subsection suggested that a HOLD recommendation is indeed neutral and has no impact on prices. Research of the recommendation changes above shows that such a conclusion is too hasty and that the interpretation of a neutral recommendation depends on the previous recommendation level. If it was lower, the recommendation is interpreted as positive; if it was higher, the recommendation has a negative tenor.

Table 4

Reaction of daily returns to analyst recommendations issued between the 1st of January 2012 and the 1st of September 2015. Groups of events were defined depending on the change in the recommendation level

PART A. Current recommendations: positive (BUY/ACCUMULATE)

t	Previous recommendation								
	Positive (Buy/Accumulate)			Neutral (Hold)			Negative (Sell/Reduce)		
	$\overline{SAR'}$	p-value		$\overline{SAR'}$	*p*-value		$\overline{SAR'}$	*p*-value	
		T_a	T_{CZ}		T_a	T_{CZ}		T_a	T_{CZ}
	106 events			50 events			13 events		
–5	–0.203	**0.046**	0.079	–0.257	0.078	0.130	–0.027	0.389	0.394
–4	0.098	0.238	0.147	0.239	0.097	0.068	–0.079	0.375	0.269
–3	0.070	0.306	0.115	–0.197	0.151	0.105	–0.261	0.247	0.239
–2	–0.035	0.373	0.374	–0.094	0.317	0.399	–0.176	0.316	0.220
–1	0.120	0.184	0.081	0.074	0.345	0.240	–0.139	0.342	0.313
0	0.147	0.127	**0.013**	0.241	0.094	**0.027**	0.077	0.376	0.230
1	0.149	0.123	**0.023**	0.222	0.117	**0.043**	0.235	0.268	0.096
2	0.201	0.048	0.014	0.056	0.367	0.202	0.375	0.156	0.132
3	0.027	0.382	0.237	0.053	0.370	0.091	–0.204	0.294	0.335
4	0.087	0.266	0.071	0.033	0.386	0.487	0.474	0.095	**0.035**
5	0.134	0.155	**0.033**	0.044	0.378	0.202	–0.136	0.344	0.431

Table 4

PART B. Current recommendations: neutral (HOLD)

t	Previous recommendation								
	Positive (Buy/Accumulate)			Neutral (Hold)			Negative (Sell/Reduce)		
	$\overline{SAR'}$	p-value		$\overline{SAR'}$	p-value		$\overline{SAR'}$	p-value	
		T_a	T_{CZ}		T_a	T_{CZ}		T_a	T_{CZ}
	66 events			121 events			47 events		
–5	0.149	0.191	**0.036**	–0.157	0.091	0.209	0.049	0.375	0.115
–4	0.037	0.380	0.217	0.031	0.376	0.276	0.100	0.313	0.067
–3	–0.018	0.393	0.494	–0.026	0.382	0.432	0.036	0.385	0.262
–2	0.089	0.306	0.201	–0.046	0.351	0.308	0.027	0.390	0.330
–1	–0.055	0.359	0.247	–0.110	0.190	0.256	–0.104	0.307	0.424
0	–0.377	0.005	0.001	0.015	0.393	0.215	0.217	0.131	0.135
1	–0.149	0.191	0.283	–0.026	0.382	0.428	0.250	0.092	**0.039**
2	0.168	0.157	**0.038**	–0.172	0.068	0.136	0.350	0.025	0.008
3	–0.039	0.378	0.363	–0.012	0.395	0.276	–0.216	0.133	0.170
4	0.124	0.238	**0.049**	–0.064	0.310	0.301	0.021	0.393	0.289
5	0.063	0.349	0.185	0.010	0.396	0.230	0.040	0.382	0.188

Table 4

PART C. Current recommendations: negative (SELL/REDUCE)

t	Previous recommendation								
	Positive (Buy/Accumulate)			Neutral (Hold)			Negative (Sell/Reduce)		
	$\overline{SAR'}$	*p*-value		$\overline{SAR'}$	*p*-value		$\overline{SAR'}$	*p*-value	
		T_a	T_{CZ}		T_a	T_{CZ}		T_a	T_{CZ}
	19 events			45 events			59 events		
–5	–0.005	0.394	0.443	–0.156	0.228	0.291	–0.074	0.337	0.375
–4	0.158	0.308	0.297	0.012	0.396	0.174	0.131	0.238	0.083
–3	–0.242	0.222	0.315	–0.029	0.389	0.429	–0.087	0.317	0.342
–2	0.312	0.156	0.146	0.085	0.337	0.321	0.082	0.325	0.123
–1	–0.008	0.393	0.362	–0.062	0.363	0.268	–0.042	0.377	0.295
0	–0.359	0.117	0.177	–0.238	0.112	0.246	–0.134	0.233	0.359
1	–0.339	0.132	0.123	–0.290	0.062	0.076	–0.009	0.396	0.337
2	–0.088	0.365	0.423	–0.268	0.080	0.059	0.045	0.374	0.370
3	0.366	0.112	0.061	–0.317	0.044	0.020	0.091	0.311	0.214
4	0.352	0.122	0.057	–0.157	0.226	0.174	–0.016	0.394	0.485
5	–0.078	0.371	0.362	0.095	0.323	0.221	–0.113	0.271	0.331

Negative recommendations are, on the whole, marked with a negative $\overline{SAR'}$ on the day of the event as well as the following day. In the group of changes from a positive recommendation to a negative, there are no significant values of test statistics. This group with the most extensive changes is simultaneously one of the least numerous, which precludes us from drawing reliable conclusions. In the case of negative recommendations preceded by neutral ones, $\overline{SAR'}$ are negative to the fourth day after the recommendation. On day $t = 3$, both test statistics are significantly negative at the 5% level. Negative information about the downgrade of a recommendation level seems to need a few days to be incorporated into the price. A second negative recommendation does not significantly influence returns, which suggests that investors expected the repetition of negative information about the security.

Analysis of recommendation changes confirms the assumption that changes in the level of recommendation poses more-important information for investors than a pure recommendation level. Downgrades are interpreted as negative events and generally lower security prices, while upgrades constitute positive information and lead to an increase in returns. This finding is consistent with the research of Buzała (2012). Excluding positive recommendations, events marked with no recommendation change have no impact on prices.

As an addendum of this research, an analogous one was conducted for the period analyzed by Buzała (2012), namely 2010–2011. The results are presented in the appendix.

3.3. Results from the linear model with categorical variables

To research the occurrence of other factors that could influence a price reaction to the recommendation, a linear regression model with categorical variables is fitted as follows:

$$CAR_i(0,2) = \alpha_0 + \alpha_{1,j_1} \cdot \chi_{j_1}(CHANGE) + \alpha_{1,j_2} \cdot \chi_{j_2}(CHANGE) + \alpha_2 \cdot SHARE + \\ + \alpha_{3,k_1} \cdot \chi_{k_1}(RANKING) + \alpha_{3,k_2} \cdot \chi_{k_2}(RANKING)$$

Since price reaction to the recommendation can be generally noticed not only on the day of the event but also on the following days, the dependent value in the model is chosen as $CAR_i(0,2)$, which stands for cumulative abnormal returns from the day of the event to the second day after it:

$$CAR_i(0,2) = \sum_{i=0}^{2} AR_{it}$$

CHANGE is the variable defining the change in the recommendation level. When we assign a value of 1 to a positive recommendation, a value of 0 to a neutral recommendation, and a value of 2 to a negative recommendation, then CHANGE ranges from form –2 to 2 (e.g., CHANGE = –2 in the case of a positive recommendation preceded by a negative one). Only the movements between the three defined groups of positive, negative, and neutral events are analyzed (e.g., if the current recommendation is BUY and the previous was ACCUMULATE, the variable takes a value of 0). The variable SHARE stands for the natural logarithm of the company's share in the WIG20 Index. RANKING is the variable that determines the reputation of the brokerage house issuing the recommendation. The variable is set to 2 when the brokerage firm is in the top position (from first to tenth) in both of two annual rankings conducted by Polish Forbes Magazine: the ranking created by individual investors and that made by institutional investors. A value of 1 is assigned to firms that are in the top position in only one of those two rankings. When the brokerage firm is not listed in either of the aforementioned two rankings or is positioned outside the top ten in both of them, the variable takes a value of 0. Function $\chi_i(x)$ means an indicator function that is:

$$\chi_i(x) = \begin{cases} 1, & x = i \\ 0, & x \neq i \end{cases}$$

The model is fitted separately to positive, neutral, and negative recommendations. In effect, variable CHANGE takes only three values in each of the groups.

The model researches significant changes in the levels of the categorical variable with reference to one fixed level. For this reason, two sub-models should be employed in each group of events to capture the potential significance of all possible level changes.

Intercept α_0 corresponds to the situation in which both categorical variables are on the reference level. Then, coefficients α_{1,j_1} and α_{1,j_2} inform us how the intercept changes when CHANGE goes to levels j_1 and j_2, respectively. Analogously, coefficients α_{3,k_1} and α_{3,k_2} represent the change in the intercept when variable RANKING goes from the reference level to levels k_1 or k_2.

Fitted linear models are presented in Table 5. In the group of positive events, two coefficients are statistically significant at the 1% level. Intercepts suggest that positive recommendations are generally seen as good news. A negative coefficient related to a variable SHARE implicates that the smaller the firm, the stronger its reaction to the recommendation. Neither the reputation of the brokerage house nor change in the level of recommendation seems to have a significant impact on the studied cumulative abnormal returns. Linear models do not confirm the remark from the event-study analysis that the reaction is the strongest in the case of the repetition of a positive recommendation.

Table 5

Coefficients of linear regression models with categorical variables fitted to positive, neutral, and negative recommendations separately. In each group, two sub-models were employed to capture potential significance of all possible level changes

Positive recommendations					
Reference levels: CHANGE: 0, RANKING: 0.			Reference levels: CHANGE: 2, RANKING: 2.		
Coefficient	**Estimate**	**p-value**	**Coefficient**	**Estimate**	**p-value**
α_0	0.0228**	0	α_0	0.0245**	0.0046
$\alpha_{1,1}$	–0.0001	0.9787	$\alpha_{1,0}$	–0.0028	0.7344
$\alpha_{1,2}$	0.0028	0.7321	$\alpha_{1,1}$	–0.0029	0.7321
α_2	–0.0087**	0.0005	α_2	–0.0087**	0.0005
$\alpha_{3,1}$	–0.0017	0.07219	$\alpha_{3,0}$	0.0011	0.8337
$\alpha_{3,2}$	–0.0011	0.8337	$\alpha_{3,1}$	–0.0006	0.9135
Multiple R^2: 0.0759					
Neutral recommendations					
Reference levels: CHANGE: –1, RANKING: 0.			Reference levels: CHANGE: 1, RANKING: 2.		
Coefficient	**Estimate**	**p-value**	**Coefficient**	**Estimate**	**p-value**
α_0	–0.0063	0.2274	α_0	0.0078	0.1967
$\alpha_{1,1}$	0.0002	0.9580	$\alpha_{1,-1}$	–0.0202**	0.0007
$\alpha_{,2}$	0.0202**	0.0007	$\alpha_{1,0}$	–0.0200**	0.0003
α_2	0.0036	0.1159	α_2	0.0036	0.1159
$\alpha_{3,1}$	–0.0023	0.6072	$\alpha_{3,0}$	0.0061	0.2675
$\alpha_{3,2}$	–0.0061	0.2675	$\alpha_{3,1}$	0.0037	0.5082
Multiple R^2: 0.0714					
Negative recommendations					
Reference levels: HANGE: –2, RANKING: 0.			Reference levels: CHANGE: 0, RANKING: 2.		
Coefficient	**Estimate**	**p-value**	**Coefficient**	**Estimate**	**p-value**
α_0	–0.0432**	0.0003	α_0	–0.0099	0.2793
$\alpha_{1,-1}$	–0.0022	0.1742	$\alpha_{1,-2}$	–0.0135	0.1742
$\alpha_{1,0}$	0.0135	0.8304	$\alpha_{1,-1}$	–0.0157*	0.0379
α_2	0.0144**	0.0008	α_2	0.0144**	0.0008
$\alpha_{,1}$	0.0160*	0.0428	$\alpha_{3,0}$	–0.0198*	0.0393
$\alpha_{,2}$	0.0198*	0.0393	$\alpha_{3,1}$	–0.0038	0.6670
Multiple R^2: 0.1312					

*, ** – significant at the 1% and 5% levels, respectively

In the group of neutral events, variable CHANGE is the only one explanatory variable that influences cumulative abnormal returns. Coefficients related to this variable are significant at the 1% level. Fitted models affirm the results from the previous subsection: the perception of neutral recommendation depends on the previous level of the recommendation. There is a significant difference in investor reaction to a HOLD recommendation when it represents a decrease from BUY or ACCUMULATE as compared to the situation when it states an increase from SELL or REDUCE. Firm size represented by its share in the WIG20 Index as well as the reputation of the brokerage firm seem to have no explanatory power.

The most statistically significant coefficients appear in models fitted to the group of negative recommendations. A negative intercept in the model with CHANGE at a reference level equal to −2 and with variable RANKING at level 0 indicate that, in such a situation, SELL or REDUCE induces a negative price reaction. Assuming that negative recommendations are generally associated with negative returns, coefficient α_2 that is positive and statistically significant at the 1% level confirms the hypothesis that smaller firms react more strongly (the bigger is the firm, the less negative the cumulative abnormal returns). Coefficients related to variable RANKING are statistically significant at the 5% level and suggest that the reputation of the brokerage house is an important factor that influences reaction to the recommendation. It does not matter whether the brokerage firm takes the top place in only one of the two considered rankings or in both of them, but the reaction is significantly different if the firm is not one of those with the best reputation. Interestingly, cumulative abnormal returns are more negative (the reaction is stronger) when the recommendation is prepared by firms with worse reputations. These findings do not confirm the results from other markets (Stickel, 1995; Jegadeesh and Kim, 2006). Since the relationship between price reaction and reputation of the brokerage firm was not proven in the two aforementioned groups of events, an unambiguous conclusion cannot be drawn. A more-detailed analysis of this topic should be conducted. In the model with variable CHANGE at a reference level equal to 0 and variable RANKING equal to 2, one more statistically significant coefficient appears; namely $\alpha_{1,-1}$. Coefficient $\alpha_{1,-2}$ is not significant, which indicates that the reaction to a negative recommendation is similar when the previous recommendation was positive and when it was neutral. If, however, the previous recommendation was also SELL or REDUCE, the negative reaction is significantly weaker. Results from the event-study analysis conducted in the previous subsection also show a dissimilarity of investor reaction to negative events according to the level of the previous recommendation. Conclusions were slightly different, but it should be noticed that the test statistics were significant on the third day after the event, while in regression models, only two days after the recommendation are counted.

To sum up, results from the linear regression model are generally consistent with those from the event-study analysis. Additionally, they confirm the findings from other markets (Womack, 1996; Barber et al., 2001; Murg and Zeitlberger, 2014; Murg et al., 2014) that the reaction to the recommendation depends on the size of the company and is stronger in the case of smaller firms. A direct and clear relationship between brokerage house reputation and reaction to the recommendation was not found.

4. Conclusions

In this paper, the impact of analyst recommendations to the prices of stocks listed on the Polish WIG20 Index is researched. The data covers the period of January 2012 to September 2015. To investigate the strength and the speed of the reaction, an event-study analysis is employed. Abnormal returns are calculated with the use of a market model based on daily data. Additionally, a linear regression model with categorical variables is employed to investigate potential factors that influence investor reaction to recommendations.

An initial study of recommendations clustered in three levels (positive, neutral, and negative) shows that the sign of abnormal returns around the day of the event is generally consistent with the information contained in the recommendation, and that the strongest reaction is observed in the group of positive recommendations. Test statistics in this cluster are significant to the fifth day after the event, which suggests that prices do not incorporate information immediately. In the case of negative events, abnormal returns remain negative to the fifth day after the recommendation. Nonetheless, the reaction is weaker than in the case of positive events, as most of these abnormal returns are statistically insignificant.

The analysis of level changes confirms the hypothesis that change in the level of recommendation (regarding the previous one) is the informative factor that influences investor behavior and, as a consequence, stock prices. Reaction to a positive recommendation is the strongest when the recommendation was also previously positive. This can be a sign that investors trust analysts more when they confirm their previous positive recommendation than when they raise their recommendation to positive. When the initial analysis does not prove any significant reactions to neutral recommendations, the study of level changes shows that reaction to this kind of event depends strongly on the level of the previous recommendation. HOLD preceded by SELL or REDUCE sets a positive event, while a previous recommendation of BUY or ACCUMULATE results in a reaction to a subsequent HOLD that is negative.

Negative recommendations do not influence stock prices as strongly as positive ones (although, they generally lead to negative abnormal returns). Statistically, nonzero abnormal returns three days after the event are found only in the group of negative events that were previously neutral.

A linear regression model with categorical variables employed separately to positive, neutral, and negative events confirms the conclusions from the event study. Particularly, the model demonstrates that the level of the previous recommendation has a strong impact on the reaction to a neutral recommendation. Furthermore, the model discloses that the size of the firm (represented by its shares in the WIG20 Index) sets an important factor in the analysis of reaction to a recommendation, as the reaction is stronger when the company is smaller. A counter-intuitive relationship between brokerage firm reputation and the reaction to a recommendation was found in the group of negative events. Results suggest that a worse reputation leads to a stronger price reaction. Since such a relationship does not appear in other groups of events, the subject needs more detailed research. A direct and clear conclusion that the reputation of a brokerage house influences reaction to its recommendation cannot be drawn.

5. Appendix

5.1. Study of the period previously researched – comparison of applied methods

To check whether the method applied in this research supports the results from foregoing studies, the period of 2010–2011 previously analyzed by Buzała (2012) is additionally researched. The dataset contains 441 absolute analyst recommendations for 23 companies listed in the WIG20 Index from the 1st of January 2010 to the 31st of December 2011.

Despite the fact that expected returns are calculated identically as those made by Buzała (2012) (with the classical market model), the studies differ significantly. Buzała (2012) analyzes average cumulative abnormal returns (ACAR) from 15 days before the recommendation to 30 days after it and checks their statistical significance. He considers nine types of events: five of them represent the recommendation level (SELL, REDUCE, HOLD, ACCUMULATE, BUY), and the other represent level changes: upgrades, downgrades, and additionally specified upgrades to BUY and downgrades to SELL. His sample is much more numerous (1185 recommendations). In this study, the sample is cut down to 441 events to eliminate those that are too close in terms of event study methodology.

5.2. The analysis of three standard recommendation levels

The results from the analysis of three standard recommendation levels are presented in Table 6. For positive and negative events, the immediate reaction of the market can be observed, which is consistent with the information contained in the recommendation. Both test statistics are significantly positive at the 1% level on the day of the publication. For positive recommendations, the non-parametric statistic in also significant at the 1% level on the following day and five days before the event. This positive reaction before publication of the recommendation could be a signal of potential information leakage, but from fourth to second days before the event, mean adjusted standardized abnormal returns $\left(\overline{SAR'}\right)$ are negative (one of them is even statistically significant in terms of the parametrical test), which suggests that there could be some confounding events. For negative recommendations, all $\overline{SAR'}$ in the event window are negative; but besides the day of the recommendation, a non-parametric test indicates abnormal returns statistically significant at the 5% level only on the fifth day after the event.

In the group of neutral events, the sign of $\overline{SAR'}$ changes again and again. The parametric test indicates statistically significant abnormal returns only before the recommendation: positive (at the 1% level) on the fifth day before the event and negative (at the 5% level) one day before it. A non-parametric test that, as previously mentioned, is more credible suggests only a positive reaction (at the 1% level) five days before the event, on the event day, and three days after it. The results from the non-parametric test could be interpreted as information that neutral recommendations are perceived as positive. However, negative returns that are not significant but still appear between those that are positive seem to deny such a conclusion. These inconsistent results are not surprising. As presented in Chapter 3 (for the period of 2012–2015), the reaction of investors to neutral recommendation can depend strongly on its previous level, and the analysis of level changes will probably lead to more-transparent results.

The aforementioned results are partly consistent with those obtained by Buzała (2012). He does not observe any market reaction to neutral events (HOLD) but finds a strong reaction to BUY and SELL recommendations (consistent with the information contained in the recommendation). For these two most-extreme levels, Buzała (2012) also observes a permanent accumulation of ACAR that starts 10 days before the event and lasts until about the 15th day after it. This drift appearing before the publication of recommendation is interpreted as a signal that recommendations were previously known to some group of investors who affect the market. Examining recommendations separately with the first submission to a limited group of investors, he finds a statistically significant ACAR even before this first publication. However, in the case of such recommendations, the reaction to their subsequent publication on the Internet is weaker. In our study, we do not observe such a significant and unequivocal market reaction anticipating the event date. Nevertheless, in the case of negative events, all $\overline{SAR'}$ are negative (although not statistically significant), which can be perceived as supporting Buzała's results.

Table 6

Reaction of daily returns to particular types of analyst recommendations (positive, neutral, and negative) issued between the 1st of January 2010 and the 31st of December 2011

t	**Positive recommendation (Buy/Accumulate)**			**Neutral recommendation (Hold)**			**Negative recommendation (Sell/Reduce)**		
	$\overline{SAR'}$	***p*-value**		$\overline{SAR'}$	***p*-value**		$\overline{SAR'}$	***p*-value**	
		T_a	T_{CZ}		T_a	T_{CZ}		T_a	T_{CZ}
	189 events			**190 events**			**63 events**		
–5	0.110	0.0639	**0.001**	0.178	**0.0063**	**0.001**	–0.045	0.3657	0.178
–4	–0.073	0.1781	0.469	0.024	0.3727	0.078	–0.100	0.2787	0.277
–3	–0.133	**0.0276**	0.261	–0.025	0.3588	0.472	–0.139	0.1482	0.258
–2	–0.019	0.3758	0.269	0.036	0.3400	0.070	–0.046	0.3721	0.437
–1	0.018	0.3812	0.118	–0.142	**0.0310**	0.247	–0.091	0.2526	0.340
0	0.206	**0.0078**	**0.000**	0.137	0.0680	**0.006**	–0.477	**0.0006**	**0.001**
1	0.102	0.1499	**0.006**	–0.069	0.2524	0.399	–0.111	0.2706	0.494
2	0.044	0.3328	0.158	–0.037	0.3484	0.266	–0.075	0.3333	0.403
3	–0.033	0.3586	0.353	0.096	0.1656	**0.009**	–0.007	0.3967	0.365
4	–0.006	0.3969	0.113	–0.130	0.0798	0.313	–0.124	0.2460	0.333
5	0.067	0.2615	0.081	0.038	0.3466	0.271	–0.256	0.0541	**0.044**

5.3. The analysis of changes in the recommendation level

The research conducted for the changes in the recommendation level differs significantly from the corresponding study of Buzała (2012). In this research, only three recommendation levels are defined (BUY and ACCUMULATE as well as SELL and REDUCE are gathered in joint groups defined as positive and negative events, respectively), and only changes between the three levels are studied. Buzała (2012) takes into account all five recommendation levels (for a change), but he does not consider each of the possible changes separately. As previously mentioned, he jointly researches all upgrades as well as all downgrades and then additionally considers upgrades to BUY and downgrades to SELL. Hence, for example, the change from ACCUMULATE to BUY is assigned by Buzała to upgrades, while in this research, such an event is clustered in the group of positive events that were previously also positive (so, to the group with no level change). Such substantial differences in definitions of recommendation levels and their changes significantly hinders a comparison of the results.

For all groups of level changes, Buzała (2012) observes progressive accumulation of statistically significant cumulative abnormal returns that starts about 5 days before the first issue of the recommendation and lasts even to the 30^{th} day after it. The sign of the return is, in each case, consistent with the information contained in the recommendation. The reaction is the strongest in the groups of upgrades to BUY and downgrades to SELL. Moreover, the values of CARs suggest that negative events (SELL and downgrades) influence prices more significantly than positive ones (BUY and upgrades).

Detailed results from conducted analysis of recommendation-level changes are not presented here since the article would be too long, but these can be provided upon request. In general, they confirm Buzała's results that stock prices react to changes in the recommendation level and that the reaction is consistent with the direction of the change. Nevertheless, they do not show clear evidence of a significant reaction before the recommendation issue and do not lead to the inference that the reaction is stronger in the case of downgrades. A direct comparison comes up against difficulties associated with differences in the applied methods of research.

An analysis of the whole period of 2010–2015 was also conducted by the author. The results do not vary considerably from those discussed in this paper and can be presented upon request.

References

[1] Barber, B.M., Lehavy, R., McNichols, M. and Trueman, B. (2001) 'Can investors profit from the prophets? Security analyst recommendations and stock returns', *The Journal of Finance*, vol. 56, pp. 531–563.

[2] Barber, B.M., Lehavy, R., McNichols, M. and Trueman, B. (2005) 'Buys, holds, and sells: The distribution of investment banks' stock ratings and the implications for the profitability of analysts' recommendations', *Journal of Accounting and Economics*, vol. 41, pp. 87–117.

[3] Brown, S.J., Warner, J.B. (1980) 'Measuring security price performance', *Journal of Financial Economics*, vol. 8, pp. 205–258.

[4] Busse, J.A. and Green, T.C. (2002) 'Market efficiency in real time', *Journal of Financial Economics*, vol. 65, pp. 415–437.

[5] Buzała, P. (2012) 'Reakcja inwestorów na rekomendacje giełdowe. Implikacje dla efektywności rynku akcji', *Zeszyty Naukowe Uniwersytetu Szczecińskiego nr 698, Finanse, Rynki Finansowe, Ubezpieczenia*, no. 50, pp. 11–20.
[6] Corrado, C.J. (2011) 'Event studies: A methodology review', *Accounting and Finance*, vol. 51, pp. 207–234.
[7] Corrado, C.J. and Zivney, T.L. (1992) 'The specification and power of the sign test in event study hypothesis tests using daily stock returns', *Journal of Financial and Quantitative Analysis*, vol. 27(3), pp. 465–478.
[8] Elton, E.J., Gruber, M.J. and Grossman, S. (1986) 'Discrete expectational data and portfolio performance', *The Journal of Finance*, vol. 41, no. 3, pp. 699–713.
[9] Glezakos, M. (2007) 'Revised Analysts' Recommendations: Are they More Valuable to Investors?', *European Journal of Economics, Finance and Administrative Sciences*, issue 7, pp. 60–70.
[10] Green, T.C. (2006) 'The value of client access to analyst recommendations', *Journal of Financial and Quantitative Analysis*, vol. 41, pp. 1–24.
[11] Gurgul, H. (2006) 'Zawartość informacyjna opinii analityków', in Analiza zdarzeń na rynkach akcji, Kraków: Oficyna Ekonomiczna.
[12] Gurgul, H. and Majdosz, P. (2004) 'The information content of analysts' recommendations: The empirical evidence from the Warsaw Stock Exchange', *Statistics in Transition*, vol. 6, pp. 1151–1163.
[13] Jegadeesh, M. and Kim, W. (2006) 'Value of analyst recommendations: International evidence', *Journal of Financial Markets*, vol. 9, pp. 274–309.
[14] Jegadeesh, M. and Kim, W. (2010) 'Do analysts herd? An analysis of recommendations and market reactions', *Review of Financial Studie*s, vol. 23, pp. 901–937.
[15] Jegadeesh, M., Kim, J., Krische, S.D., Lee and C.M.C. (2004) 'Analyzing the analysts: When do recommendations add value?', *The Journal of Finance*, vol. 59, pp. 1083–1124.
[16] Kim, S.T., Lin, J.-C. and Slovin, M.B. (1997) 'Market Structure, Informed Trading and Analyst's Recommendations', *The Journal of Financial and Quantitative Analysis*, vol. 32, pp. 507–524.
[17] Lin, H. and McNichols, M. (1998) 'Underwritting relationship, analysts, earnings forecasts and investment recommendations', *Journal of Accounting and Economics*, vol. 25, pp. 101–127.
[18] Loh, R.K. and Stulz, R.M. (2009) 'When are analyst recommendation changes influential?', The National Bureau of Economic Research Working Paper, [Online], Available: http://www.nber.org/papers/w14971, no. 14971.
[19] Mielcarz, P. Podgórski, B. and Weremczuk, P. (2007) 'Positive recommendations and abnormal return on the Warsaw Stock Exchange in 2005–2006', in Urbańczyk E. (ed.) The Problems of Company Value Management, Szczecin: Wydawnictwo Naukowe Uniwersytetu Szczecińskiego.

[20] Murg, M. and Zeitlberger, A. (2014) 'Impact of analyst recommendations on Austrian and German stocks: Information leaks, overreaction, and the influence of firm size', *Journal of Banking and Financial Research*, vol. 62(11), pp. 826–837.
[21] Murg, M., Pachler, M. and Zeitlberger, A. (2014) 'The impact of analyst recommendations on stock prices in Austria (2000–2014): evidence from a small and thinly traded market', *Central European Journal of Operations Research*, pp. 1–22, DOI 10.1007/s10100-014-0360-9.
[22] Podgórski, B. and Mielcarz, P. (2008) 'Wpływ negatywnych i neutralnych rekomendacji na osiąganie ponadprzeciętnych stóp zwrotu na GPW w latach 2005–2006', Studia i Prace Wydziału Nauk Ekonomicznych i Zarządzania Uniwersytetu Szczecińskiego, no. 7, pp. 533–543.
[23] Stickel, S.E. (1995) 'The anatomy of the performance of buy and sell recommendations', *Financial Analysts Journal*, vol. 55, issue 5, pp. 25–39.
[24] Womack, K.L. (1996) 'Do brokerage analysts' recommendations have investment value?', *The Journal of Finance*, vol. 51, no. 1, pp. 137–167.
[25] Yazici, B. Muradoglu, G. (2002) 'Dissemination of stock recommendations and small investors: who benefits?', *Multinational Finance Journal*, vol. 6, no. 1, pp. 29–42.

Managerial Economics
2016, vol. 17, no. 1, pp. 149–162
http://dx.doi.org/10.7494/manage.2016.17.1.149

Tomasz Wójtowicz*

Intraday patterns in time-varying correlations among Central European stock markets[1]

1. Introduction

Existence and strength of relationships between various markets is an important issue examined in the economic and econometric literature in recent years. An increasing number of papers investigate short- and long-term linkages between returns and volatility on different stock exchanges. Such studies have been also performed for European stock markets, however, some of their results still lack consensus.

On the basis of daily data, Voronkova (2004) shows the existence of long-term linkages between European developed markets and three CEE stock markets. Additionally, Syriopoulos (2004, 2007) indicates that relationships between CEE and developed markets are stronger than among CEE countries themselves. On the other hand, Černy and Koblas (2005) as well as Égert and Kočenda (2007) do not find long-term relationships between intraday data of emerging and developed European stock markets.

Investigation of short-term relations, particularly Granger causality, leads to more common results. Hanousek et al. (2009) prove significant spillover effects on three CEE emerging markets, namely, Prague, Budapest and Warsaw. Their main indices influence each other, but they are also significantly influenced by returns of DAX, the main index of the Frankfurt Stock Exchange (FSE). The impact of FSE is even stronger than the impact of any of the emerging markets. Similar results are evidenced by Černý and Koblas (2005).

An important role of developed European markets for CEE emerging markets is also indicated by Égert and Kočenda (2007). On the basis of intraday data, they

* AGH University of Science and Technology, Faculty of Management, Department of Applications of Mathematics in Economics, e-mail: twojtow@agh.edu.pl

[1] Financial support for this paper from the National Science Centre of Poland [Research Grant DEC--2012/05/B/HS4/00810] is gratefully acknowledged.

show significant causalities between returns of CEE markets and causal relations from developed to emerging markets. An analogous pattern is observed for volatility. Interestingly, there is also evidence of the opposite relations from the volatility of BUX and WIG20 to that of DAX and UKX.

Interdependencies and co-movement of European stock markets have been also analyzed via multivariate GARCH models. Using CCC and STCCC models, Savva and Aslanidis (2010) show that the largest CCE markets (in Czech Republic, Hungary and Poland) exhibit stronger correlations with the euro area than smaller markets (such as Slovenia and Slovakia).

Syllignakis and Kouretas (2011) show that correlations between developed and emerging European markets have increased over time. The largest shift was caused by the 2007–2009 financial crisis. Also Gjika and Horvath (2013) confirm strong correlations between CEE markets and markets in the euro area. They show that the accession of CEE countries to the EU increased correlations. On the other hand, Égert and Kočenda (2011) show something opposite. They find very little positive time-varying correlations among intraday returns of BUX, PX50 and WIG20. Correlations between these markets and Western European stock markets also are very weak.

In this paper, we focus on relationships on an intraday scale. We study time-varying co-movement of prices on three Central European stock exchanges in Frankfurt, Vienna and Warsaw. These stock markets differ considerably. The Frankfurt Stock Exchange (FSE) is an example of a large developed market. In fact, it is one of the largest and the most important stock markets in Europe. Taking into account capitalization, the Vienna Stock Exchange (VSE) is about eighteen times smaller and the Warsaw Stock Exchange (WSE) is about eleven times smaller than FSE[2]. Despite these differences, the stock exchanges in Frankfurt and Vienna are both developed markets, while WSE is still seen as an emerging market. Hence, in this paper, we analyze relationships between large (FSE) and smaller stock markets (VSE and WSE) and also between developed (FSE and VSE) and emerging (WSE) stock markets. We study how these similarities and differences are reflected in the correlations between the markets and how they impact relationships between them.

We focus our attention on intraday patterns in conditional correlations between these stock markets. We describe and compare intraday correlations on different days of the week. We also analyze the impact of US macroeconomic news announcements on the strength of interrelations between stock exchanges

[2] At the end of July 2015, capitalization of FSE was at the level of 1,625,718 mln € compared to 147,417 mln € of capitalization of WSE and 90,932 mln € capitalization of VSE [Source: Federation of European Securities Exchanges, www.fese.eu].

in Frankfurt, Vienna and Warsaw. Announcements of various US macroeconomic indicators were shown to be very important to European stock markets (Nikkinen and Sahlström, 2004; Harju and Hussain, 2011; Gurgul and Wójtowicz, 2014, 2015).

In order to analyze the evolution of time-varying intraday linkages between the markets, we apply the dynamic conditional correlation model (DCC) introduced by Engle (2002) to 5-minute data from the period between March 22, 2013 and July 31, 2014. It allows us to describe the evolution of short-time linkages between the stock markets under study as well as study intraday patterns in these relationships. During the estimation of the appropriate VAR model, we also examine the existence of Granger causalities between intraday returns. Results of this study will contribute to a better understanding of linkages between European stock markets, particularly in the CEE region.

The rest of the paper is organized as follows. In the next section we give short description of the DCC-GARCH models. In Section 3 we present and analyze in detail the data that we use in the empirical study. Section 4 contains the main empirical findings. A short summary concludes the paper.

2. DCC-GARCH Models

The dynamic conditional correlation (DCC) model introduced by Engle (2002) is one of the multivariate volatility models. It is a generalization of the constant conditional correlation (CCC) model of Bollerslev (1990). The DCC model allows a quite simple description of a time-varying variance-covariance matrix between return series. The model assumes that n-dimensional vector of returns $r_t = (r_{1,t}, \ldots, r_{n,t})'$ has conditional multivariate normal distribution with zero mean and covariance matrix H_t, i.e. $r_t | \Omega_{t-1} \sim N(0, H_t)$, where Ω_{t-1} is the information set available at time $t - 1$[3]. In the DCC model, the covariance matrix can be decomposed into:

$$H_t = D_t R_t D_t \tag{1}$$

where $D_t = diag(h_{1,t}, \ldots, h_{nn,t})$ is a diagonal matrix of conditional standard deviations from univariate GARCH models and R_t is the time-varying conditional correlation matrix of the following form:

$$R_t = diag\left(Q_t\right)^{-\frac{1}{2}} Q_t diag\left(Q_t\right)^{-\frac{1}{2}} \tag{2}$$

Dynamics of conditional correlations depends on the definition of Q_t.

[3] In practice, returns are replaced by residuals from the appropriate VAR model.

In the DCC model with one lag, Q_t evolves in time according to the formula (3):

$$Q_t = (1-a-b)\bar{Q} + a\varepsilon_{t-1}\varepsilon'_{t-1} + bQ_{t-1} \tag{3}$$

where $\varepsilon_t = D_t^{-1}r_t$ are standardized returns, $\bar{Q} = E(\varepsilon_t\varepsilon'_t)$ is the unconditional covariance of standardized returns, and a and b are nonnegative parameters such that $a + b < 1$. If this condition is satisfied, Q_t reverses back to $\bar{Q}$.

The parameters of the DCC-GARCH model are estimated via a two-step procedure. In the first step parameters of univariate GARCH models are estimated and returns are standardized. Then, a and b are estimated by maximizing the following likelihood function (Engle, 2002):

$$L_C(\theta) = -\frac{1}{2}\sum_t \left(\ln|R_t| + \varepsilon'_t R_t^{-1}\varepsilon_t\right) \tag{4}$$

where $\theta = (a, b)$. Computation of R_t in (4) (via computation of Q_t) is made recursively with starting value Q_0. When the DCC model is estimated on the basis of intraday data pooled together (like in this paper), the first value of Q_t for each day is computed on the basis of the last value from the previous day.

In this paper, to take into account possible different dynamics of conditional correlations during days with and without US macroeconomic news announcements, we also consider a regime-switching DCC model with a covariance matrix of the following form:

$$Q_t = \begin{cases} Q_t^{\mathrm{I}} = (1-a_{\mathrm{I}}-b_{\mathrm{I}})\bar{Q}_{\mathrm{I}} + a_{\mathrm{I}}\varepsilon_{t-1}\varepsilon'_{t-1} + b_{\mathrm{I}}Q_{t-1} \\ Q_t^{\mathrm{II}} = (1-a_{\mathrm{II}}-b_{\mathrm{II}})\bar{Q}_{\mathrm{II}} + a_{\mathrm{II}}\varepsilon_{t-1}\varepsilon'_{t-1} + b_{\mathrm{II}}Q_{t-1} \end{cases} \tag{5}$$

where regimes I and II correspond to days without and with US macroeconomic news announcements, respectively, and $\bar{Q}_{\mathrm{I}}$ and $\bar{Q}_{\mathrm{II}}$ are unconditional covariances of standardized returns in each set of days. Additionally to the above dynamic structure, we assume that the initial value for each day is equal to $\bar{Q}_{\mathrm{I}}$ and $\bar{Q}_{\mathrm{II}}$, depending on the regime. With these assumptions, the sum in the likelihood function (4) can be separated into two sums: $L_C^{\mathrm{I}}(\theta)$ and $L_C^{\mathrm{II}}(\theta)$ for each regime, respectively. It follows that the estimation of such DCC model for days with news announcements does not depend on returns for days without news, and vice versa.

3. Data

The analysis presented in this paper is based on 5-minute returns of DAX, ATX, and WIG20, the main indices of stock exchanges in Frankfurt, Vienna, and Warsaw. The returns cover the period from March 22, 2013 to July 31, 2014. Data come from Bloomberg, the Vienna Stock Exchange and the Warsaw Stock Exchange.

In the analysis, we consider only those returns from days when all of the markets were open. However, trading hours on the stock markets must be also taken into account, because the stock markets were open at different hours in the period under study. In 2013 and 2014, continuous trading started at 8:55 on VSE and at 9:00 on FSE and WSE. It ended at 16:50 (WSE), 17:30 (FSE), and 17:35 (VSE). Moreover, on FSE and VSE, there were intraday auctions at 13:00 and 12:00, respectively. Due to these differences in trading hours on the markets, and to the fact that the first 5-minute intraday return is observed at 9:05 and is accompanied by very high volatility, intraday relations are analyzed only in the common periods between 9:10 and 16:50. To model (or filter) intraday data, we must take into account the well-known fact that intraday volatility increases at the beginning and end of each trading session. Figure 1 shows a U-shaped pattern observed in intraday return volatility. It also shows a strong impact of news about the US economy, which is usually announced at 14:30. This strong impact of various US macroeconomic news announcements on the European stock market is widely confirmed by empirical works (e.g., Harju and Hussain, 2011; Gurgul and Wójtowicz, 2015).

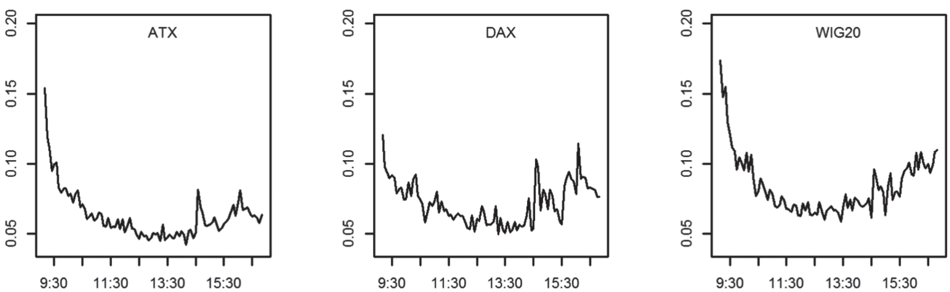

Figure 1. Cross-sectional standard deviations (in percentages) of 5-minute returns of ATX, DAX, and WIG20

To deal with periodic patterns in volatility as well as the impact of US news announcements, we apply a method of Flexible Fourier Form (FFF) adopted to intraday data by Andersen and Bollerslev (1997). Specifically, we decompose 5-min returns $R_{t,n}$ at time n on day t as:

$$R_{t,n} - E\left(R_{t,n}\right) = s_{t,n}\sigma_{t,n}Z_{t,n}$$

where $Z_{t,n}$ is $i.i.d(0,1)$, $\sigma_{t,n}$ is a daily volatility factor and $s_{t,n}$ is an intraday (diurnal) seasonal component such that $\ln(s^2_{t,n})$ can be estimated from the following FFF regression (6).

$$2ln\frac{\left|R_{t,n}-\bar{R}\right|}{\left(\hat{\sigma}_t N^{0.5}\right)}=c+\sum_{k=1}^{D}\lambda_k I_k\left(t,n\right)+\delta_1\frac{n}{N_1}+\delta_2\frac{n^2}{N_2}+ \\ +\sum_{p=1}^{P}\left(\delta_{c,p}\cos\left(\frac{2\pi p}{N}n\right)+\delta_{s,p}\sin\left(\frac{2\pi p}{N}n\right)\right)+\varepsilon_{t,n} \tag{6}$$

where refers to the number of returns per day (here N = 94), $N_1=\frac{N+1}{2}$, $N_2=\frac{\left(N+1\right)\left(N+2\right)}{6}$, $I_k(t,n)$ allows for the inclusion of weekdays and US macroeconomic news announcement dummies. In this paper, we use six dummy variables to model intraday volatility up to a half an hour after news announcements. The daily variance component $\sigma_{t,n}$ is approximated by volatility forecasts from the appropriate GARCH model with skewed Student's *t*-distribution constructed for daily returns.

On the basis of literature (e.g., Nikkinen et al., 2006; Harju and Hussain, 2011; Gurgul and Wójtowicz, 2014, 2015), we include regression dummy variables in the FFF describing the impact of announcements of the following US macroeconomic indicators: Consumer Price Index (CPI), Producer Price Index (PPI), Industrial Production (IP), Retail Sales (RS), Durable Goods Orders (DGO), Nonfarm Payrolls (NFP), Existing Home Sales (EHS), Housing Starts (HS), and New Home Sales (NHS). The majority of them (CPI, PPI, RS, DGO, NFP, and HS) are released at 8:30 EST[4] (14:30 CET). EHS and NHS are released at 10:00 EST (16:00 CET). Only IP is released at 9:15 EST (15:15 CET). Due to the differences in the introduction of Daylight Saving Time in the US and Europe, some of announcements reach European stock markets one hour earlier in March and October; i.e., at 13:30 CET, 15:00 CET, and 14:15 CET, respectively.

4. Empirical Results

We start the analysis with a computation of unconditional Spearman correlations between the returns of ATX, DAX, and WIG20 in the entire period of March 2013 – July 2014. This will be a background for further analysis of time-varying intraday co-movements. Results in Table 1 give very general information about the average strength of relations between the indices. All computed values of correlation coefficients are significantly positive and indicate rather mild interdependencies between the markets, particularly between FSE and VSE. The smallest, but still significant, correlation is between ATX and WIG20. This is in contrast

[4] EST – Eastern Standard Time; CET – Central European Time

with the results of Égert and Kočenda (2011) of very weak intraday correlations between CEE markets and European developed markets.

Table 1

Spearman correlations between 5-min returns of ATX, DAX, and WIG20

	ATX-DAX	ATX-WIG20	DAX-WIG20
Correlation	0.336	0.154	0.254

More in-depth analysis of intraday relations is made on the basis of the DCC-GARCH model described in Section 2. First, we estimate the trivariate VAR model to filter out autocorrelation observed in intraday returns. On the basis of the Akaike information criterion, we chose the VAR model with 7 lags. Its estimation provides a perfect opportunity to study Granger causalities between the stock markets. The significance of past DAX returns in equations for WIG20 and ATX returns[5] indicates a strong one-directional intraday Granger causality from the stock exchange in Frankfurt to markets in Vienna and Warsaw. This is in line with previous results indicating the strong impact of large developed European markets on the stock exchanges in the CEE region. Moreover, it indicates that such an impact is observed not only for emerging markets but also for mature markets (like VSE).

To model the conditional variance of the univariate series of residuals from the VAR model, we first remove diurnal periodicity from the 5-min return volatility. The application of FFF confirms the conclusions from Figure 1 regarding a very high variance of returns at the beginning of the trading session. It also indicates a strong and significant impact of US macroeconomic news announcements on intraday volatility. For each index, dummy variables are significant in the first 5-minute period after news announcements (irrespective of the time of the announcement). This is clearly visible in Figure 2, where we present examples of intraday volatility components for days with US announcements at 14:30. After removing the daily and intraday seasonality components of volatility we filter out 5-min returns with GARCH(1,1) models with conditional skewed Student's *t*-distribution.

Time-varying correlations of the standardized residuals are modeled via the DCC model with 1 lag and multivariate normal distribution. The estimation results reported in Table 2 are typical for a financial time series: a small value of and a value of significant and close to 1 indicate very strong persistence of time-varying intraday correlations between ATX, DAX, and WIG20.

[5] The F statistics in the significance tests of joint impact of historical DAX returns in equations for ATX and WIG20 returns is significant at any reasonable level.

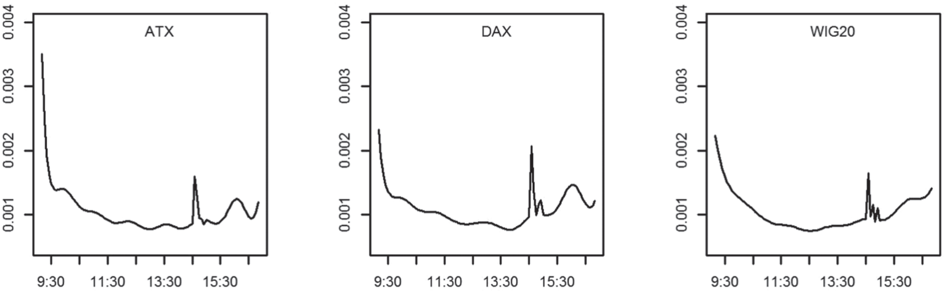

Figure 2. Intraday seasonal component $s_{t,n}$ of ATX, DAX, and WIG20 on days with US macroeconomic news announcement at 14:30

Table 2

Parameters of DCC-GARCH model for 5-minute returns of ATX, DAX, and WIG20

	Estimate	**Std. error**	***t*-statistics**	***p*-value**
a	0.0068	0.0071	0.956	0.339
b	0.9894	0.0147	67.46	0

The strength and evolution of interrelations between the stock markets is captured by intraday conditional correlations (the off-diagonal elements of matrices R_t) presented in Figure 3. They are in line with values of Spearman correlations in Table 1. In general, intraday correlations vary around their unconditional values in the period under study. The strongest relation is observed between ATX and DAX, where conditional correlations change between 0.1 and 0.6. The weakest relations are observed between ATX and WIG20, where intraday correlations are smaller than 0.4 during the whole period. Changes in intraday relations between the markets are similar. In the first part of the period, intraday correlations decrease, while the lowest correlations are generally observed in the central part of the sample.

Additionally, to the analysis of the whole period, we study changes in correlations during the trading day. For each time t from set 9:10, 9:15, ..., 16:50, we compute cross-sectional average $\bar{R}_{t,ij}$ of conditional correlations between indices i and j at time t. Changes in the averages presented in Figure 4 indicate the existence of an intraday pattern in the relationships between stock markets. In general, correlations are stronger at the beginning and at the end of a trading session, while they are weaker in the middle of the day. It is important to note that conditional correlations start to increase about 14:30; i.e., when the majority of important US data is announced and the US derivative market opens. The averages reach the highest values in the final part of the trading session (around 16:30). These observations lead to the question about the impact of US data announcements on the strength of interrelations between European stock markets.

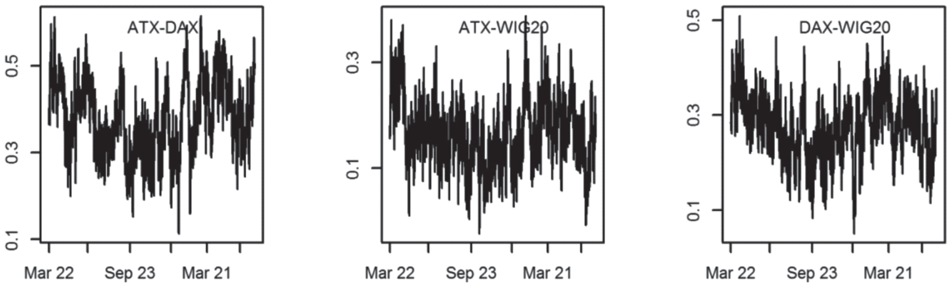

Figure 3. Intraday conditional correlations between ATX, DAX, and WIG20 during the period of March 22, 2013 – July 31, 2014

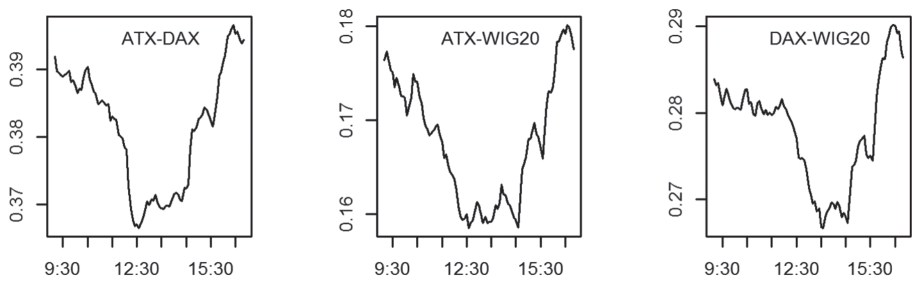

Figure 4. Cross-sectional averages of conditional correlations between ATX, DAX, and WIG20

4.1. Correlations during days with US macroeconomic news announcements

As a first insight into relationships during days with and without US macroeconomic news, we compare Spearman correlations during these days. Values in Table 3 indicate that stock markets are more closely related when important news about the US economy is to be announced. During these days, correlation coefficients are about 11–18% higher than during days without scheduled information from the US.

Table 3

Spearman correlations between 5-min returns of ATX, DAX, and WIG20 during days with and without US macroeconomic news announcements

	ATX-DAX	ATX-WIG20	DAX-WIG20
Days without announcements	0.319	0.147	0.246
Days with announcements	0.376	0.170	0.273

To compare correlations between the markets during days when US data is announced as well as days without such important announcements, we estimate a regime-switching DCC model. Estimation results are reported in Table 4. The values of a_{I} and b_{I} are close to the values of a and b from Table 2. In fact, the differences between the respective parameters are insignificant. But, when we compare the right panel of Table 4 with Table 2, we can notice that b_{II} is significantly smaller than b for the whole sample. Moreover, a_{II} and b_{II} are significantly different from a_{I} and b_{I}, respectively. This confirms that the dynamics of intraday conditional correlations during days with and without US news announcements differ significantly. When US news is announced, the conditional correlation is less persistent, and the impact of previous returns is a little stronger than during days without new information.

Table 4

Parameters of regime switching DCC-GARCH models for ATX, DAX, and WIG20 during days with and without US macroeconomic news announcements.

Parameter	Days without US news			Parameter	Days with US news		
	Estimate	Std. error	*p*-value		Estimate	Std. error	*p*-value
a_{I}	0.008	0.0007	0	a_{II}	0.014	0.0013	0
b_{I}	0.979	0.0035	0	b_{II}	0.958	0.0055	0

As before, we compare not only the estimated values but also the intraday seasonality of conditional correlations during both types of days. In Figure 5, we can observe that the cross-sectional averages of conditional correlations during days with announcements are above the averages from days without announcements during the whole trading day. To be more precise, for each time t, we use the Kruskal-Wallis test to compare the distributions of the both groups of conditional correlations. The shadowed regions in Figure 5 indicate t for which the null hypothesis about equality of distributions cannot be rejected. The clearest interpretation is for ATX and DAX where, for each time t, the conditional correlations during days with announcements are significantly greater than correlations during days without US macroeconomic news. For pairings ATX-WIG20 and DAX-WIG20, such significant differences are visible at the beginning of the trading session (until 10:50 for ATX-WIG20 or until 12:20 for DAX-WIG20) and at the end of the trading session (after 14:40 for ATX-WIG20 and after 14:00 for DAX-WIG20). A comparison of the left and right panels in Figure 5 shows the difference in relations of stock exchanges in Warsaw and Vienna with the stock market in Frankfurt. Scheduled US macroeconomic news announcements significantly strengthen the relationships between VSE and FSE over the whole day. In the case of WSE, an increase in correlation is visible in the presence of new information.

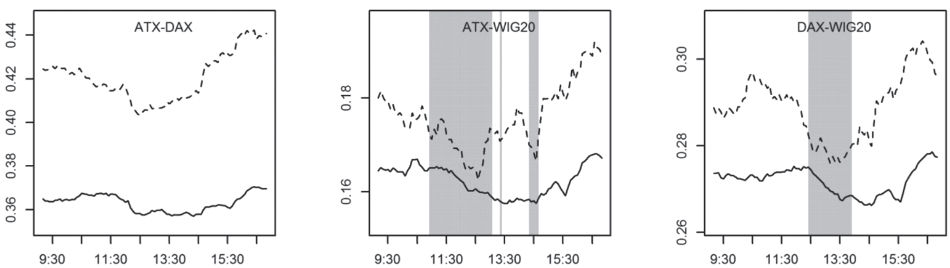

Figure 5. Cross-sectional averages of conditional correlations between ATX, DAX, and WIG20 during days with US macroeconomic news announcements (dashed lines) and without (solid lines)

4.2. Correlations in different days of the week

In order to compare the heterogeneity of interrelations between stock markets in Frankfurt, Vienna, and Warsaw throughout the week, we compare correlations between these markets on different days of the week. From Table 5 (where we report unconditional Spearman correlations), we can notice that the lowest correlations are at the beginning of the week (on Mondays for correlations with DAX and on Tuesdays for correlations between ATX and WIG20). The highest correlations are observed on Wednesdays (DAX-WIG20) or Thursdays (ATX-DAX and ATX-WIG20). These differences confirm that the strength of intraday interrelations between the stock markets under study may depend on the day of the week. They also suggest the existence of the 'day-of-the-week effect' in conditional correlations. To examine this, we estimate a regime-switching DCC model with five regimes corresponding to the days of the week. In the last two columns of Table 5, we report estimated values of model parameters $a_{\text{Mon}},\ldots, a_{\text{Fri}}$ and $b_{\text{Mon}},\ldots, b_{\text{Fri}}$. All of them are significantly greater than 0 at the 1% level.

Table 5

Spearman correlations between 5-min returns of ATX, DAX, and WIG20 on different days of the week during days with and without US macroeconomic news announcements

	ATX-DAX	ATX-WIG20	DAX-WIG20	*a*	*b*
Mondays	0.322	0.144	0.222	0.0081	0.9845
Tuesdays	0.327	0.137	0.245	0.0070	0.9649
Wednesdays	0.333	0.165	0.272	0.0102	0.9708
Thursdays	0.352	0.167	0.258	0.0121	0.9705
Fridays	0.346	0.156	0.271	0.0118	0.9662

Similar to section 4.1, we also compare cross-sectional distributions of intraday conditional correlations for each day of the week. As before, at each time *t*, we apply the Kruskal-Wallis test to verify the significance of the difference between the distributions of conditional correlations from different days of the week.

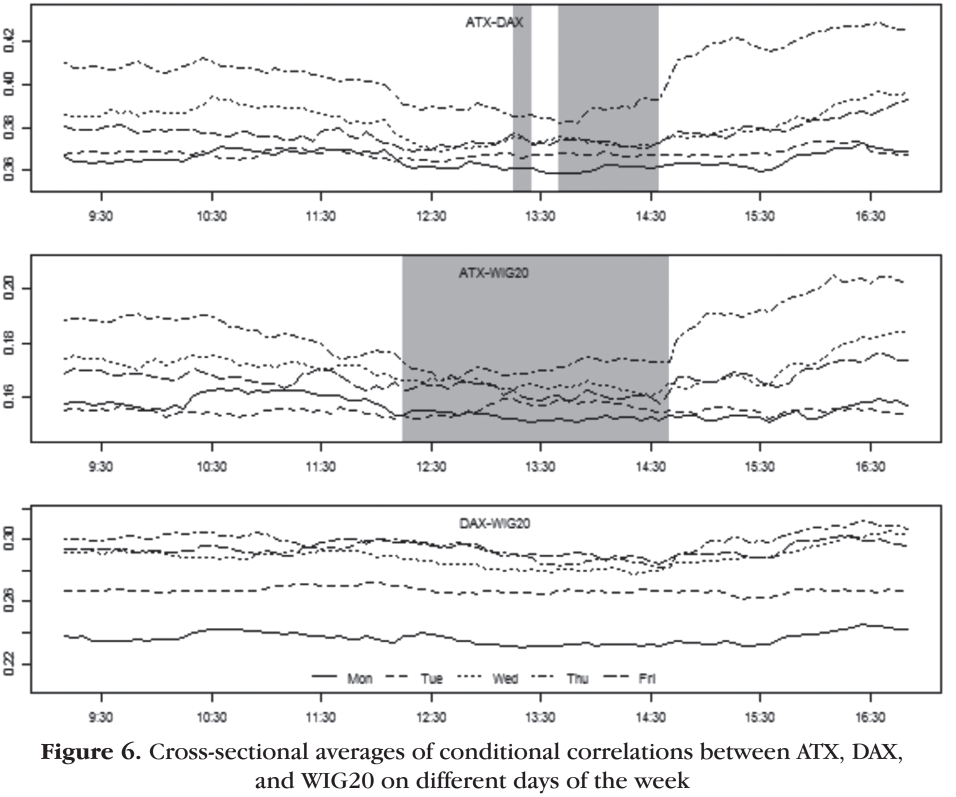

Figure 6. Cross-sectional averages of conditional correlations between ATX, DAX, and WIG20 on different days of the week

The shaded regions in Figure 6 indicate values of t where the null hypothesis is not rejected. From the bottom panel of Figure 6, we can conclude that the relationships between DAX and WIG20 strongly depends on the day of the week. It is significantly the weakest in Mondays than on other days. These differences are significant for each time t during the trading session. This is similar to the impact of US macroeconomic news announcements on the correlation between ATX and DAX. In the case of correlations with ATX, the rejection of the null hypothesis in the morning and afternoon is caused by significantly higher correlations on Thursdays rather than by low correlations on Mondays. This indicates a difference in the relationships between the stock exchange in Frankfurt and VSE and WSE. The strength of relationships between WSE and FSE is very weak in Mondays, while relationships on Wednesdays, Thursdays, and Fridays are similar.

5. Conclusions

In this paper, we analyze and compare interrelations between stock markets in Frankfurt, Vienna, and Warsaw. The analysis is performed on the basis of 5-minute data from the period of March 22, 2013 – July 31, 2014. The application of an appropriate VAR model confirms the previous results about Granger causality running from a large, developed stock exchange in Frankfurt to stock markets in Central and Eastern Europe. Further analysis indicates significant intraday correlations between the stock markets under study. The strongest relationships is observed between both developed markets in Frankfurt and Vienna. The application of DCC models shows the difference between intraday relations of ATX and WIG20 with DAX. US macroeconomic news announcements have a stronger impact on relationships between stock markets in Vienna and Frankfurt, while the day-of-the-week effect is more pronounced in the relationships between stock exchanges in Warsaw and Frankfurt.

References

[1] Andersen, T. and Bollerslev, T. (1997) ‘Intraday periodicity and volatility persistence in financial markets’, *Journal of Empirical Finance*, vol. 4, pp. 115–158.

[2] Bollerslev, T. (1990) ‘Modelling the Coherence in Short-Run Nominal Exchange Rates: A Multivariate Generalized ARCH Model’, *The Review of Economics and Statistics*, vol. 72, No. 3, pp. 498–505.

[3] Černý, A. and Koblas, M. (2005) ‘Stock Market Integration and the Speed of Information Transmission: The Role of Data Frequency in Cointegration and Granger Causality Tests’, *Journal of International Business and Economics*, vol. 1, pp. 110–120.

[4] Engle, R.F. (2002) 'Dynamic Conditional Correlation', *Journal of Business and Economic Statistics*, vol. 20, No. 3, pp. 339–350.
[5] Égert, B. and Kočenda, E. (2007) 'Interdependence between Eastern and Western European Stock Markets: Evidence from Intraday Data', *Economic Systems*, vol. 31, No. 2, pp. 184–203.
[6] Égert, B. and Kočenda, E. (2011) 'Time-varying synchronization of European stock markets', *Empirical Economics*, vol. 40, No. 2, pp. 393–407.
[7] Gjika, D. and Horváth, R. (2013) 'Stock Market Comovements in Central Europe: Evidence from Asymmetric DCC Model', *Economic Modelling*, vol. 33, pp. 55–64.
[8] Gurgul, H. and Wójtowicz, T. (2014) 'The impact of US macroeconomic news on the Polish stock market. The importance of company size to information flow', *Central European Journal of Operations Research*, vol. 22, pp. 795–817.
[9] Gurgul, H. and Wójtowicz, T. (2015) 'The Response of Intraday ATX Returns to U.S. Macroeconomic News', *Finance a úvěr – Czech Journal of Economics and Finance*, vol. 65, No. 3, pp. 230–253.
[10] Hanousek, J., Kočenda, E. and Kutan, A.M. (2009) 'The reaction of asset prices to macroeconomic announcements in new EU markets: evidence from intraday data', *Journal of Financial Stability*, vol. 5, No. 2, pp. 199–219.
[11] Harju, K. and Hussain, S.M. (2011) 'Intraday seasonalities and macroeconomic news announcements', *European Financial Management*, vol. 17, pp. 367–390.
[12] Nikkinen, J. and Sahlström, P. (2004) 'Scheduled Domestic and US Macroeconomic News and Stock Valuation in Europe', *Journal of Multinational Financial Management*, vol. 14, pp. 201–245.
[13] Nikkinen, J., Omran, M., Sahlström, M. and Äijö, A. (2006) 'Global stock market reactions to scheduled U.S. macroeconomic news announcements', *Global Finance Journal*, vol. 17(1), pp. 92–104.
[14] Savva, C.S. and Aslanidis, C. (2010) 'Stock Market Integration between New EU Member States and the Eurozone', *Empirical Economics*, vol. 39, No. 2, pp. 337–351.
[15] Syllignakis, M.N. and Kouretas, G.P. (2011) 'Dynamic Correlation Analysis of Financial Contagion: Evidence from the Central and Eastern European Markets', *International Review of Economics & Finance*, vol. 20, No. 4, pp. 717–732.
[16] Syriopoulos, T. (2004) 'International portfolio diversification to Central European stock markets', *Applied Financial Economics*, vol. 14, pp. 1253−1268.
[17] Syriopoulos, T. (2007) 'Dynamic linkages between emerging European and developed stock markets: Has the EMU any impact?', *International Review of Financial Analysis*, vol. 16, No. 1, pp. 41−60.
[18] Voronkova, S. (2004) 'Equity Market Integration in Central European Emerging Markets: A Cointegration Analysis with Shifting Regimes', *International Review of Financial Analysis*, vol. 13, No. 5, pp. 633–647.

Summaries

Anna Czapkiewicz, Marta Stachowicz: **The long-run relationship between the stock market and main macroeconomic variables in Poland** ▪ Managerial Economics 2016, vol. 17, no. 1

JEL Classification: G01, G11, G15

Keywords: *long-run relationship, model VECM, stock market, macroeconomic data*

The investigation concerns the problem of whether some macroeconomic variables and the EUR/PLN exchange rate might affect the performance of the Warsaw Stock Exchange. The answer to this question can be obtained from a cointegration analysis. The advantage of testing for cointegration is the identification of a stable long-run relationship between the stock price index, some macroeconomic variables, and the EUR/PLN exchange rate, which can be implemented using various cointegration methodologies. Analysis of the response of one variable to an impulse of another variable is also performed to show the importance of a given variable in a system.

Henryk Gurgul, Łukasz Lach: **Comparative advantage of the EU in global value chains: How important and efficient are new EU members in transition?** ▪ Managerial Economics 2016, vol. 17, no. 1

JEL Classification: F1, C67, D57

Keywords: *value added, productivity, capital efficiency, CEE economies, international input--output matrices, transition*

We suggest original modifications and extensions of the recently presented methodological developments in ex-post accounting framework in global value chains in order to obtain empirical results both for the analyzed group of ten CEE economies as well as at a country-and-sector-specific level. The empirical results confirm that the role of the selected CEE economies in transition in creating value added with respect to the total value added in the European Union in the GVC framework was biggest in the cases of agriculture-, wood-products-, metal-production, and travel-and-tourism-related sectors. We also found that, after two decades of transition, the measures of productivity in the examined economies in 2009 were still much lower as compared to the EU average for most of the sectors. Moreover, in the transition period, these indexes were increasing, especially after EU accession. In contrary, after two decades of transition, the measures of capital efficiency in the ten CEE economies in 2009 were comparable to the EU average for most of the sectors. Moreover, during this period, the growth rates of these indexes were, in general, positive. However, their growth rates dropped after EU accession.

Henryk Gurgul, Artur Machno: **The impact of asynchronous trading on Epps effect. Comparative study on Warsaw Stock Exchange and Vienna Stock Exchange** ▪ Managerial Economics 2016, vol. 17, no. 1

JEL Classification: F36, G15

Keywords: *VSE, WSE, market microstructure, Epps effect, asynchronous trading, correlation estimation, asynchronous time series*

The novelty/value added of this paper is the comparison of the Epps effect between developed and emerging stock markets from Central Europe by means of the correction formula derived by the authors. The main goal of the study is to test whether or not asynchrony in transaction times is a considerable source of the Epps effect in the case of the Warsaw and Vienna stock exchanges for the most-liquid assets from these markets. Among all analyzed stock pairs on the WSE, asynchrony turns out to be the main cause of the Epps effect. However, the corrected correlation estimator seems to be more volatile than the regular estimator of the correlation. In the case of the VSE, evidence of the Epps effect is not unique. For the most-liquid and most-correlated pair (namely, ANDR-EBS), the analysis delivers similar results as for Polish stocks. However, the Epps effect could not be detected for the remaining pairs on the VSE. The presented analysis can be reproduced for the same data or replicated for another dataset; all R codes used in the computation within this paper are available upon request.

Henryk Gurgul, Robert Syrek: **The logarithmic ACD model: the microstructure of the German and Polish stock markets** ▪ Managerial Economics 2016, vol. 17, no. 1

JEL Classification: G15, G19

Keywords: *intraday data, microstructure, duration, ACM models, Frankfurt Stock Exchange, Warsaw Stock Exchange*

The main goal of this paper is to compare the microstructure of selected stocks listed on the Frankfurt and Warsaw Stock Exchanges. We focus on the properties of duration on both markets and on fitting the appropriate ACD models. Because of the quite different levels of capitalization of stocks on these markets, we observe essential discrepancies between these stocks. While for most German companies on the DAX30, the Burr distribution fits better than generalized gamma distribution, the latter distribution is superior in the case of the largest Polish companies. Analyzing series by hazard function, we note the similarity of hazard functions for companies on both markets, which tend to have a U-shaped pattern.

Magdalena Kludacz-Alessandri: **Non-financial dimensions of measurement and assessment in the performance model for hospitals** ▪ Managerial Economics 2016, vol. 17, no. 1

JEL Classification: I18, M41

Keywords: *non-financial dimensions of measurement and assessment in the performance model for hospitals*

This paper presents one of the important stages of a research project carried out in Poland, the aim of which was to develop an integrated performance measurement and assessment system

designed for Polish public hospitals and for the stakeholders of different healthcare systems. The goal for the stage described in this article was to present a flexible and comprehensive tool (a performance model) for measuring hospital performance that could be used by hospital management, the founding bodies, and other stakeholders at the regional and central levels of healthcare systems, to improve the quality of medical services, the availability of healthcare resources, and the organization of the healthcare system in Poland. This article describes the non-financial dimensions that were identified in this model.
The research involves the construction of a model that consists of defined and selected dimensions (patients, internal processes, development, and finance) with key performance indicators for the analysis, measurement, and assessment of hospital performance. The model takes into account the following three levels of the Polish healthcare management system: the central level, represented by the Ministry of Health; the regional level, represented by the regional governor, the marshal's office, and the regional offices of the National Health Fund; and the local level, represented by hospitals and their funding bodies.
The performance model exceeds the scope of the financial dimension and enriches it with three non-financial dimensions: patients, internal processes, and development. It allows one to concentrate not only on the tasks of the hospital but also on the objectives of other stakeholders operating in the Polish healthcare system.

Milena Suliga: **The reaction of investors to analyst recommendations of stocks listed on the WIG20 index** ▪ Managerial Economics 2016, vol. 17, no. 1

JEL Classification: G14

Keywords: *abnormal returns, event-study methodology, recommendation changes, linear regression with categorical variables*

Analyst recommendations are one of the types of information whose appearance on the market can have an influence on security prices. In this paper, I study the impact of analyst recommendations on stocks listed on the WIG20 Index, using event-study methodology and linear regression models. The dataset contains 576 absolute recommendations published from the 1st of January 2012 to the 1st of September 2015 by various analyst houses. The prefatory study researches price reaction to positive, neutral, and negative recommendations separately. Subsequently, to check if investor reaction depends on a change in the level of recommendation, corresponding research is repeated for events clustered in nine groups defined in terms of possible level changes. Linear regression models with categorical variables are used in search of additional factors affecting investor reactions. Changes in the level of recommendation, size of the company, and reputation of brokerage house represent explanatory variables. Preliminary results point out that the direction of investor reaction is generally consistent with the information contained in the recommendation, and that the reaction of the market seems to be stronger in the case of positive events than in the case of negative ones. The analysis of recommendation changes reflects more-detailed dependents. In particular, the interpretation of a neutral recommendation depends strongly on the level of the previous recommendation. If it represents growth from SELL or REDUCE, the reaction is positive, while in the case of a drop from ACCUMULATE or BUY, it leads to negative abnormal returns. This relationship is additionally confirmed by results from the linear regression models. The models show the size of the firm as a significant factor that has an influence on the reaction to a recommendation: the smaller the firm, the stronger the reaction.

Tomasz Wójtowicz: **Intraday patterns in time-varying correlations among Central European stock markets** ▪ Managerial Economics 2016, vol. 17, no. 1

JEL Classification: G15

Keywords: *CEE stock markets, DCC-GARCH model, emerging markets, intraday data*

In this paper we investigate intraday relationships between three Central European stock exchanges: those in Frankfurt, Vienna and Warsaw. They represent different types of stock markets: two of them are developed, while the last is an emerging market. Via DCC-GARCH models we analyze and compare time-varying conditional correlations of intraday returns of the main indices of the stock exchanges. We study the impact of important public information, US macroeconomic news announcements, on the strength of interrelationships between the markets. Additionally, we analyze diurnal patterns in time-varying correlations on different days of the week.

Instruction for authors

Articles which are to be published in *Managerial Economics* should be prepared according to the following guidelines.

Length of the paper should not exceed 22.000 characters (figures and tables included).

The running text of the body of the paper should be followed by figures and tables. In the main text the authors should indicate places where each table/figure should appear.

Title page should include a footnote, giving the author's affiliation.

Figures must be prepared in a form suitable for direct reproduction. Digital artwork at least 300 dpi resolution is accepted. Photographs, on glossy paper (9 by 13 cm or larger), should display sharp contrast. Figures, tables and photographs should be numbered according to their reference in text.

Illustrations should be edited in CorelDraw (*.CDR), DrawPerfect (*.WPG) or in any other vector graphics form e.g. HPGL, Encapsulated PostScript (*.EPS), Computer Graphics Metafile *.CGM) or bitmaps (*.TIF, *.PCX).

Mathematical equations within the text should be written in separate lines, numbered consecutively (numbers within round brackets) on the right-hand side. Greek characters must be written out clearly.

Summary (no more than 300 words, in English) and keywords should be submitted in separate file containing the name of the author, title of the paper with the heading "Summary".

Authors using Word are requested to employ, as far as possible, text form of mathematical symbols leaving graphic form for the equations longer than single line.

Reference style

In general, the authors should use the Harvard style of referencing. References to literature within the text should be given in the form: the name of the author(s) and the year of publication (in parentheses), e.g. "Smith (1990) underlines.", "As shown in Smith (1990).". In case of more than two authors of the cited publication the "et al." shortcut should be used.

Lists of references should be written in alphabetical-chronological order, numbered and follow the rules:

- JOURNAL ARTICLE

Muller, V. (1994) 'Trapped in the body: Transsexualism, the law, sexual identity',

The Australian Feminist Law Journal, vol. 3, August, pp. 103–107.

- BOOKS

Book with one author

Adair, J. (1988) Effective time management: How to save time and spend it wisely, London: Pan Books.

Book with two authors

McCarthy, P. and Hatcher, C. (1996) Speaking persuasively: Making the most of your presentations, Sydney: Allen and Unwin.

Book with three or more authors

Fisher, R., Ury, W. and Patton, B. (1991) Getting to yes: Negotiating an agreement without giving in, 2nd edition, London: Century Business.

Book – second or later edition

Barnes, R. (1995) Successful study for degrees, 2nd edition, London: Routledge.

Book by same author in the same year

Napier, A. (1993a) Fatal storm, Sydney: Allen and Unwin. Napier, A. (1993b) Survival at sea, Sydney: Allen and Unwin.

Book with an editor

Danaher, P. (ed.) (1998) Beyond the ferris wheel, Rockhampton: CQU Press.

A chapter in a book

Byrne, J. (1995) 'Disabilities in tertiary education', in Rowan, L. and McNamee, J. (ed.) Voices of a Margin, Rockhampton: CQU Press.

- WORLD WIDE WEB PAGE

Young, C. (2001) English Heritage position statement on the Valletta Convention, [Online], Available: http://www.archaeol.freeuk.com/EHPostionStatement.htm [24 Aug 2001].

- CONFERENCE PAPERS

Hart, G., Albrecht, M., Bull, R. and Marshall, L. (1992) 'Peer consultation: A professional development opportunity for nurses employed in rural settings', Infront Outback - Conference Proceedings, Australian Rural Health Conference, Toowoomba, pp. 143–48.

- NEWSPAPER ARTICLES

Cumming, F. (1999) 'Tax-free savings push', Sunday Mail, 4 April, p. 1.

All the items cited in the main text, and no other items, must be placed in the list of references.

Authors should include 2-3 JEL codes with manuscript during submission. For more details on the JEL classification system CLICK:
http://www.aeaweb.org/econlit/jelCodes.php?view=jel.

Information about the journal and the deadlines for submitting articles for next issues are presented at

http://www.managerial.zarz.agh.edu.pl

IMPORTANT NOTE: Instead the traditional e-mail-based procedure, the authors are encouraged to use the new Online Submission System that has just been activated: https://journals.agh.edu.pl/manage.

Double blind peer review procedure

1. In order to assess a quality of submitted publication the Editorial Board consults at least two outside referees (not affiliated to any of the authors' institutions) which are recognized experts in the specific field.

2. At least one of the referees must represent a foreign institution (i.e. an institution located in other country than the home institution of each author).

3. The journal uses double blind peer review policy, i.e. neither the author nor the referee knows the identity of the other. In addition, each referee signs a declaration of no conflict of interests, whereby the conflict of interest is defined as a direct personal relationship (kinship to the second degree, legal relationships, marriage) or a professional scientific cooperation between the referee and the author which took place at least once in the past two years preceding the year of accepting the invitation to review.

4. Only written referee reports are considered (journal does not accept face-to-face or phone-call-based reports). Each report should clearly express the referee's final recommendation (i.e. whether the article should be accepted for publication, revised or rejected). The referees are kindly requested to fill the review form which can be found in "For reviewers" section. In general, the referees are asked to:

– assess:
  • the scientific importance of the submission's topic,
  • the quality of research;

– verify whether:
  • the Abstract is concise and informative,
  • the facts and interpretations are satisfactorily separated in the text,
  • the interpretations and conclusions follow from the data,
  • the length and structure of the paper is appropriate,
  • the paper can be shortened without loss of quality,

- all the tables and figures are necessary,
- the diagrams and photographs are of good quality,
- there are all essential figures that should be prepared,
- all the references are exact,
- the manuscript requires proof reading by native speaker,
- there is sufficient attention given to previous research.

5. The names of the referees of particular articles are classified.

6. Once a year the journal publishes the complete list of referees.