Impact of Selected Determinants of Innovation on the Economic Growth of the Visegrad Group Countries

Katarzyna BROŻEK

Kazimierz Pulaski University of Technology and Humanities in Radom, Radom, Poland
k.brozek@uthrad.pl

Abstract. The proposed article deals with the innovativeness of the Visegrad Group countries in the context of economic growth. Because of the clarity of the considerations, the work was divided into several basic parts. In the theoretical part, economic growth was discussed in a synthetic way, while further considerations were devoted to econometric modeling. The whole was crowned with a summary and conclusions. There were two research methods used in the study: statistical data analysis and econometric analysis. The general purpose of the considerations was to identify statistically significant determinants of innovation affecting the economic growth of the V4 countries. The estimation was carried out using the classic method of least squares. As the explained variable, GDP was chosen, while the statistically significant variable, with the 0.01 significance level, turned out to be the variable unemployed with higher education at 10,000. residents. In contrast, statistically significant variables at the 0.05 significance level turned out to be internal R & D expenditure of the enterprise sector at 10,000. residents and internal government spending on R & D at 10,000 residents. The examined econometric model also includes state budget funds for industrial production and technologies, employed in science and technology as well as zero-one variables. Statistical data used in the study were taken from the European Statistical Office, the research period was set for 2005-2016.

Keywords: Econometric Analysis, Innovation, Visegrad Group.

1   Introduction

In the modern world, various types of phenomena, e.g. economic ones [more broadly: 1; 16; 20], social or natural are almost always conditioned by the action of other phenomena. Therefore, the existence of relationships between phenomena is often the subject of scientific inquiries. Thus, in the presented article, it was decided to examine one of many ever-striking phenomena, namely the problem of economic growth [see 6; 12].

The paper analyzes the impact of selected factors on economic growth in countries belonging to the Visegrad Group, i.e. in Poland, Slovakia, the Czech Republic and Hungary [cf. 14]. Therefore, the study covered four countries with diversified economic level, while the research period was limited to twelve years, i.e. from 2005 to 2016. The

study did not include the last year, i.e. 2017 due to the lack of data at the time of preparing the proposed analysis. It is worth noting that work on the model used GNU Regression Econometric and Time-Series Library - gretl, which provides advanced econometric methods.

The main research goal was to identify statistically significant determinants of innovation affecting the economic growth of four countries belonging to the Visegrad Group. However, it was decided to extend the analysis to include specific objectives, which are an attempt to find answers to the following research questions:
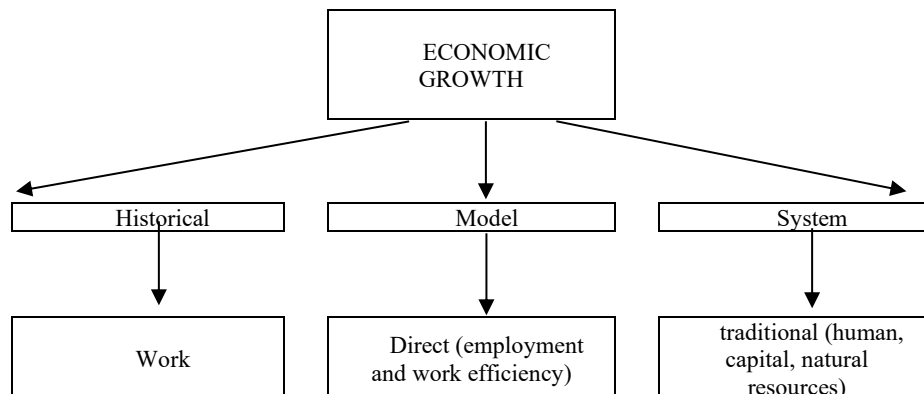
- Is the panel analysis helpful in solving problems related to the presentation of factors determining economic growth in the Visegrad Group countries?
- Did all of the examined explanatory variables interact in a statistically significant way with the explained variable?
- Which of the studied variables had a strong influence on the shaping of the explained variable?
- Which of the zero-one variables turned out to be statistically significant for the estimated model?

Based on the results obtained from the conducted analysis, which consisted in building and estimating the econometric model, all the research goals were achieved.

2   Economic Growth in a Synthetic Theoretical Approach

Economic growth can be seen as increasing the capacity of a given society to produce goods and services within the entire economy. Therefore, in the theory of economics, the term economic growth is used to describe quantitative changes. However, it should be noted that in the case of defining economic development, in addition to the aforementioned aspect, qualitative changes and changes in the structure of the economy are also taken into account [5]. The most frequently used quantitative reflection of economic growth is the increase in the real value of GDP in the economy.

In the subject literature one can find three basic accounts of factors affecting economic growth. This aspect is presented in Figure 1.
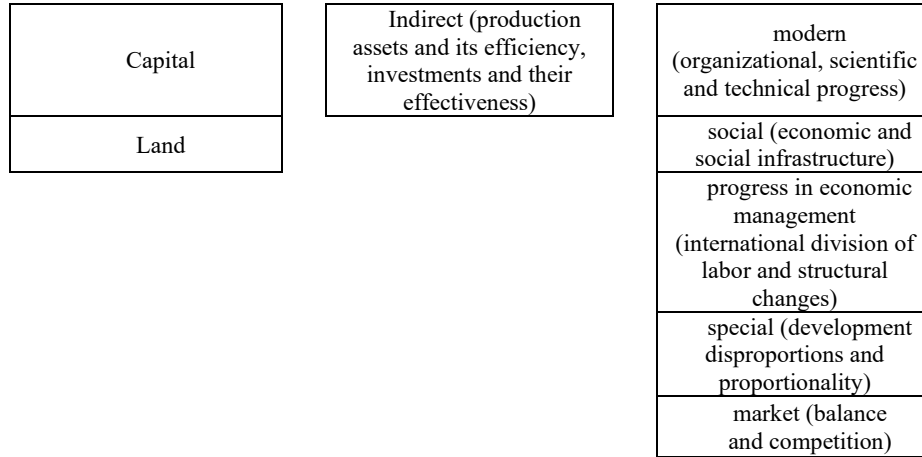
Capital	Indirect (production assets and its efficiency, investments and their effectiveness)	modern (organizational, scientific and technical progress)
Land		social (economic and social infrastructure)
		progress in economic management (international division of labor and structural changes)
		special (development disproportions and proportionality)
		market (balance and competition)

Fig. 1. Three approaches to economic growth factors [13, 17].

The fundamental aspects determining economic growth are labor productivity and employment growth. These elements constitute the national income. Growth also generates a combination of such factors as the process of creating a convergent to a country-specific path of sustainable growth, showing the development of *per capita* production; technological development; changes in expenditures on human capital, changes in investments and paths of sustainable growth [10].

Observing a specific industry, it is noticeable that the average increase in productivity can be presented as the merging of the following elements [10]:

- increase in participation in the market of enterprises with high productivity;
- increase in productivity inside enterprises already operating;
- building new business entities that replace entities with lower productivity.

3   Econometric Analysis

3.1   Constructing the Model

The econometric analysis was carried out on panel data, i.e. those that are observed in at least two dimensions [11]. In other words, this type of data is a two-dimensional variable, conditioned in time and space.

The study decided to examine selected determinants of innovation [more 7; 21] in the context of economic growth of the Visegrad Group countries in 2005-2016. Value gains and zero-one variables were used. Assuming that the index $i = 1, 2, ..., N$ marks the following areas (states V4), and the index $t = 1, 2, ..., T$ marks time units (see Table 1), then the constructed model will be in the form:

$$Y_{it} = \alpha_1 X_{1it} + \alpha_2 X_{2it} + \alpha_3 X_{3it} + \alpha_4 X_{4it} + \alpha_5 X_{5it} + Cz + H + P + v_{it}, \qquad (1)$$

where:

explained variable:

Y_{it} – Gross Domestic Product at current market prices at 10,000 residents,

explanatory variables:

X_{1it} – Internal expenditure on R & D of the enterprise sector at 10,000 residents,

X_{2it} – Internal government spending on R & D at 10,000 residents,

X_{3it} – State budget funds for industrial production and technologies for 10,000 residents,

X_{4it} – Unemployed people with higher education at 10,000 residents,

X_{5it} – Employed in science and technology for 10,000 residents,

Cz – zero-one variable for the Czech Republic,

H – zero-one variable for Hungary,

P – zero-one variable for Poland,

v_{it} – total random error (consisting of a purely random part ε_{it} and individual effect u_i, therefore $v_{it} = \varepsilon_{it} + u_i$) [15].

Table 1. Assignment of indices to individual countries of the Visegrad Group and respective periods.

	i	t
1	Czech Republic	2005
2	Hungary	2006
3	Poland [vide 18]	2007
4	Slovakia [vide 2; 18]	2008
		2009
		2010
		2011
		2012
		2013
		2014
		2015
		2016

When starting work related to econometric modeling, a set of statistical data should be first developed. Each data set is classified into one of three types: cross-sectional data, time series and panel data. In the estimated case, the latter ones were used, in other words, space-time, i.e. observed in many areas for many periods of time [9]. It is worth emphasizing that the data type declaration is an important element during process modeling, because a particular type of data determines the right methods and tools for the selected data type [15]. The empirical studies used statistical data taken from the Eurostat database [8].

In the next step, we should focus on the organization of panel data. The indices of units and (for individual V4 countries) and time t (2005-2016) were used.

The results of the model described above are presented in the following tables. The results of the most important and, at the same time, necessary tests are put below them.

3.2   Results of the Model Estimation

Estimation using the classical least squares method (CLSM) is considered acceptable when the individual effect does not occur and the panel is treated as a cross-sectional data set. This kind of situation happens in the examined model [vide 19].

Using the GRETL econometric program, the estimation presented in table 2 was obtained. It contains the values characterizing and describing the results of least squares estimation. The test for distribution normality is presented in Figure 2.

Table 2. Model 1: Panel estimation LSM, using 48 observations, 4 units of cross-section data were included, Time series length = 12, Dependent variable (Y): Yit [own elaboration based on the GRETL program].

	Factor	Std. error	Student's t-	P value	
const	6,03717	2,03689	2,964	0,0052	***
X1it	39,9913	17,1090	2,337	0,0246	**
X2it	39,5526	17,3784	2,276	0,0284	**
X3it	8,01838	10,7498	0,7459	0,4602	
X4it	-401,124	144,454	$-2,777$	0,0084	***
X5it	172,846	108,150	1,598	0,1181	
Cz	-4,20596	2,84193	$-1,480$	0,1469	
H	-5,21639	2,72466	$-1,915$	0,0629	*
P	-3,25356	2,73545	$-1,189$	0,2415	
The arithmetic mean of the dependent variable	5,189866	Standard deviation of a dependent variable			7,430614
The sum of residual squares	1666,281	Standard error of residues			6,536449
R-squared determination coefficient	0,357902	Adjusted R-square			0,226190
F(8, 39)	2,717303	The p-value for the F test			0,017578
Logarithm of credibility	$-153,2406$	Akaike information criterion			324,4812
Schwarz Bayesian criterion	341,3220	Hannan-Quinn criterion			330,8454
Autocorrelation of residues - rho1	$-0,078460$	Durbin-Watson statistics			2,004700

The first step in assessing the quality of the econometric model is to assess the significance of the impact of individual independent variables on the dependent variable. For this purpose, the Student's t-test of significance of the α parameter was used. If a given variable is marked with three asterisks (***), the variable is significant at the 1% significance level, if two (**), at 5%, and if one (*) is at 10% [15].

In table 2 in the penultimate column there is an empirical significance level p (test probability), which additionally allows to indicate the weakest model variable. However, the F-Snedecor test (Statistics F) enables a comprehensive assessment of the suitability of the econometric model.

After the panel estimation, the main model took the following form [vide 3;4]:

$$Y = \underset{(2,03)}{6} + \underset{(17,10)}{39} X_1 + \underset{(17,37)}{39} X_2 + \underset{(10,74)}{8} X_3 - \underset{(144,45)}{401} X_4 + \underset{(108,15)}{172} X_5 - \underset{(2,84)}{4} Cz$$
$$- \underset{(2,72)}{5} H - \underset{(2,73)}{3} P \tag{2}$$

The equations describing particular countries of the Visegrad Group are as follows:

- Poland:
$$Y_P = 2{,}79 + 39X_1 + 39X_2 + 8X_3 - 401X_4 + 172X_5 \tag{3}$$
- Hungary:
$$Y_H = 0{,}82 + 39X_1 + 39X_2 + 8X_3 - 401X_4 + 172X_5 \tag{4}$$
- Czech Republic:
$$Y_{Cz} = 1{,}83 + 39X_1 + 39X_2 + 8X_3 - 401X_4 + 172X_5 \tag{5}$$
- Slovakia:
$$Y_S = 6{,}04 + 39X_1 + 39X_2 + 8X_3 - 401X_4 + 172X_5 \tag{6}$$

On the basis of the conducted research, it can be concluded that three of the analyzed variables, i.e.: unemployed with higher education at 10,000. inhabitants, internal expenditure on R & D of the enterprise sector for 10,000 residents and internal government spending on R & D at 10,000 people are good stimulators for the Gross Domestic Product in the Visegrad Group countries. This is confirmed by the level of their significance: X_{4it} - 0,01; X_{1it} and X_{2it} - 0,05.

Excluding the constant, the largest p value is for variable 6 (X_{3it}).

The evaluation of the normality of the distribution of residuals was carried out using the Doornik-Hansen test verifying the hypothesis on the normality of the residual distribution.

Test for normality of residual distribution:

Null hypothesis: the random component has a normal distribution.

Test statistics: Chi-square (2) = 0.701282, with a p value of 0.704237.

The Chow test for structural changes in the distribution of the sample in the observation 2:12

Null hypothesis: no structural changes.

Test statistics: F(6, 33) = 0,871212,

with p value = P(F(6, 33) > 0,871212) = 0,52633.

Evaluation of homogeneity of the variance of the random component - heteroskedasticity of the random component was done with the White test. This test assumes verification of the significance of regression determined for residual squares with a set of model variables, their squares and products.

White's test on heteroscedasticity of residues (variability of residual variance):

Null hypothesis: heteroscedasticity of residues does not occur.

Test statistics: LM = 41,1026,

with p value = P(Chi-square(38) > 41,1026) = 0,336254.

Cumulative distribution of Waltz's heterokedasticity test:

Null hypothesis: units have a common residual variance.

Asymptotic statistics: Chi-square(4) = 2,83326, with p value = 0,586105.

Panel residual variance = 34,7142.

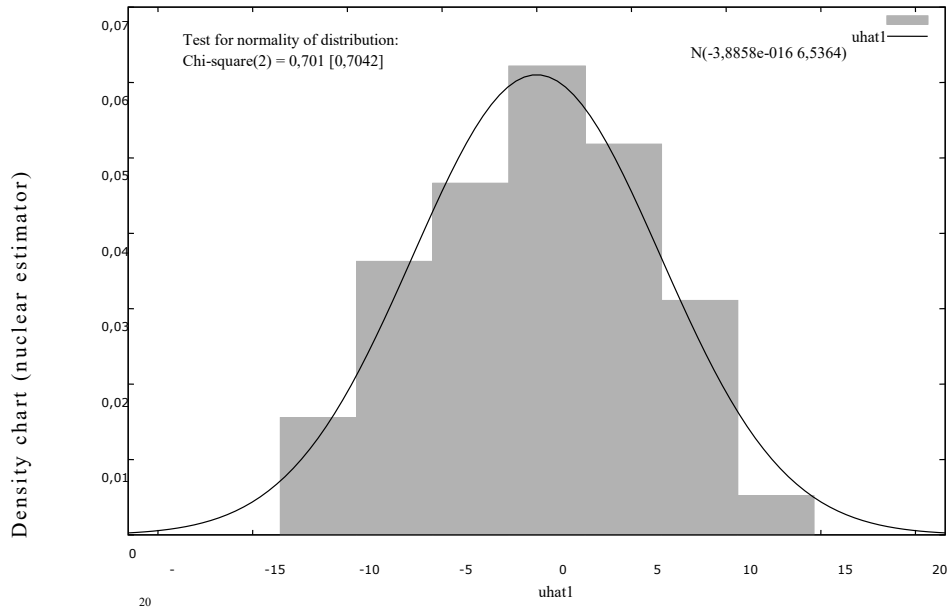| unit	variance
1	48,7280 (T = 12)
2	36,3393 (T = 12)

3   29,5124 (T = 12)
4   24,2771 (T = 12)

Fig. 2. Test for normality of distribution [own elaboration based on the GRETL program].

The evaluation of the normality of the distribution was carried out using the Doornik-Hansen test, i.e. a test using the transformed value of skewness and kurtosis parameters.

In the next part of the analysis it was decided to work out the frequency distribution (Table 3) and make the alignment assessment.

R-squared determination coefficient = 0,856305.

Test statistics: TR^2 = 41,102638,

with p value = P(Chi-square(38) > 41,102638) = 0,336254.

Frequency distribution for uhat1, observations 1-48, number of intervals = 7, mean = -3,88578e-016, standard deviation = 6.53645.

Table 3. Frequency distribution [own elaboration based on the GRETL program].

Intervals	Average	Number	Frequency	Cumulated	
< -9,5064	-11,518	3	6,25%	6,25%	**
-9,5064 - -5,4838	-7,4951	7	14,58%	20,83%	*****
-5,4838 - -1,4612	-3,4725	9	18,75%	39,58%	******
-1,4612 - 2,5614	0,55006	12	25,00%	64,58%	*********
2,5614 - 6,5840	4,5727	10	20,83%	85,42%	*******
6,5840 - 10,607	8,5953	6	12,50%	97,92%	****
>= 10,607	12,618	1	2,08%	100,00%	

Null hypothesis: an empirical distributor has a normal distribution.

The Doornik-Hansen Test (1994) - transformed skewness and kurtosis:
Chi-square (2) = 0.701, with a p value of 0.70424.

Another very important issue in econometric modeling is the study of the collinearity of explanatory variables. This is extremely important, because the possible collinearity of variables is an undesirable feature. Because if there is an exact collinearity, the model will not be estimated because the determinant of the X^TX matrix is zero. The high correlation of explanatory variables causes that the determinant value is close to zero, thus the standard errors of parameter evaluations, derived from the variance and covariance matrix, have relatively large values, which in turn leads to understating the value of Student's t-statistics in the parameter significance assessment. If the value of VIF_j equals one, it means that the variable x_j is uncorrelated in relation to the other explanatory variables of the model. A dramatically different situation occurs when VIF_j > 10 then it is a sign of collinearity, which significantly disturbs the quality of the analyzed model [15].

Collinearity Rating VIF (j) - factor of variation distortion:
VIF (Variance Inflation Factors) - the minimum possible value = 1.0.
Values greater than 10.0 may indicate a collinearity problem - distortion of variance.
X_{1it} - 1,227; X_{2it} - 1,383; X_{3it} - 1,293; X_{4it} - 1,036; X_{5it} - 1,052; Cz - 1,701; H - 1,564; P - 1,576.

The results indicate that there is no collinear problem in the model under study because all tested values are less than 10.
$VIF(j) = 1/(1 - R(j)^2)$,
where $R(j)$ is the multiple correlation coefficient between the variable 'j' and the other independent variables of the model.
Matrix features X'X:1-norm = 88,267132, determinant = 0,0016567627, indicator of the conditioning of the CN matrix = 2,2254459e-005.

The next stage of the panel model research was devoted to taking into account individual effects, which can be twofold, which is why they are divided into fixed effects and random effects. However, due to the breadth of the proposed analysis, it was decided to limit only to the study of established effects. The results obtained are shown in Table 4.

Table 4. Estimated effects (non-random effects) taking into account the diversity of the free expression according to units in the cross-section, [own elaboration based on the GRETL program].

	Coefficient	Standard error	p value	
const	2,8682	1,4041	0,04789	**
X1it	39,991	17,109	0,02464	**
X2it	39,553	17,378	0,02842	**
X3it	8,0184	10,75	0,46020	
X4it	-401,12	144,45	0,00839	***
X5it	172,85	108,15	0,11807	
The arithmetic mean of the dependent variable	5,189866	Standard deviation of a dependent variable	7,430614	

The sum of residual squares	1666,281	Standard error of residues	6,536449
LSDV R-square	0,357902	Within R-square	0,325046
LSDV F(8, 39)	2,717303	The p-value for the F test	0,017578
Logarithm of credibility	−153,2406	Akaike information criterion	324,4812
Schwarz Bayesian criterion	341,3220	Hannan-Quinn criterion	330,8454
Autocorrelation of residues - rho1	−0,078460	Durbin-Watson statistics	2,004700

4 group mediums including data:
Residual variance: 1666,28 / (48 - 9) = 42.7252.
Total significance of grouped inequalities:
F (3, 39) = 0, with the value of p 1.
A low p value means rejecting the H0 hypothesis that the LSM panel model is correct, against the H1 hypothesis, that the model with the determined effects is more appropriate.
Average residuals for section units in the LSM panel estimation:
  unit 1: 1,4803e-015
  unit 2: 6,6613e-016
  unit 3: 2,3685e-015
  unit 4: -5,9952e-015
It should be noted that in order to determine whether a given panel model can be estimated using LSM, one should verify the hypothesis about the existence of an individual effect or, concurrently, whether the variance of the individual effects component is equal to zero. Breusch-Pagan's test serves this purpose.

Breusch-Pagan test statistics:
LM = 2.18182 with the value of p = probe (chi-square (1) > 2.18182) = 0.139649.
A low p value means rejecting the H0 hypothesis that the LSM panel model is correct, against the H1 hypothesis, that the random effect model is more appropriate.

Summary and Conclusions

After the analysis, it can be concluded that the chosen model is suitable for forecasting. Because the share of the standard error of the residues to the arithmetic mean of the dependent variable does not exceed 10%.

After analyzing the value of the coefficients of determination to the minimum value adopted from above, it can be stated that the tested model has a high fit.

The distribution of residues is consistent with the normal distribution, moreover, the constructed model does not contain autocorrelations of residues.

The relationship between variables is linear.

In addition, on the basis of the arguments quoted below, it is considered that the specificity of the model is correct because:

- a set of explanatory variables was selected appropriately;
- there is correctness of the functional form;
- zero-one variables were used.

To sum up, the aim of the research was to identify statistically significant determinants of innovation affecting the economic growth of the V4 countries. Therefore, an analysis of the relationship between selected variables regarding the level of innovation and changes in economic growth in 2005-2016 was conducted. A panel model was built, using annual data describing four selected for analysis states. The estimation of the model was made using the classic least squares method using the GRETL program.

The results of the analysis presented in the paper allow to formulate the following conclusions:

- panel analysis is useful for solving problems related to the search for determinants shaping economic growth in the V4 countries;
- a statistically significant variable with a 0.01 significance level turned out to be a variable unemployed with a higher education of 10,000. residents;
- two variables should be included in the factors determining the determinants of economic growth in the analyzed countries: $X1_{it}$ - internal expenditures on R & D of the enterprise sector and $X2_{it}$ - internal expenditures on R & D of the government sector. The increase in the level of these variables has a positive impact on economic growth in the Visegrad Group countries. This identification makes it possible to improve national policies and to focus actions on those that most significantly improve the socio-economic situation of the V4 countries;
- the variables $X1_{it}$ and $X2_{it}$ had a statistically significant effect on the change in the level of economic growth, with the $X4_{it}$ variable having a stronger impact on the shaping of the explained variable;
- the zero-one variable for Hungary turned out to be a statistically variable being at the significance level of 0.1.

References

1. Beugelsdijk, S., Brakman S., Garretsen, H., Marrewijk, Ch.: International Economics and Business Nations and Firms in the Global Economy, pp.392-408. Cambridge University Press, Cambridge (2014).
2. Biea, N.: Econometric Growth in Slovakia: Past Successes and Future Challenges. EU Economic Brief, No. 008. Luxemburg: Publications Office of the European Union (2006).
3. Brożek, K., Kogut, J.: Ekonometryczna analiza zarządzania działalnością B+R w polskich przedsiębiorstwach. Education of Economists and Managers no 2(44) 2017 pp. 145-160 (2017).

4. Bukowski, S.I., Garlińska-Bielawska, J.: Eksport do Niemiec jako czynnik wzrostu gospodarczego w Polsce w latach 2000-2012, Acta Universitatis Lodziensis Folia Oeconomica, Lodz (2014).

5. Bywalec, C.: Konsumpcja a rozwój gospodarczy i społeczny, p. 61. Wydawnictwo C.H. Beck, Warsaw (2010).

6. Chao, X.: The Dynamic Evolution of the Distribution of the Quality of Economic Growth in China: 1998-2014, pp.310-317. In: Hradec Economy Days 2017 vol. 7 (1), University of Hradec Králové, Hradec Králové (2017).

7. Dymitrowski, A.: The Role of Innovations. Created in the Internationalization Process for Company Performance, pp. 95-120. PWN, Warsaw (2014).

8. European Statistical Office, https://ec.europa.eu/eurostat, last accessed 2018/08/10.

9. Franc-Dąbrowska, J.: Praktyczne zastosowanie wybranych modeli panelowych do oceny sytuacji finansowej przedsiębiorstw rolniczych, https://wne.sggw.pl, last accessed 2018/11/26.

10. Gnypowa, B. (eds.): Zrozumieć wzrost gospodarczy. Analiza na poziomie makroekonomicznym, poziomie branży i poziomie firmy, pp. 19, 87. Oficyna Ekonomiczna, Cracow (2005).

11. Górecki, B. R.: Ekonometria podstawy teorii i praktyki, p. 17. Wydawnictwo Key Text, Warsaw (2013).

12. Kang, M.: Factor analysis and quality evaluation for economic growth with investment-growth model. The Journal of Quantitative & Technical Economics, (2) pp. 153-160 (2006).

13. Korez-Vide, R. & Tominc, P.: Competitiveness, Entreneurship and Economic Growth. In: P. Trąpczyński, L. Puślecki, M. Jarosinski (Eds.), Competitiveness of CEE Economies and Businesses, pp. 25-44. Basel. Springer International Publishing AG (2016).

14. Kowalska, A., Kovárník, J.: The Innovativeness and Competitiveness of the Visegrad Group Countries in the years 2011-2016 - Selected Indicators, pp. 460-470. In: Hradec Economy Days 2018 vol. 8 (1), University of Hradec Králové, Hradec Králové (2018).

15. Kufel, T.: Ekonometria, Rozwiązywanie problemów z wykorzystaniem programu GRETL, Wydanie trzecie zmienione, p. 33, pp. 57-58, pp. 64-65, p. 173. Wydawnictwo Naukowe PWN, Warsaw (2013).

16. Marakova, V., Dyr, T., Wolak-Tuzimek, A.: *Factors of Tourism's Competitiveness in the European Union Countries.* In: Ekonomika a Management 2016 no. 3, pp. 92-109 (2016).

17. Masoud, N., A contribution to the Theory of Economic Growth: Old and New. Journal of Economics and International Finance, 6(3), pp. 47-61 (2014).

18. Meyer, N., Jongh, J.: Entrepreneurial Activity and Its Relationship to Economic Growth and Development: A Comparative Study between Slovakia and Poland, In: M. Frankovsky, J. Dobrovic, R. Fedorko (Eds.), Management and the World in Motion, Challenges, Opportunities and Threats, Bookman s.r.o., Presov (2018).

19. Meyer, N., Meyer, D.F. An Econometric Analysis of Entrepreneurial Activity, Economic Growth and Employment: The Case of the BRICS Countries. International Journal of Economic Perspectives, Vol. 11(2), pp. 429-441 (2017).

20. Özcelik, E., Taymaz, E.: Does Innovativeness Matter for International Competitiveness in Developing Countries? The Case of Turkish Manufacturing Industries. Research Policy 33(1), pp. 409-424 (2004).

21. Sipa, M.: Diversification of Indexes Determining Innovation of Economies – The Visegrad Group Countries. In: Hradec Economy Days 2015 vol. 5, University of Hradec Králové, Hradec Králové (2015).