Factor Analysis as an Alternative Approach to Measurement of Human Development

I. Majerova, J. Nevima

***Abstract*—** The alternative indicators are increasingly used for measuring the socio-economic level of individual economies. One of them is the Human Development Index, which measures this level not only by GDP per capita, but also trough the indicators in the field of education and healthy life. This index was originally designed for comparing the level of human development in the economies as a whole. Because of differences in this development are not only at the national level, it is an effort to measure them also at level of regions. The aim of this article is, using the quantification of regional human development, to determine the most significant indicators of human through factor analysis. The regions of the Visegrad Group countries at NUTS II level were selected for this purpose. The research was made in the period from 2004 till 2013. There is initially set the assumption that the most important factor of human development is economic level, measured by gross domestic product per capita. It was not confirmed and was found that the most important role is played by another factor, namely lifelong learning.

***Keywords*—**Factor analysis, human development, Human Development Index, modified Human Development Index, Visegrad Group.

I. Introduction

The most widely used indicator for measuring economy´s state of affairs is GDP [1], [1], neither it does not include the social, political, cultural and environmental aspects of development. Many alternatives can be applied for measurement of socio-economic development, the best-known and most used is an index called the Human Development Index [3].

The Human Development Index (HDI) has been used by the United Nations since 1990, clearly brings a different perspective on development issues. This index should be better able to emphasize the effect of other than just monetary (economic) factors of economy of a country. The basis of the HDI index is a greater explanatory power, which is to follow economic development or sustainable development in general. The measurement of human development through the HDI is an alternative to the GDP/GNI per capita as a measure of human well-being in the last years.

According to [4] HDIs are primarily nation level indicators, estimated for the country as a whole. The constructions of the HDI do not express the differences in regions of countries. However, the regional disparities exist in developing and developed countries as well and they influence regional development. The north-west/south-east division exist in the European regions and such differences in regional development among them would weaken social cohesion in Europe [5].

For the purpose of this paper, we decided to analyse this issue for a group of countries of the Visegrad Group (V4) at the NUTS II level. This group includes the countries, as Czech Republic, Hungary, Poland and Slovakia. There are 35 regions on the NUTS II level – eight in the Czech Republic, seven in Hungary, sixteen in Poland, and four in Slovakia. Firstly we have chosen the indicators for the construction of the modified Human Development Index (NHDI). Secondly we used the factor analysis for reveal a certain structure in data which is not visible at first side. In case of NHDI, at first we will search what is characteristic for NUTS II regions, resp. which variable is dominant for them.

The research was made in the period from 2004 till 2013. There is initially set the assumption that the most important factor of human development is economic level, measured by gross domestic product per capita.

II. Data of Human Development

The beginning of the Human Development Index dates back to 1990 when the UN Development Programme (UNDP) published the first report on human development (Human Development Report) which established the need of human development measurement. Human development has two forms, which should be in balance, the formation of human capabilities in terms of improving health, increasing knowledge and skills to meet human need and their own skills and competences, free time, job security, cultural, social and political events. Basically, human development is clearly and directly dependent on income. It is therefore necessary to examine other variables that point out the potential of a country much better as well as the options currently appear in human development [6].

This work was supported by the students´ grant project "Influence of Selected Macroeconomic and Microeconomic Determinants on the Competitiveness of Regions and Firms in Countries of the Visegrad Group Plus". Project registration number is SGS/13/2015.

I. Majerova, is with the Department of Economics and Public Administration, Silesian University in Opava, School of Business Administration in Karvina, Univerzitní namesti 1934, 73401 Karvina, Czech Republic (corresponding author to provide phone: 00420-59638256; e-mail: majerova@opf.slu.cz).

J. Nevima, is with the Department of Economics and Public Administration, Silesian University in Opava, School of Business Administration in Karvina, Univerzitni namesti 1934, 73401 Karvina, Czech Republic (corresponding author to provide phone: 00420-59638258; e-mail: nevima@opf.slu.cz).

A. Data of Human Development Index

The Human Development Index (HDI) is a summary measure of achievements in key dimensions of human development: a long and healthy life, an access to knowledge and a decent standard of living [7.

These three dimensions have four parts - health and standard of living has one part each and education has two parts:

1. Health dimension – life expectancy at birth (interval 20-85 years)
2. Education dimension – expected year of schooling (0-18 years) and mean years of schooling (0-15 years)
3. Standard of living dimension – GNI per capita in USD/PPP (100-75.000).

Because of the need to improve their explanatory power, the calculation method of two of three dimensions (health indicator index is the only one which has remained unchanged) has changed over time. The literacy rate of population has been replaced by an indicator of expected years of schooling, the combined gross enrolment by the mean number of years of education (knowledge dimension). The dimensions of living standards are now measured by GNI per capita in purchasing power parity to the USD.

HDI index calculation required the values in the range from 0 (the lowest level of human development) to 1 (the highest human development), and therefore they were determined for each dimension of the minimum and maximum values (more in [8]) based on historical evidence.

B. NUTS Human Development Index

For the purpose of the paper, we adopted the same principle of HDI creating for the national level – the health dimension, knowledge dimension and dimension of a living standard. Components of each dimension, however, had to be modified because of the lack of data at the regional level (NUTS II level). Data were used from a regional database of Eurostat and converted to the number of inhabitants representing the given group

The construction of the HDI of V4 regions (NHDI) was as follows:

1. Health with the value of life expectancy (LE) at birth that represents, according to Eurostat, the mean number of years that a newborn child can expect to live if subjected throughout his life, to the current mortality conditions

2. Knowledge, which includes two components:

   a. Tertiary educated people in the age of 25-64 (TE), where the indicator is defined as a number of population aged 25-64 who have successfully completed tertiary studies (e.g. university, higher technical institution, etc.).

   b. Lifelong learning (LL) in the form of amount of people participate in education and training that covers participation in formal and non-formal education and training for the age group of 25-64 are presented.

3. Standard of living, measured through GDP per capita in PPS (GDP).

These indicators were chosen for their greatest explanatory power in relation to human development. The life expectancy at birth reflects the level of health and quality of life and measures the qualitative aspects of living a healthy life. It correlates positively with human development – the higher the healthy life expectancy of region, the more developed it is.

The share of tertiary educated people in productive age on the population in this age group is connected with the ability of people to reflect the needs of knowledge of economy and to contribute to it and human development. Lifelong learning, in the form of participation in education and training, encompasses all learning activities undertaken throughout life (after the end of initial education) with the aim of improving knowledge, skills and competences, within personal, civic, social or employment-related perspectives [9]. Due to lifelong learning people extend their possibilities for increasing their incomes. As a dimension of health, both indicators of education are positively correlated with human development.

The last but not least dimension is the GDP per capita. The implementation of this indicator was influenced by the opinion of [10] who considered the income (product) as a primarily mean to achieve human development. The GDP per capita reflects the economic level better than its absolute value. The indicator is measured by an artificial European currency unit, the purchasing power standard (PPS).

III. Methodology

Factor analysis (FA) belongs to so called data-mining method. It is a posterior method [11], which is applied with the aim to reveal a certain structure in data which is not visible at first side. In case of NHDI, at first we will search what is characteristic for NUTS 2 regions, resp. which variable is dominant for them. From mathematical point of view, we explain observed variables as linear combination of factors plus certain error at factor analysis.

According to [12] this relation we can describe asset of regression equations (1)

$$\begin{aligned} Y_1 &= \beta_{11}X_1 + \beta_{12}X_2 + \ldots + \beta_{1p}X_p + e_1 \\ Y_2 &= \beta_{21}X_1 + \beta_{22}X_2 + \ldots + \beta_{2p}X_p + e_2 \\ &\vdots \\ Y_k &= \beta_{k1}X_1 + \beta_{k2}X_2 + \ldots + \beta_{kp}X_p + e_k \end{aligned} \quad (1)$$

where Y_i is the i-manifest variable, X_j is j- extracted factor, β_{ij} is estimated regression coefficient of load for the j-factor and i-manifest variable and e_i is the residual variable, ie. the unexplained part of the variance for the i-manifest variable.

Among advantages of FA using the fact that it comes out of factually found correlations among observed phenomena

belongs and does not provide their clustering according their outer similarity. The certain disadvantage can also be counting outliers values which disrupt integrity of data file. Practical calculations for needs of the case study will be provided by SW product SPSS (version 15.0.1).

However, at first we will introduce steps when applying FA. At first we will verify adequacy of data selection, resp. ratio of measure and ideal value. It can be provided by Kaiser-Meyer-Olkin rate (KMO), where a value 0.6 is recommended. Simultaneously, in this phase we verify if there is multi-correlation (except for correlation itself we verify the relationships toward the other variables) in the model.

Factor analysis has a sense if there is a significant multi-correlation in the model. FA at first standardizes initial values. It means that from measured values total average is subtracted and the difference is divided by standard deviation. Each of the factors is calculated gradually and extraction of factors is a way how to select factors from a set of variables. In our case, we will use Principal Component Analysis [13]. To match a variable to a factor correctly, its factor loading over 0.3 is required.

In the next part of FA factor optimization is provided. We will reach it by so called factor rotation [14]. The procedure how to determine final number of factors comes out of implementation of so called Kaiser Normalization. The rule says that into FA selection the factors are involved whose variance is over 1, which means that the factor contents at least one strongly differentiating variable [15]. The variance of factor can be determined by Cattel scree plot. Naming of found factors is derived from our ability to penetrate into substance of solved research problem and reach certain generalization among input variables which reached high value of variance. Now we will continue to practical application of factor analysis in selected regions.

IV. Results

When applying factor analysis we will observe which factor for each of 35 NUTS 2 regions is dominant. This is the main motivation of further steps. Data used for this analysis are shown in the Appendix – there is an example of the years 2004, 2009 and 2013 in the table.

At first the sense of provided FA is verified through KMO ratio and established result is recorded in the first column of the Table I. For each of evaluated regions it was found out that FA has a sense. For all calculations method of main components was used, which serves to factor extraction, as we tend to aim the first factor to gain the most of variance.

With regard to nature of data file we have chosen Oblimin method. The Oblimin method leads to the simplest structure of factors and is most often used when it is obvious that factors with certain rate of probability cannot be dependent on each other.

The second column of the Table I illustrates which factor was crucial for the region. The number of was found out by application of Kaiser Normalization and proved by Cattel scree plot. The name of factor complies with regard to illustrative view of FA with input variables – it is not necessary to name it. In the case of multifactor solution is important given the first place.

TABLE I. RESULTS OF FACTOR ANALYSIS OF V4 REGIONS

KMO	Component	Statistical name of region
0.631	TE, LL, GDP	CZ01
0.633	TE, GDP	CZ02
0.777	LL, GDP	CZ03
0.623	TE, GDP, LL	CZ04
0.674	TE	CZ05
0.778	GDP, LL	CZ06
0.832	LE	CZ07
0.648	GDP, LL	CZ08
0.776	LE	HU10
0.813	GDP, LL	HU21
0.621	GDP, LL	HU22
0.645	TE, LL	HU23
0.685	TE, GDP	HU31
0.704	TE, LL	HU32
0.734	GDP, LL	HU33
0.638	LE, LL	PL11
0.664	LE, LL	PL12
0.780	LE, LL	PL21
0.627	LE, LL	PL22
0.766	TE, LL	PL31
0.692	LE, LL	PL32
0.677	TE, LL	PL33
0.639	GDP, LL	PL34
0.749	TE, LL	PL41
0.737	TE, LL	PL42
0.775	TE, LL	PL43
0.850	GDP	PL51
0.611	GDP, LL	PL52
0.716	TE, LL	PL61
0.686	TE, LL	PL62
0.787	TE	PL63
0.778	LE, LL	SK01
0.617	TE, LL	SK02
0.717	TE, LL	SK03
0.780	GDP, LL	SK04

Extraction Method: Principal Component Analysis
Rotation Method: Oblimin with Kaiser Normalization

There are shown very interesting results in the table. The factors of NHDI play various role in the regions of V4. Whereas the GDP per capita is the dominant factor in the Czech Republic, followed closely by lifelong learning, in other regions it is precisely the lifelong learning. The result of factor analysis is surprising in regions of Poland and Slovakia, where the importance of the economic level in terms of GDP per

capita is insignificant, and educational site leads (both the rate of lifelong education and rate of tertiary education).

Taking into account the Visegrad Group countries as a whole, the most important factor in human development in our comparison is lifelong learning - in 29 regions, further the indicator of GDP per capita and tertiary education, and so in 15 regions. Life expectancy at birth plays the least important role (in eight regions).

V. Conclusion

The Human Development Index is one of the indicators which can measure the socioeconomic development. This indicator has been used since 1990, it measures the above mentioned development at the national level and it is used to compare differences between economies. However, there are not only disparities between economies, but also within them.

For the purpose of our paper, we decided to analyse the problem of human development for a group of countries of the Visegrad Group at the NUTS II level, for 35 regions. We have chosen the indicators for the construction of the modified Human Development Index (NHDI). We used three components - the health dimension (life expectancy at birth), the knowledge dimension (tertiary educated people and people participate in education and training) and the dimension of living standard (GDP per capita). Than we used the factor analysis and for this reason that this analysis reveals a certain structure in data which is not visible at first side.

The results of the factor analysis show that the most important factor in human development is lifelong, further GDP per capita and tertiary education, and last life expectancy at birth.

The initially set assumption that the most important factor of human development is economic level, measured by gross domestic product per capita was not confirmed and was found that the most important role is played by another factor, namely lifelong learning. This confirms that the education of the population is a very important variable both regional and national significance.

Our future research will focus on the creating a cluster analysis for the values of individual components of the index, including the comparison of the both (factor and cluster) results.

APPENDIX I

THE VALUE OF NHDI COMPONENTS OF THE NHDI COMPONENTS IN REGIONS V4 IN THE YEAR 2004, 2009 AND 2013

Stat. name	Region	2004				2009				2013			
		LEB	TE	LL	GDP/c	LEB	TE	LL	GDP/c	LEB	TE	LL	GDP/c
CZ01	Praha	77.7	183.141	77.112	39.700	79.1	226.8864	80.8704	43600	80.1	283.6992	64.2756	46.000
CZ02	Střední Cechy	75.7	64.7163	27.4554	17.200	77.2	100.1052	47.8764	18300	78.2	148.0162	67.6858	19.500
CZ03	Jihozápad	76.4	75.2752	30.2445	17.100	77.6	99.0102	42.8342	17800	78.4	123.966	72.3135	19.400
CZ04	Severozápad	74.5	44.1388	31.1568	14.900	75.6	56.1204	47.4351	16100	76.4	80.5125	51.528	16.500
CZ05	Severovýchod	76.7	78.9694	42.005	15.600	77.9	111.0272	53.7788	16500	78.6	142.4136	111.8964	18.000
CZ06	Jihovýchod	76.5	125.9768	59.2832	16.200	78.1	161.8513	61.2928	18200	79.1	215.1972	87.6024	20.600
CZ07	Střední Morava	76.1	76.527	34.0893	14.400	77.2	90.7776	38.2968	16100	78.1	112.2984	54.0696	17.700
CZ08	Moravskoslezsko	75.2	71.5671	33.9763	15.600	76.2	101.6046	41.0742	16400	77.1	122.7072	74.6004	18.400
HU10	Közép-Magyarország	74.5	411.8196	95.772	23.300	75.9	483.3066	64.1121	26200	77	553.4108	70.0098	28.700
HU21	Közép-Dunántúl	72.9	80.9305	23.123	13.600	74	92.563	12.26	13100	75.5	114.893	12.094	15.600
HU22	Nyugat-Dunántúl	73.7	79.6367	17.2639	14.300	75	95.404	11.224	14700	76.1	99.9648	10.1088	17.900
HU23	Dél-Dunántúl	72.8	71.214	19.422	10.000	74	80.5086	11.0502	10700	75.2	93.4764	13.2808	11.900
HU31	Észak-Magyarország	71.5	85.05	23.1336	9.500	72.9	99.5241	15.1593	9500	74.2	106.5627	13.4001	10.500
HU32	Észak-Alföld	72	103.1814	22.9292	9.200	73.8	122.88	20.48	10100	75.4	144.536	27.608	11.300
HU33	Dél-Alföld	72.7	94.4667	21.2367	9.900	74.1	126.3325	15.1599	10300	75.6	128.219	17.6125	11.900
PL11	Lódzkie	73.4	231.694	79.8405	10.700	74.1	350.2074	67.6143	13300	75.4	389.818	51.4228	16.700
PL12	Mazowieckie	75.6	557.8207	150.6909	17.800	76.2	893.2398	219.4992	22500	77.7	1059.239	200.4774	28.500
PL21	Malopolskie	76.2	258.6465	69.0795	10.200	77.2	356.8999	67.4327	12800	78.5	458.185	80.9575	15.800
PL22	Slaskie	74.5	385.1188	134.5278	12.500	75.2	534.148	109.4352	15400	76.3	663.558	122.3775	18.600
PL31	Lubelskie	74.6	191.824	74.3318	8.100	75.6	267.8624	68.2534	9900	77.1	321.7845	61.8331	12.600
PL32	Podkarpackie	76	132.9548	33.7348	8.300	77.2	242.487	35.7957	10200	78.6	255.9249	28.8054	12.700
PL33	Swietokrzyskie	75.2	111.804	29.766	8.900	76.3	159.852	31.9704	11400	77.1	200.694	23.9289	13.100
PL34	Podlaskie	75.5	86.658	28.3305	8.600	76.6	134.246	24.976	10600	77.1	154.4337	21.8929	13.000
PL41	Wielkopolskie	75.2	248.5104	81.036	12.500	76	293.1775	61.9861	15500	77.2	415.5372	65.7046	19.300
PL42	Zachodniopomorskie	74.7	130.2132	45.0738	10.500	75.4	175.7946	46.3538	12500	76.7	196.8643	26.3584	15.100
PL43	Lubuskie	74.6	86.3786	29.8067	10.500	75.2	95.3876	19.3116	12300	76.3	118.4088	16.0944	15.000
PL51	Dolnoslaskie	74.8	247.0646	95.3857	11.900	75.3	312.8796	80.5902	15700	76.9	374.241	61.131	20.100
PL52	Opolskie	75.7	61.8975	22.68	9.600	76.9	85.8114	25.9553	12100	77.2	99.9647	14.8273	14.500
PL61	Kujawsko-Pomorskie	74.8	141.7	58.9472	10.000	75.4	187.9362	44.0838	12000	76.9	218.5713	41.1801	14.800
PL62	Warminsko-Mazurskie	74.5	97.8066	23.3568	8.700	75.3	156.3983	34.9316	10500	76.3	163.6542	22.1368	12.900
PL63	Pomorskie	75.7	165.584	50.304	11.400	76.5	225.8122	57.7908	14100	77.9	327.7959	71.2066	17.300
SK01	Bratislavský kraj	75.4	94.3525	42.2013	33.700	75.3	120.4863	27.9498	42500	78.1	141	26.696	49.000
SK02	Západné Slovensko	74.4	102.14	30.642	13.000	75.3	142.012	21.848	16200	76.8	182.5834	28.5974	18.800
SK03	Stredné Slovensko	73.7	87.4709	36.145	10.700	74.8	115.9982	17.6686	13900	76.2	147.8056	19.655	15.900
SK04	Východné Slovensko	73.9	85.707	15.219	9.900	75.1	109.6137	17.262	11800	76.2	158.0775	17.1627	13.800

Notice: LE in the years, TE and LL in the thousands of inhabitants, GDP/c in PPS

REFERENCES

[1] J. E. Stiglitz, A. Sen, and J. P. Fitoussi, "Report by the commission on the measurement of economics performance and social progress, " Commission on the Measurement of Economic Performance and Social Progress, 2009.

[2] J. C. J. M. Van den Bergh, "The GDP paradox, " *Journal of Economic Psychology*, vol. 30, pp. 117-135, 2009.

[3] M. P. Todaro, and S. C. Smith, *Economic Development*. 11th ed. Essex: Pearson Education, 2011.

[4] S. Basu, and P. Basu, "Regional Disparity in Australia: Analysis of Gender Development Index," *International Review of Business Research Paper*, vol. 1, no. 2, pp. 56-66, 2005.

[5] P. Nijkamp, and M. Abreu, (2009). *Regional Development Theory*. Submitted paper PN218MA-EOLSS. [Online]. Available : ftp://zappa.ubvu.vu.nl/20090029.pdf

[6] I. Majerova, "Least Developed Countries in Indexes of Human Development and Poverty," in *Proceedings of the 18th IBIMA Conference Innovation and Suistanable Economic Competitive Advantage: From Regional Development to World Economies.* Norristown: IBIMA, 2012, pp. 1210-1224.

[7] UNDP, *Human Development Report. Work for Human Development*. New York: UNDP, 2015.

[8] S. Anand, and A K. Sen, "Human development index: methodology and measurement," Occasional Paper 12. New York : Human Development Report Office, 1994.

[9] EUROSTAT, (2015). *Regional Database.* [Online]. Available: http://ec.europa.eu/eurostat/web/regions/data/database

[10] A. Sen, *Development as Freedom*. Oxford: Oxford University Publishing, 1999.

[11] A. S. Mulaik. "Blurring the Distinctions between Component Analysis and Common Factor Analysis," *Multivariate Behavioral Research*, vol. 25, iss. 1, pp. 53-59, 1990.

[12] J. Nevima, *The Competitiveness of Visegrad Group Regions*, Prague : Professional Publishing, 2014.

[13] J. W. Osborne and A. B. Costello, "Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis," *Pan-Pacific Management Review*, vol. 12, no. 2, pp. 131-146, 2009.

[14] P. R. McDonald, *Faktorová analýza a příbuzné metody v psychologii*, Praha: Academia, 1991.

[15] F. H. Kaiser, "The application of electronic computers to factor analysis," *Educational and Psychological Measurement*, vol. 20, pp. 141-151, 1960.

Ingrid Majerová, Dr. is working as an Assistant Professor at the Silesian University in Opava, School of Business Administration in Karvina, Czech Republic. She specializes in the area of economics, international economics and international development assistance. She is author of 2 monographs, 14 textbooks, published in various journals, such as Prague Economic Papers, WSEAS Transaction on Business and Economics and others, and presented her research at various international scientific conferences.

Jan Nevima, PhD is Associate Professor at the Silesian University in Opava, School of Business Administration in Karvina, Czech Republic. His research interests are econometric and multi-criteria analysis for assessment of regional and national competitiveness and convergence. He is the author (co-author) of 2 monographs, 1 textbook and several papers published in home and foreign journals and conference proceedings. He is actively involved in preparation and management of European Social Fund projects, home and foreign developing projects.