THE ASSESSMENT OF SOCIAL PROGRESS REGIONS OF SLOVAKIA

Dana JAŠKOVÁ

Abstract

Regional policy is characterized as a set of objectives, measures and decisions in the development activities of stakeholders. Regional development is defined as a system of economic, cultural and environmental processes. In the article is construction a composite index of social progress for the regions of Slovak Republic, the Social Progress Index. The Social Progress Index is an aggregate index number of social and environmental indicators that capture three dimensions of social progress: Basic Human Needs, Foundations of Wellbeing, and Opportunity. Each of this three dimensions including four components of the Social Progress Index. The main aim of the paper is to construction of the by Social Progress Index for condition of the Slovak regions. Input indicators included in the Social Progress Index are analysed using statistical methods. Internal data consistency within each component is verified by Principal Component Analysis. The normalized data are aggregated into a composite indicator and compared.

Key words

Regional policy, Social progress, Composite indicator, statistical methods

JEL Classification: C34, C52, R23

Introduction

Regional policy is characterized as a set of objectives, measures and decisions in the development activities of stakeholders (Habánik a kol, 2014). Regional development is defined as a system of economic, cultural and environmental processes. These processes take place in the region. Regional development contributes to its competitiveness, sustainable economic, social and territorial development. The region is defined as an administrative unit of the national and local levels (Cooke, Piccaluga, 2006).

In most studies, the starting inter-comparison is analysis of their economic level. The level is expressed by macroeconomic indicators GDP (Annoni, Kozovska, 2010). Looking for suitable measures of well-being to assess people´s quality of life is becoming more important on the agendas of governments and central institutes of statistics in several countries. An increasing number of programmes are being implemented in European countries. Since its introduction, GDP has been the most widely used indicator of country´s economic performance. However it is also highly criticized as a measure of people´s wellbeing. Indeed, GDP is measure of production, but it ignores the undesirable side effects, such as pollution, environment, which often accompany production growth. GDP does not include in its calculation a number of factors which significantly affect people´s quality of life. Included there are, for example, the quality of education, health care, environment, social relations, personal safety, decent housing (Ferrara, Nistico, 2015).

The European Union Regional Social Progress Index (EU-SPI) was a joint project of several European institutions. Index was constructed by Global Social Progress Index developed by the Social Progress Imperative. This organization is non-profit, non-governmental. The Global Social Progress Index has been published in 2014 and 2015 for over 130 countries in the world. The Social Progress Imperative defines concept social progress as the capacity of society to meet the basic human needs of its citizen. The definition further includes three broad elements of social progress: Basic Human needs Foundations of wellbeing and Opportunity. The EU-SPI provides a consistent and comparable measurement of the Regions of the EU social and environmental area. The EU-SPI is based on a different set of indicators but set of dimensions and components is the same.

1 Index construction methodology

The indicator can be considered as a special subset of the statistical results. A general definition of the concept, which would be applicable in all areas of official statistics, does not exist. There are several approaches to this definition. By the first approach the indicator is characterized as a combination of statistical results using a defined algorithm in the form of derived measurements. The second principle uses normative interpretation with the possibility of determining categories. The third principle involves mainly social statistics such as health, education, quality of work. In this sense, indicator includes something wider than is actually measured. The fourth approach is engaged in synthetic indicators. They are formed by combining of individual indicators, while using different methods "weighting" of each group.

The indicator is a statistical tool that monitors the nature and level of phenomena and processes monitor their development, changes and trends. This results in certain properties of the indicator:

- significant, relevant, understandable,
- transparent,
- analytical,
- complete,
- credible,
- internally comparable ,
- externally comparable ,
- intertemporal (Michálek, 2013).

The composite indicator is an indicator that is constructed from sub-indicators. The indicators are often presented in the different units have different levels and have different variability (Minařík, 2013). The EU-SPI is an aggregate index of fifty social and environmental indicators that capture three dimensions of social progress: Basic Human Needs, Foundations of Wellbeing, and Opportunity. Each of three dimensions including four components of the SPI:

Basic Human needs: Nutrition and Basic Medical Care, Water and Sanitation, Shelter, Personal Safety.

Foundation of Wellbeing: Access to Basic Knowledge, Access to Information and Communications, Health and Wellness, Ecosystem Sustainability

Opportunity: Personal Rights, Personal Freedom and Choice, Tolerance and Inclusion, Access to Advanced Education.

Each component is measured through several indicators. One of the main differences with other Wellbeing indexes is that the regional SPI includes only social and environmental indicators. SPI excludes regional GDP or income-based indicators. This is because the aim is in fact to express social progress directly. By excluding economic indicator, the SPI can systematically analyse the relationship between economic development (measured foe example by regional GDP) and social development.

Regional index EU-SPI allows regions comparing to any degree of economic levels. Helps regions with a lower SPI learn from regions with higher SPI. All components included in EU-SPI will identify significant differences, for example, of access to health care, quality of housing, personal safety, higher education, access to ICT, environmental pollution. In the design of EU-SPI have been used 50 indicators.

Data source are EUROSTAT, EU-SILC, European Environmental Agency, the Gallup World Poll, the Quality of Government Institute of the University of Gothenburg and Eurobarometer. Comparison was 272 EU regions. The study concluded alignment of EU regions at NUTS 2 level by EU-SPI values. Best Rated region was the region Övre Norrland in Sweden (1/272). Worst rated Yugoiztochen region was a region in Bulgaria. Of the 272 regions of the Slovak regions placed as follows: Region of Bratislava (181/272), West Slovakia (229/272), Central Slovakia (221/272) and East Slovakia (243/272).

Construction of EU-SPI consisted of the following steps:

1. Selection of observational units (assessing of the best possible geographical coverage given data availability and reliability).
2. Checking for statistical internal consistency within each component.
3. Normalizing.
4. Aggregating indicators.
5. Regional comparison score.
6. Testing scores and rankings through an extensive robustness analysis.

The achievement of main aim of this paper is basic research of regional differences in the social field of Slovak regions. The comparison is made using aggregate indicators the Slovak Regional Social Progress Index S-SPI. The process design is identical to the steps described above.

2. Construction of S-SPI

This section is designed summary indicator. The process of S-SPI (Slovak Social Progress Index) construction is in accordance with the methodology published in document “The EU Regional SPI: a measure of social progress in the EU regions, methodological paper” (Annoni, Dijkstra, 2013). The advantage of the summary indicator is a simple comparison of regions. The disadvantage is the different interpretations using different methods. Custom design composite indicator is described in several sub-sections and steps.

2.1 Selection of observational units

Those requirements must be respected in selection of appropriate indicators. The number of indicators should not be not even small (distorted the real situation) or too large (loss of clarity and transparency of interpretation). Indicators need to be regularly measured and officially published.

EU-SPI was constructed for all regions at the NUTS 2 level. This paper provides a description of the regional S-SPI calculated for all NUTS 3 regions in the Slovak Republic (eight self-governing regions): Bratislava, Trnava, Trenčín, Žilina, Nitra, Banská Bystrica, Prešov a Košice. Appropriate division of observational units is fairly debated issue. For observation of the RD, it is recommended to divide into district or functional regions. But for this kind of division, some important indicators are not available (Sloboda, 2006).

2.2 Selection of appropriate indicators

In the design of the S-SPI have been used 20 indicators. The composite indicator was calculated from the data for 2015. Selection of appropriate indicators was based on the official availability on NUTS 3 level. Data were retrieved from Slovak Statistical Office, Eurostat, Ministry of Interior, and Ministry of Education. It is clear that some of the relevant data were not available. Due to the mutual comparability of the data, they were divided by mid-year population in that region in 2015.

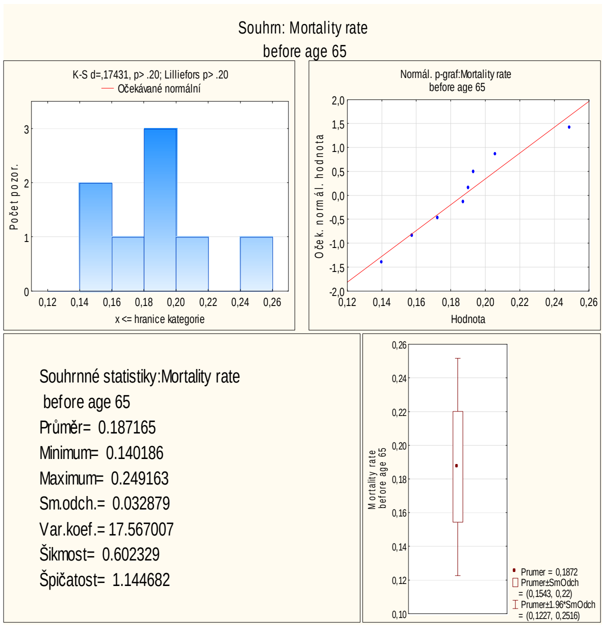
At level of self-governing regions, we encounter the problem of missing data in official databases. The literature on the analysis of missing data is extensive and rapid development more comprehensive methods can be found in Little, Rubin (2002). Whenever one or more indicators are observed at the country level only, an imputation method is adopted which imputes data by statistical estimates using available data. The formula for the calculation of the indicator values $y_{j,r}$ at NUTS 3 level, from y_{nat} at NUTS 1 level, is described in the next:

$$y_{j,r} = \frac{y_{nat}}{\frac{1}{k}\sum_{i=1}^{n}\frac{x_{inat}}{x_{i,j}^{reg}}},$$

where x_{inat} is the value of indicator x_i at the country level and $x_{i,j}^{reg}$ is value of x_i for region j.

For Basic human needs dimension we selected indicators: x_1 - Mortality rate before age 65 (MR), x_2 - Infant mortality (IM), x_3 - Beds in health facilities (HF), x_4 - Water supply from public water supply (WS), x_5 - Sewage treatment (ST), x_6 - Living area (LA), x_7 - Burdensome cost of housing (BCH), x_8 - Number of offenses (NO), x_9 - Number of forfeited (NFF), x_{10} - Homicide rate (HR) and x_{11} - Number of fires (NF). For dimension of Foundations of wellbeing: x_{12} - Secondary enrolment rate (SE), x_{13} - Number of Posts (NP), x_{14} - Internet at home (IH), x_{15} - Risk of poverty (RP), x_{16} - Life expectancy at birth (LEB), x_{17} - Environmental quality (EQ), x_{18} - Production of pollutant emissions (PPE). In the dimension of opportunity was selected following indicators: x_{19} - Gender gap (GG), x_{20} - Tertiary education attainment (TE). Given the direction, it went into the analysis of eleven positive indicators and neigh negative indicators.

Graph 1 Exploratory analysis



Source: own calculation

2.3 Components Internal consistency

The issue of aggregating indicator into a single composite indicator is an increasingly discussed topic. The aggregation process always implies the choice of weights or use aggregation method. Both issues play crucial role when assessing regional disparities.

Internal consistency is verified by classical multivariate method, Principal Component Analysis (PCA). PCA is useful statistical technique for finding patterns in data of high dimension. Using method is based on the properties of the correlation matrix of variables. Initial variables (manifest) will be replaced by a smaller number of new variables, called latent variables – the main components. This process consisted of several steps.

Step 1: Exploratory data analysis

Table 1 Descriptive statistics

Variable	Descriptive statistics (indicators in 2015)				
	Mean	Std. dev.	Coeff. of variance	Skewnes	Kurtosis
MR	0,19	0,03	17,57	0,60	1,14
IM	4,94	3,20	64,67	1,52	1,67
HF	7,87	1,37	17,46	-1,02	0,83
WS	88,60	5,01	5,66	-0,12	0,50
ST	12,46	4,43	35,57	0,26	0,07
LA	72,24	5,02	6,95	-0,08	-0,55
BCH	375,75	55,04	14,65	1,97	4,75
HC	1,38	0,63	45,62	1,52	2,47
NFF	10,09	2,63	26,06	0,20	0,11
NO	803,74	314,69	39,15	1,96	4,44
NF	200,25	39,06	19,51	0,91	0,70
SE	0,82	0,32	38,85	1,94	4,36
NP	30,66	6,02	19,64	-0,83	1,33
IH	77,91	3,61	4,63	-0,05	0,53
RP	11,98	3,63	30,30	-0,11	-1,65
LEB	73,25	1,06	1,44	1,19	0,21
EQ	107,46	90,12	83,87	2,26	5,28
PPE	0,70	0,27	38,15	0,13	-1,97
GG	0,80	0,06	7,76	1,21	0,66

Source: own calculation

Exploratory data analysis is a critical first step in analysing data from an experiment. The purpose of the analysis is to detect the presence of particularities between the data and verify the assumptions for further statistical processing. For this purpose, were calculated descriptive statistics (position, variability, and asymmetry). By graphical methods we have identified the presence of outliers, data independence (ACF), homogeneity (Box Plot) and normality (K-S, N-E, Lilliefors's test). The picture exploratory data analysis shows one of the outputs of analysis variable Mortality rate before age 65. The following table contains some statistics all tested variables.

The value of the coefficient of variation was used in subsequent analyzes as a decision criterion for the selection of appropriate indicators. The exploratory analysis shows that the data meet the required minimum prerequisites for further analysis.

The next step consists of variable´s transformation for some indicators, due to the value of the coefficient of skewness, where the absolute value of this coefficient was higher than 1.

Step 2: Correlation analysis

Table 2 Part of the correlation matrix

Proměnná	Mortality rate before age 65	Infant mortality	Beds in health facilities	Water supply from public water supply	Sewage treatment	Living area	Burdensome cost of housing	H
Mortality rate before age 65	1,00	-0,41	-0,04	0,22	0,12	-0,64	0,12	
Infant mortality	-0,41	1,00	0,12	-0,92	0,58	0,44	-0,65	
Beds in health facilities	-0,04	0,12	1,00	-0,18	0,39	-0,35	0,02	
Water supply from public water supply	0,22	-0,92	-0,18	1,00	-0,75	-0,38	0,86	
Sewage treatment	0,12	0,58	0,39	-0,75	1,00	-0,12	-0,74	
Living area	-0,64	0,44	-0,35	-0,38	-0,12	1,00	-0,46	
Burdensome cost of housing	0,12	-0,65	0,02	0,86	-0,74	-0,46	1,00	
Homicide rate	0,45	-0,64	0,03	0,79	-0,47	-0,63	0,87	
Number of forfeited	0,57	0,10	-0,29	-0,06	-0,19	-0,18	0,15	
Number of offenses	0,27	-0,35	-0,12	0,60	-0,68	-0,34	0,86	
Number of fires	0,07	0,69	0,04	0,60	0,08	0,24	0,23	

Source: own calculation

After one-dimensional analysis of variables we performed the correlation analysis. For the strong correlation between the indicators we considered while the correlation coefficient applied $|r| > 0{,}9$. These values have been diagnosed by the inverse correlation matrix and subsequent VIF-factor. The Variance Inflation Factor (VIF) measures the impact of collinearity among the variables in a regression model. If $|VIF| > 10$, multicollinearity is high. For further analysis we considered as a key indicator one who had the greatest variability (coefficient of variation) and seemed to be more appropriate for the description of interregional disparities.

The correlation analysis shows that from the structure of composite indicator should be removed five indicators: Water supply from public water supply, Burdensome cost of housing, Number of offenses, Secondary enrolment rate and Tertiary education attainment.

To further analysis it went fifteen indicators.

Step 3: Principal Components Analysis

Fifteen indicators were analysed by PCA. Its aim was to identify the key indicators and transform the original data to new latent variables. The suitability of selected indicators was statistically assessed by Kaiser-Meyer-Olkin´s criterion. Kaiser-Meyer-Olkin (KMO) test is a measure of how suited your data is for PCA. The test measures sampling adequacy for each variable in the model and for the complete model. The result of the PCA is shown in next several different outputs (tables and graphs).

Table 3 Eigenvalues

Order of Eigenv.	Eigenvalues of correlation matrix and related statistics (only active variables)			
	Eigenv.	% of total variance	Cumulative of eigenvalue	Cumulative % %
1	4,98	49,81	4,98	49,81
2	1,91	19,15	6,90	68,95
3	1,44	14,43	8,34	83,38
4	1,10	10,96	9,43	94,34
5	0,39	3,94	9,83	98,28
6	0,17	1,69	10,00	99,97
7	0,00	0,03	10,00	100,00

Source: own calculation

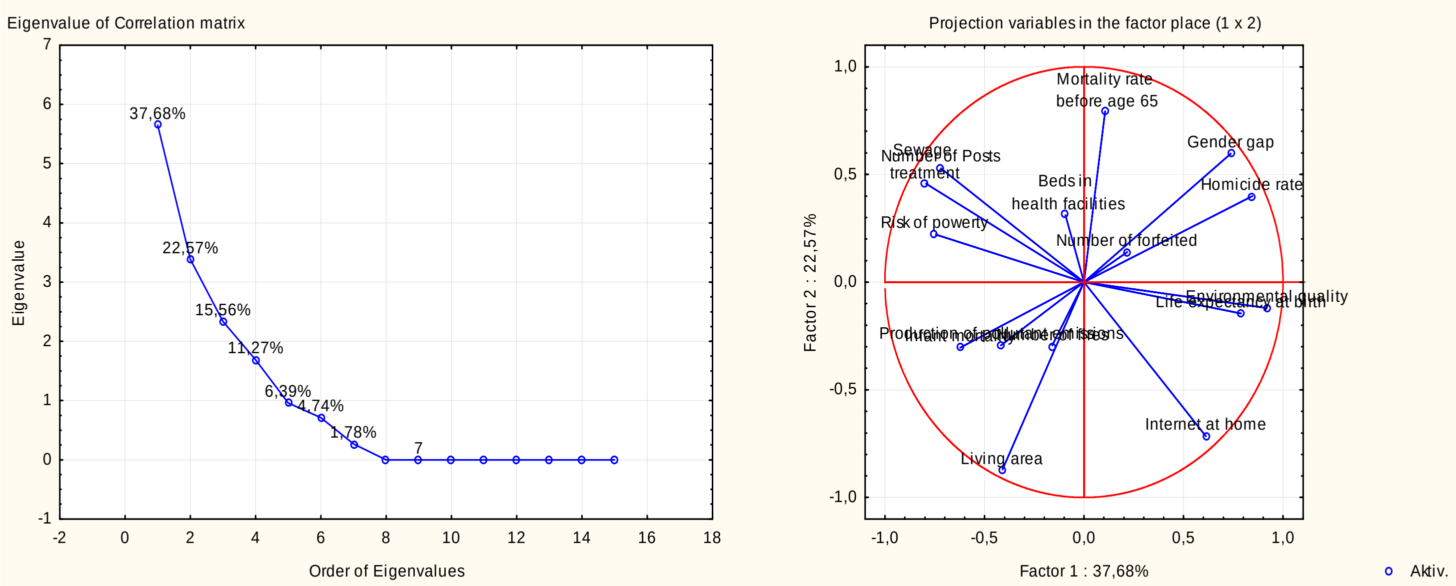
Since the covariance matrix is square, we can calculate the eigenvectors and eigenvalues for this matrix. These are rather important, as they tell us useful information about our data.

Table 4 Factor coordinate variables by correlations

Variable	Factor coordinate variables by correlations			
	Factor 1	Factor 2	Factor 3	Factor 4
Mortality rate before age 65	0,11	0,79	-0,39	-0,39
Infant mortality	-0,62	-0,30	-0,41	0,54
Beds in health facilities	-0,09	0,32	0,31	0,72
Sewage treatment	-0,80	0,46	0,01	0,33
Living area	-0,41	-0,87	-0,02	-0,19
Homicide rate	0,84	0,40	-0,02	-0,05
Number of forfeited	0,22	0,14	-0,94	-0,12
Number of fires	-0,16	-0,30	-0,81	0,41
Number of Posts	-0,72	0,53	0,12	-0,30
Internet at home	0,61	-0,71	-0,07	-0,17
Risk of powerty	-0,75	0,22	-0,22	-0,02
Life expectancy at birth	0,79	-0,15	0,39	0,40
Environmental quality	0,92	-0,12	-0,35	0,07
Production of pollutant em.	-0,42	-0,29	0,21	-0,20
Gender gap	0,74	0,60	-0,05	0,24

Source: own calculation

Graph 2 Outputs of the PCA



Source: own calculation

For further analysis we recommend to retain only those components that have their eigenvalue is greater than 1. Subsequently we selected the first fourth components that are explaining $94{,}34\ \%$ of the total variance. This stems from the Kaiser criteria. In another analyse they are preserved only those components that have modified a custom number greater than 1 (Meloun at al., 2012). Proper selection of components can be assessed according to the Cattell index chart (Scree Plot). It is a histogram of eigenvalues. From it can be to identify the main components. On the right is a chart, Plot components weights, of the first and second components. Each point on the graph is analyzed indicator. The chart compared the distance between the indicators. Indicators of the variability between regions that have a high correlation values are in the graph near the unit circle.

In terms of further reduction indicators and finding key indicators are further analyzed only those indicators that have a value of the correlation coefficient above 0.7 (Hrach, Mihola, 2006). The values of the correlation coefficients with the main components of the indicator (factor 1-4) are shown in Table 4. Omitted indicators are: Infant mortality and Production of pollutant emissions. Thirteen other indicators will be used later in the step on "Weighting and aggregation" to construct weights for the *SPI* composite indicator table.

Step 4: Weighting and aggregation

The results of PCA analysis allows to determine of q-th indicator in any time as:

$$w_i = |r_{i,j}|.var_j$$

where $r_{i,j}$ is value of correlation coefficient of the i-th indicator $(i = 1, \dots, 13)$ of the j-the component $(j = 1, \dots, 4)$. var_j is proportion of variability explained by j-th component. The values of weights that are assigned to each indicator are shown in Table 5

Table 5 The values of weights

indicator	MR	BHF	ST	LA	HR	NFF	NF	NP	IH	RP	LEB	EQ	GG
weight	0,30	0,08	0,30	0,20	0,32	0,15	0,13	0,27	0,16	0,28	0,30	0,35	0,28

Source: own calculation

For PCA analysis it shows that the highest weight is assigned to the indicator Environmental quality and lowest weight indicator Beds in health care.

Normalisation of data is required prior to any data aggregation as the indicators in a data set often have different measurement units. The method Min-Max was used. This method normalises indicators to have an identical range $\langle 0; 1 \rangle$ by subtracting the minimum (maximum) value and dividing by the range of the indicator values. If $x_{i,r}$ is positively oriented we followed the formula:

$$I_{i,r} = \frac{x_{i,r} - min_r(x_i)}{max_r(x_i) - min_r(x_i)},$$

In the case of negative force of $x_{i,r}$ the normalization is realized through the formula:

$$I_{i,r} = \frac{max_r(x_i) - x_{i,r}}{max_r(x_i) - min_r(x_i)}.$$

Table 6 Regional Slovak SPI

Region (r)	SPI_r	Order
Trenčín	1,273115	1
Bratislava	1,104446	2
Prešov	1,079353	3
Trnava	1,056612	4
Žilina	1,035773	5
Nitra	0,894041	6
Banská Bystrica	0,796455	7
Košice	0,760206	8

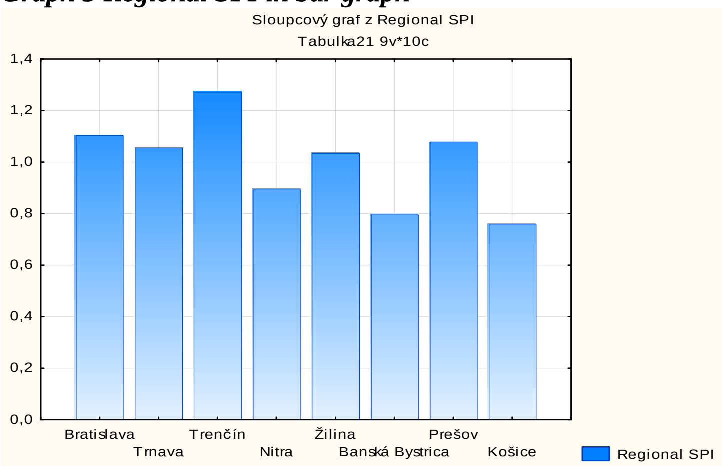
Source: own calculation

For calculation of SPI was used the Additive aggregation method. The Composite indicator SPI for each region SPI_r was calculated according formula:

$$SPI_r = \frac{\sum_{i=1}^{n} I_{i,r} w_i}{\frac{\sum_i^n \sum_r^m I_r}{m}},$$

where w_i is weight of q-th indicator. If SPI_r = 1, the region is assesed as an average. SPI_r> 1 means the above average appreciation of region, SPI_r< 1 means, that region is evaluated as a below average.

Graph 3 Regional SPI in bar graph



Source: own calculation

The resulting values of the regional SPI_r were as follows table 6.

Best Rated region was Trenčín. SPI_r reached around 0,17 better than Bratislava region. It's interesting because the assessments which take into account economic indicators Bratislava region dominates. These two regions as only amounted SPI_r greater than 1. They can therefore be considered as regions with above-average SPI_r. Indicators of quality of life in these regions are excellent. The second group may include regions of Prešov, Trnava and Žilina. SPI_r values are closer to one. These regions can be considered in terms of SPI_r for average. The last third group consists of regions of Nitra, Banská Bystrica and Košice. In these cases, the SPI_r less than 1 and therefore consider them in terms of SPI_r as below average.

SPI_r index values for individual comparisons are shown in bar graph third. The practice is often determined by regression relationship between SPI and regional GDP. Official figures for 2015 are not yet published.

Conclusion

The paper applied approach comparing regions of the EU in 2015. The European Union Social Progress Index (EU-SPI) builds on feedback from public and experts in the field like. According to the methodological procedure it is designed composite indicator SPI_r. When its construction was the use of advanced multivariate statistical methods. From the database were selected twenty key indicators at NUTS 3. The source was the official website of the Slovak Government. Using correlation analysis has identified key indicators. Further reduction of dimensions has been using PCA analysis. Of the fifteen indicators using weighted aggregation method min-max was designed composite indicator SPI_r. Using SPI_r, the individual regions ranked.

Best Rated region in the field of social progress in 2015 was Trenčín region. Bratislava region ended up in second place. It is interesting. When compared regions in terms of economic progress (Regional Competitiveness Index - RCI) clearly dominates the Bratislava region. The other three regions, Prešov, Trnava and Žilina region were rated similarly. He was sixth in Nitra region. The worst were evaluated regions of Banská Bystrica and Košice.

There are many positive but also negative reviews to assess the social progress of the region using a single value. Benefits include a summary of multidimensional data simple and comparability. Described methodological approach could in the future serve for comparison evolution of social progress in the regions of Slovakia.

Acknowledgements

Tento príspevok vznikol v rámci realizácie projektu VEGA 1/0953/16 Hodnotenie miery vplyvu klastrov na rozvoj regiónov Slovenskej republiky.

References

Annoni, P., Kozovska, K. (2010). EU Regional Competitiveness Index 2010. *JRC scientific and technical reports*. Luxembourg.

Annoni P., Dijkstra, L. (2013). The EU Regional SPI: a measure of social progress in the EU Regions. Methodological paper. *Economic Analysis Unit of the EC-DG for regional and Urban Policy.*

Cooke, P., Piccaluga, A. (2006). Regional development in the Knowledge Economy. *Routledge, Taylor & Francis Group.*Great Britain.

Ferrara, AR., Nistico, R. (2015). Regional well-being indicators and dispersion from a multidimensional perspective: evidence from Italy. *Annals of regional science.*Springer55 (2-3),373-420.

Habánik, J. a kol. (2014). *Regionálna ekonomika a regionálny rozvoj*. Trenčín: FSEV TnUAD.

Haviernikovǻ, K. (2016). The basic conditions for clusters creation in the Trenčín region. *In: Ukraina - Višegradska četvirka (V4) - es mechanizm mižnarodnogo spivrobitnictva:* Zbirnik naukovich prac. - Luck : BIEM, 30-42.

Little, R.J.A., Rubin, D. B., (2002). Statistical Analysis with Missing Data. *Wiley Interscience*. J. Wiley & Sons, Hoboken, new Jersey.

Meloun, M., Militký, J., Hill, M. (2012). Statistická analýza vícerozměrných dat v příkladech. *Academia, Praha* Michálek, A. (2013). *Vybrané metódy merania regionálnych disparít.* [online] [cit. 2017-7-1].. Available at: http: https://www.sav.sk/journals/uploads/12121204Michalek.pdf

Minařík, B., Borůvková, J., Vystrčil, M. (2013). *Analýzy v regionálním rozvoji*. Příbram: Professional Publishing.

Sloboda, D. (2006). *Slovensko a regionálne rozdiely*. [online] [cit. 2017-277-1]. Available at: http: http://www3.ekf.tuke.sk/re/Disparity%2520a%2520perifernost/Regionalne%2520disparity/Slovens...

Vojtovic, S., Krajnakova, E. (2013). Trends in Economic Growth and Unemployment in Slovakia. *Internatiomal Conference on Education, Management and Social Science (ICEMSS)*. China, 189-191.

Štatistický úrad Slovenskej republiky (2017). *Databázy* [online] [cit. 2017-03-27]. Available at: http://datacube.statistics.sk/TM1WebSK/TM1WebLogin.aspx

Contact

Dana Jašková, PhD., RNDr.
Alexander Dubček University of Trenčín
Faculty of Social and Economic Relations
Department of Economy and Economics
Študentská 3
911 50 Trenčín
Slovak Republic
e-mail: dana.jaskovanuni.sk