

Regional Sustainable Development – Czech Region Ranking Using Multi-criteria Decision Analysis

Lenka Hudrlikova¹, Ludmila Petkovova¹

ABSTRACT:

The aim of the paper is to provide a ranking of the Czech NUTS 3 regions based on sustainable development indicators. The original list of indicators was published by the Czech Statistical Office in 2008 and reviewed in 2010. In the analysis the same set of indicators with the latest data was used. The indicators in each pillar are merged by means of linear aggregation with weights derived from the principal component analysis. Because three pillars of sustainable development (environmental, economic and social) are assumed to be non-compensable, the multiple-criteria decision analysis is applied on a pillar level in the final composite indicator. Both two main approaches – Borda and Condorcet were considered. Since the Borda approach leads to the compensability of the indicators, the Condorcet approach was in the spotlight. Advanced rules and adjustment for Condorcet approach were employed. Advantages and disadvantages of the methods are provided. As a result more final rankings exist. The deep discussion about the results is provided. The special attention is paid to the capital city Prague, border regions, and industrial regions. In addition, the correlation between final ranking and other indicators is tested.

1. Introduction

Although the total area of the Czech Republic is small, the development in its regions differs. It is determined by different environment and history. In the past decades the Czech Republic has been transformed, in this context the development on region level attracted attention. The environmental issues arose and so the sustainable development has been in the spotlight. Regional sustainable development is a very actual issue for Czech policymakers; however it is startling that there is no multidimensional comparison across regions based on indicators of sustainability. Some regional comparisons are available, but mostly in the form of quality of life assessments (Mederly, Topercer, & Nováček, 2004) or statistical overviews without a deeper analysis and regional comparison (Czech_Statistical_Office, 2010). Indeed, measuring of sustainable development is questionable due to the vague definition of sustainable development indicators. If measuring is poor, benchmarking or ranking can be a useful tool for assessment of the regions. The question left is a chosen method. The

| ¹ University of Economics, Prague

sustainable development is characterised by deep complexity (Funtowicz, Munda, & Paruccini, 1990). The main assumption is that sustainability assessment needs a set of multidimensional indicators (Munda, 2005). However, for benchmarking or ranking those indicators have to be aggregated into one number. The well-known concept of three pillars (economic, social and environmental of sustainability development) was followed. The equal importance (or weight) was assigned to each pillar. In pursuit of final ranking, the ranking of separate pillars was assessed by means of multi-criteria decision analysis. The aim of the paper is provide comparison across regions because it shows which regions are on a very low level in sustainable development. It gives an indication of the need for an improvement. Ranking facilitates an analysis of the strengths and weaknesses of each region. Also the comparisons in separate pillars pinpoint them. Last (but) not least advantage is that ranking is easily understandable for policymakers, public and other stakeholders.

2. Data

The possibilities of statistical data processing at the local level are rather problematic due to a small number of official – no matter how thoroughly conducted – surveys. In spite of the limited amount of data and different data requirements in comparison with the national level, questionnaire surveys are often launched to find out about public opinion on particular issues. But it is difficult to make spatial and time comparison of the country regions. Only local surveys of a limited scope are carried out. For meaningful statistical analyses, it is necessary to have a high-quality initial data set at one's disposal. In order to ensure a better comparability it is useful to select such indicators that are broadly available in regions not only for single country, but eventually for other countries as well. Only on NUTS 3 level official data sources provide reliable data. This does not allow us to study sustainable development in the Czech Republic at a lower regional level in detail, which would be more suitable for the identification of regional disparities in some sustainable development aspects. A greater number of observations would be also more appropriate so that statistical methods could be employed for further analyses. For the above mentioned reasons the level of NUTS 3 territorial administration unit was chosen which means analysis is performed on 14 Czech regions (for the list see Table 1). The original set of sustainable development indicators was created by the Czech Statistical Office in 2008 and reviewed in 2010. Since then, no more reviews were carried out.

In this analysis the latest available data were used. Some adjustments had to be made with respect to data availability and the scope of this analysis. Final set of indicators after all changes is listed in Appendix. All the indicators are

divided into three sustainable development pillars – economic (12 indicators), social (12) and environmental (12). Full list with detailed clarification of the changes is available in (Fischer, Petkovová, Helman, Kramulová, & Zeman, 2013).

3. Methodology

In order to assess the sustainable development in Czech regions, two steps of aggregation of indicators were carried out. The first step consists in aggregation on a pillar level which means separately for each pillar (economic, environmental social) the regions ranking was performed. The indicators within the pillar were compiled by means of weighted arithmetic mean. This set of indicators was created by the Czech Statistical Office from the point of view of definition of sustainable development. No statistical measures or analysis were provided. But if we want to use it for ranking, at least scrutiny of correlations between indicators is essential. However there is no general consensus on correlation in the case of composite indicators (Saltelli, 2012). On the one hand, high correlation among indicators could be seen as a problem which should be corrected by creating appropriate weights. Weight should be set inversely proportional to the strength of the overall correlation for a given indicator. On the other hand, high correlation should not be corrected because it is a feature of the problem. Highly correlated indicators could indicate non-compensable different dimensions. The correlation implies indicators measuring the same phenomenon, but there should be checked whether the correlation is due to redundancy of the information (double counting of the dimension). Within the pillar correlation is very strong. Weights derived from the Principal Component Analysis (hereunder PCA) aim at correcting for correlation (i.e. two highly correlated indicators will weighted lower). Because we used the data for Czech regions where Prague (the capital city) is a significant outlier, robust PCA was used instead of generally known PCA. Outliers can seriously distort the estimated means and covariances and robust estimation techniques that decrease or completely remove the influence of observations that are outliers. The EM algorithm uses weights that are inversely proportional to how “outlying” the observation is (NCSS, 2007). The usual estimates of the means and covariances are modified to use these weights. The process is iterated until it converges. Weights depend on eigenvalues and on the optimal numbers of components. In accordance to Kaiser Criterion, which suggests choosing all components which are associated to an eigenvalues higher than one, five components were selected in economic and environmental pillar (and four components in social pillar respectively). Weights are normalized by squared factor loading, which is the share of the variance of the factor explained by the variable. Finally weights are

scaled to unity sum. In practice linear aggregation (LIN) is the most widespread. The simplest method is weighted average:

$$CI_c = \sum_{q=1}^Q I_{qc} \cdot w_q, \quad (1)$$

subject to

$$\sum_q w_q = 1 \text{ and } 0 \leq w_q \leq 1,$$

where I_{qc} is normalized indicator q ($q=1, \dots, Q$) for region c ($c=1, \dots, M$) and w_q weight for indicator q ($q=1, \dots, Q$). Note that prior to aggregation the normalization is needed (in this case z-scores). Linear aggregation implies full compensability. Poor performance in some indicator can be compensated by sufficiently high values of others indicators. In this case trade off between indicators is not undesirable. Within one pillar, the indicators can be compensable because they can measure similar phenomenon. Using these weights and normalized indicators, score (and ranking) for each regions in each pillar separately is available.

In the second step, the final ranking by means of multi-criteria decision analysis is provided. There are two main approaches: (i) Borda and (ii) Condorcet. The first one is considered to be compensatory, the latter is non-compensable (OECD/JRC-European Commission, 2008). Compensability, which means trade-offs between indicators or pillars, is a crucial issue. When weights should express importance of the coefficient, Condorcet approach should be chosen. While Borda scoring is applicable when weights are meaningful in the form of trade-offs (Munda, 2010). These two approaches are main streams but there are many adjustments of these methods with respect to used rules. Borda rule assigns no point if the alternative is ranked last and one point if it is ranked next to the last. Given N alternatives, the process continuing like this up to $N-1$ points awarded to the alternative rank first. The winner is the region with the highest total score. Borda approach uses a frequency matrix. It is based on points expressing how many times the alternatives appear as 1st, 2nd, ..., n -th place. It means Borda method is based on ordinal information and therefore interval level information is lost. But the advantage is that impact of outliers is eliminated. Weights between the three pillars are assumed to be equal.

On contrast, Condorcet approach is based on outranking matrix where for each pair of regions a concordance index is computed by counting how many individual indicators are in favour of each region. Because of transitivity, only pairs whose concordance index is higher than 50% of criteria are selected. In this way final ranking is obtained. Because this approach is based on pair wise comparisons between all alternatives considered, it might be very computational exacting. There are several rules which have been developed to overcome the computational problems (Munda & Nardo, 2009). The results for Copeland rule

are presented in the section of results. Another drawback of this method is possible indecisiveness. The Copeland rule assures of consistency. Copeland rule is very simple. Outranking matrix is transformed into matrix with values 0, 1, -1. Basically if alternative $A > B$ positive one point is assigned, if $B > A$ negative one point is assigned and zero in the case of tie. The Copeland score for a given is sum in a corresponding row.

Other properties are the same as in Borda, i.e. used ordinal information and interval level information is lost, independent to outliers. Note that mathematical formulation, axioms, theorems etc. can be found in (Munda, 2010).

4. Results

The results of rankings are showed in Table 1. First three columns show ranking in every particular pillar. The strong/weak performance in individual pillar can be observed. Prague (the capital city) is a centre for business, government, various institutions, education. There is diversified economy with the creation of skilled jobs. It follows that Prague ranks first in economic and social pillars. However its performance in environmental pillar is the worst. Actually this region is only one big city agglomeration without much green environment.

Table 1. Ranking in each pillar and final rankings

	Economic pillar	Social pillar	Environmental pillar	Borda		Condorcet	
				score	rank	score	rank
Prague	1	1	14	26	3	13	1
Central Bohemia	3	7	13	19	7	1	7
South Bohemia	11	5	7	19	7	1	7
Plzen	9	2	3	28	1	7	3
Karlovy Vary	13	12	2	15	9	-7	10
Usti	12	14	4	12	14	-11	14
Liberec	6	9	1	26	3	7	3
Hradec Kralove	8	4	8	22	5	3	5
Pardubice	10	8	10	14	12	-5	9
Vysocina	14	3	11	14	12	-9	13
South Moravia	2	6	6	28	1	9	2
Olomouc	4	11	12	15	9	-7	10
Zlin	7	10	5	20	6	3	5
Moravia-Silesia	5	13	9	15	9	-7	10

Source: authors' calculations

By applying multi-criteria decision analyses, the Borda and Condorcet final ranking were performed. Note that score in table 1 are not comparable. Only ranks can be compared. Maximum variation between Borda and Copeland rankings amounts to 3 places (the case of Pardubice region).

Plzen region and South Moravian region achieved the first place in Borda ranking. But in Condorcet ranking Prague is the leader. To compare the methods one can see that the top three regions are the same nevertheless their ranking is reversed. These top three are regions with the large cities Prague, Brno, Plzen (Prague - the capital city, South Moravian region, Plzen region). Universities are situated there and therefore high number of university graduates is available. Their economies are diversified with the potential creation of skilled jobs. The density of roads, highways and railway tracks is high. Plzen region is known for traditional machinery industry (e.g. car industry Skoda). South Moravian region is area with significant development of academic and research infrastructure. The region is well-known for production of wine. In general the agriculture in this region is on very high level. These favourable conditions together with very good environment provided them leading positions in regional comparison.

In both rankings, Usti region was identified as a laggard. This region struggles with environmental problems because energetic (especially coal industry), chemical and some other heavy industries are located there. This region faced also economic and social problems (high unemployment, low regional GDP) since the restructuralization of heavy industry in the Czech Republic. Usti region is labelled structurally affected. This region suffers from outflow of skilled labour force because of lack of job opportunity. The young people leave the region because they study at universities in different cities (Prague, Brno, Olomouc,..). After graduation they stay in the city of their studies where they can easily find a job. It causes backwardness in economic pillar and also a poor performance in social pillar. The similar situation is in Karlovy Vary region and Moravian-Silesian region. Both regions struggle with high unemployment rate and low level of education. No university is situated in Karlovy Vary region. Karlovy Vary region and Usti region were marked by expulsion of the part of the population after the Second World War. The original way of lifehood was ruined and the mining and heavy industry was concentrated there. But at the end of the past century, majority of manufacturing plants and works were closed because of ineffectiveness. The heavy industry and mining in these regions caused damages to environment and air pollution. It influenced negatively their performance in every pillar.

5. Discussion

In order to verify the rankings, correlation between chosen representatives of each pillar and rankings was measured by means of Spearman rank correlation coefficient (see Table 2).

Table 2. Correlation matrix

	Economic	Social	Environmental	Borda	Condorcet	GDP per capita	Unemployment rate	Emission of pollutants CO
Economic	1,00	0,13	-0,41	0,57	0,62	0,51	0,20	-0,04
Social	0,13	1,00	-0,32	0,55	0,59	0,59	0,83	0,05
Environmental	-0,41	-0,32	1,00	0,21	0,07	-0,37	-0,30	0,22
Borda	0,57	0,55	0,21	1,00	0,94	0,53	0,53	0,17
Condorcet	0,62	0,59	0,07	0,94	1,00	0,58	0,63	0,30
GDP per capita	-0,51	-0,59	0,37	-0,53	-0,58	1,00	-0,60	0,31
Unemployment rate	0,20	0,83	-0,30	0,53	0,63	0,60	1,00	0,15
Emission of pollutants CO	-0,20	-0,02	0,31	0,07	0,18	-0,47	0,08	1,00

Source: authors' calculations

First three rows (or columns) show the correlation between separate pillar and the other indicators. It is no surprise that correlation between environmental and economic pillars is negative. The correlation coefficient between Borda and Condorcet ranks accounts for 94% which signal very strong correlation. Not surprisingly indicator of GDP per capita is positively correlated with economic and social pillars and negatively with environmental pillars. Correlation between unemployment rate and social pillar is very strong. The selection of indicator of emission of carbon monoxide CO might not be the great representative of environmental area because its correlation coefficients are low. In the Czech republic the environmental statistics was poor in the past. The environmental indicators were not regularly monitored, especially data on regional level were not available. However in the situation is getting better.

Conclusion

Two main approaches from the group of multi-criteria decision analysis were compared. The final rankings are similar, and the correlation is very high. Borda and Condorcet approaches are based on different ideas and in other examples they can produce completely different results. Therefore the selection of the ranking rule matters. Nevertheless the similar rankings enable to draw a conclusion about the sustainable development in Czech Republic. The importance of economic prosperity was revealed. In the Czech Republic the economic growth is linked with social welfare. It was also proved by correlation. As already mentioned, top three regions are excellent in economic pillar which is rewarded also in social pillar. Note the pillars were weighted equally, therefore the environmental pillar might be disadvantaged. The evidence is a correlation matrix. As an example it is noticeable for the capital city Prague.

Even though Prague is on the last place in environmental pillar, top performance in economic pillar and subsequently social area lead to the first and the third place respectively. To conclude, the drawbacks and advantages of each region were taken into account in ranking. Sustainable development target positive progress in all three pillars simultaneously. The progress of one area must not be redeemed by deterioration of the others.

The used rule can have significant effects on the final output. Each method has some pros and cons. Therefore it is up to the constructor of ranking select to the most suitable method and clarifies the motivation.

Acknowledgement

This paper has been prepared under the support of the University of Economics, Prague - project No. 11/2012 “Construction and verification of sustainable development indicators in the Czech Republic and its regions”.

Appendix**Final set of indicators (source: Czech Statistical Office, 2010, authors' adaption)****Economic pillar**

1. Change in Gross Domestic Product (Development of GDP in constant prices)
2. Labour Productivity (Development of GDP per 1 employed)
3. Local Government Deficit/Surplus
4. Gross Value Added in Services (Share of the Tertiary Sector in Gross Value Added in %)
5. Investment Rate in %
6. Net Disposable Income of Households per inhabitant in thousands of CZK
7. Small and Medium-sized Enterprises (Share of Small and Medium-sized Enterprises in the Total Employment in %)
8. Transport Infrastructure – Density of the Motorway Network in km per 100 km²
9. Transport Infrastructure – Railway Lines Density in km per 100 km²
10. Freight Transport (Excluding Transit, including Road, Rail and Water Transport per thousand of CZK GDP, in kg)
11. Passenger Transport (within the Region by Public Road and Rail Transport per Capita)
12. Research & Development Expenditures to GDP in %

Social pillar

13. Households with Net Income below Subsistence Minimum
14. General Unemployment Rate in % (Aged 15+)
15. Employment of Elderly Workers (Employment Rate of People Aged 55-64 in %)
16. Employment of Women in %
17. Mortality (Standardized Mortality Rate - Number of Deaths per 1000 mid-year Population)
18. Life Expectancy (of men at birth in years)
19. Life Expectancy (of women at birth in years)
20. Highest Level of Education Attained (Share of the Population with Tertiary Education in the Population Aged 15 and Over in %)
21. Internet Access (Share of Households connected to Internet in %)
22. Local Government Expenditures on Culture per inhabitant in CZK
23. Civil Society – Political Participation (Turnout in Elections to Municipal Councils in %)
24. Civil Society – Civil Participation (Mid-year Population to Non-profit Organization)

Environmental pillar

25. Arable Land in %
26. Consumption of Industrial Fertilizers in Pure Nutrients in kg/ha of Arable Land
27. Coefficient of Ecological Stability
28. Share of organically farmed land in the total area of agricultural land in %
29. Share of Broadleaved Species in %
30. Areas with Deteriorated Air Quality in %
31. Nitrogen Oxide Emissions (REZZO 1-4) in tonne per km²
32. Sulphur Dioxide Emissions (REZZO 1-3) in tonne per km²
33. Waste Generated by Enterprises in kg per thousand CZK of GDP
34. Municipal Waste Generated in kg per inhabitant
35. Acquired Investment Expenditures on Environment Protection according to Location of Investment in CZK per inhabitant

36. Non-investment Expenditures on Environment Protection according to Region of Residence of the Investor per million CZK of Regional GDP

References

- Czech_Statistical_Office. (2010). *Vybrané oblasti udržitelného rozvoje v krajích České republiky 2010*. Praha: Český statistický úřad.
- Fischer, J., Petkovová, L., Helman, K., Kramulová, J., & Zeman, J. (2013). Sustainable development indicators at the regional level in the Czech Republic. *Statistika*, 50(1).
- Funtowicz, S., Munda, G., & Paruccini, M. (1990). The aggregation of environmental data using multicriteria methods. *Environmetrics*, 1(4), 353-368.
- Mederly, P., Topercer, J., & Nováček, P. (2004). Indikátory kvality života a udržitelného rozvoje - kvantitativní, vícerozměrný a variantní přístup.
- Munda, G. (2005). "Measuring Sustainability": A Multi-Criterion Framework. *Environment, Development and Sustainability*, 7(1), 117-134.
- Munda, G. (2010). *Social Multi-Criteria Evaluation for a Sustainable Economy*: Springer.
- Munda, G., & Nardo, M. (2009). Noncompensatory/nonlinear composite indicators for ranking countries: a defensible setting. *Applied Economics*, 41(12), 1513-1523.
- NCSS. (2007). User's Guide IV
Multivariate Analysis, Clustering, MetaAnalysis, Forecasting / Time Series, Operations Research, Mass Appraisal
- OECD/JRC-European Commission. (2008). Handbook on Constructing Composite Indicators. *Methodology and User Guide*, Brussels: OECD, 148.
- Saltelli, A. (2012). *Composite Indicators: An introduction*. Paper presented at the 10th JRC Annual Seminar on Composite Indicators.