Warsaw Forum of Economic Sociology 10:2(20) Autumn 2019
Warsaw School of Economics; Collegium of Socio-Economics; Institute of Philosophy, Sociology and Economic Sociology

Comparison of Healthcare Performance and its Determinants in European Countries Using the TOPSIS Approach

Agata Sielska*

Abstract

The paper provides a comparison of European countries with respect to their healthcare. Two dimensions were taken into account. The first set of criteria contains the resources provided by the healthcare sector to the inhabitants, such as the number of practising physicians and the number of hospital beds per one hundred thousand inhabitants, and the resources provided to the sector by financing entities, such as health care expenditures. The second dimension reflects the output sphere which results from the quality of healthcare system performance; this set of criteria includes life expectancy, the share of people with good or very good perceived health, and the infant mortality rate. Countries were ranked using the TOPSIS method. Results reveal that resources are crucial to the assessment.

Keywords: healthcare, ranking, multicriteria methods, TOPSIS

JEL Classification: I10, I14, I19, C44

Introduction

The organization of healthcare systems, its resources and performance vary between European countries. The provision of best possible healthcare for inhabitants is still a valid topic. The main goal of the paper is to assess the performance of healthcare in European countries based on both resources used in the sector and health status of the population. The paper consists of four parts. In the first one the background for the study is presented. In the second part we discuss the possibility to use multiple criteria decision making methods for multidimensional analysis,

* Warsaw School of Economics (SGH), asiels@sgh.waw.pl, ORCID 0000-0001-6880-8899.

while in the third part one of those methods, the TOPSIS is described. Further, data and criteria weights are presented. Finally, the last part provides the discussion of the results.

Background

Health is one of the most important factors taken into account in the quality of life and well-being. Governments' spending on healthcare sector is constantly rising and while it becomes a burden to the society it opens new possibilities to improve quality of services and efficiency. According to Slottje it had been assumed, that 'basic needs such as health and education would be taken care of as a by-product of growth in GNP' (Slottje 1991: 684), however later infant mortality and life expectancy were included the quality of life assessment. Dasgupta (2000) also uses life expectancy and refers to the 'World Development Report' published in 1995 by the World Bank in which i.a. life expectancy at birth and infant survival rate were taken into account. Žmuk (2015) applies Health Care Index which is calculated based on the characteristics of the health system such as health care equipment and staff. In consequence the index is considered to reflect the overall quality of the health care system in a given country (Numbeo, n.d.).

Some studies on regional quality of life and living standards also include criteria connected to the healthcare such as: outpatient healthcare facilities per 10,000 population, doctors per 1,000 population and the number of inhabitants per 1 public pharmacy (Przybyła 2015) or budget health protection expenditure per capita and stationary social welfare facilities per 1,000 population (Majka 2015). Dąbrowa (2011) provides additional examples of inclusion of healthcare dimension based on the literature study. One of quite recent big studies was conducted by van Zanden et al. (van Zanden et al. 2014) who measured health status for the period from 1820s to 2000s. Their research in this field was based on life expectancy and height – i.e. the data that were available (or estimated) for the analysed period and also reflected the bodily health and nutrition.

The health situation of a given population depends on various resources. Hospital equipment and access to professional medical staff depend in turn on expenditures and the higher education system, as well as general socio-economic conditions. EU countries differ in the way the healthcare systems are financed and organised. The question as to whether the system influences the performance of the healthcare model is still valid (Pelone et al. 2008). In the literature, many examples of comparing countries based on healthcare sectors and related features can be found. Jaworzyńska (2016) presents a comparison of the health system financing model in Poland and other countries. Many comparisons of healthcare and its overall or specific performance that are based on the quantitative data use cluster analysis (Nixon 2000; Miszczyńska 2013; Girginer 2013, Pezer 2018; Walczak et al., 2018; Yilmaz Işikhan, Güleç 2018). Data Envelopment Analysis (DEA) is one of the methods often used for measuring the efficiency of the healthcare and comparing performance among countries (Grausová, Hužvár, Štrangfeldová 2014; Kujawska 2015; 2018; Kriksciuniene, Sakalauskas 2017).

General comparisons are based upon following criteria: number of beds (Nixon, 2000; Miszczyńska 2013; Asandului et al. 2014; Grausová, Hužvár, Štrangfeldová, 2014; Önen, Sayın 2018); number of doctors (Nixon 2000; Miszczyńska 2013; Asandului et al. 2014; Grausová, Hužvár, Štrangfeldová 2014; Kujawska 2018; Önen, Sayın 2018); healthcare expenditure as a percentage of GDP or per capita (Nixon 2000; Miszczyńska 2013; Asandului et al. 2014; Grausová, Hužvár, Štrangfeldová 2014); perceived health status (Miszczyńska 2013); life expectancy at birth (Nixon 2000; Miszczyńska 2013; Asandului et al. 2014; Önen, Sayın 2018); infant mortality rate (Nixon 2000; Asandului et al. 2014; Grausová, Hužvár, Štrangfeldová 2014) and *health adjusted life expectancy* (Asandului et al. 2014).

Apart from overall comparisons, some specified fields are also analysed and compared among countries. Pezer (2018), for example, compares maternity support policies in the EU, while Kringos et al. (2013) focus on primary care, based on 77 indicators.

Multiple Criteria Decision Making Methods for Multidimensional Analysis

If objects are to be ranked or compared based on multiple criteria, multidimensional methods are usually employed. However, methods originating from MCDM (*Multiple Criteria Decision Making*) can be used as well, as shown in a number of papers. These methods are dedicated to, among other things, rank objects (further referred to as 'alternatives', in accordance with MCDM terminology), according to a decision-maker's preferences. In cases where a single, specified decision-maker does not exist, weights used in the analysis can be derived based on criteria values. The general rule is that the most differentiated characteristics among the alternatives should play the most important role in the comparison, as they convey the most information.

Similarly, criteria which are strongly correlated should play a less important role, in order to avoid redundant information. There are a number of weight elicitation methods. Weights can either be assigned to each criterion based on its statistics (Diakoulaki et al. 1995), or based only on their relative importance, i.e. the ranking of criteria can be constructed and specific weights can be elicited further by using different approaches like Rank Order Centroid (Roszkowska 2013).

Although MCDM methods were created for another purpose, they are used for multidimensional analysis in various fields. Roszkowska and Filipowicz-Chomko (2019) indicate the usefulness of MCDM methods as evaluation tools. In the literature, other examples of such use of MCDM techniques can be found. In some of those, the weighting is omitted (Baer-Nawrocka, Markiewicz 2010; Ertman 2011); in such cases it is assumed that all the weights were set equally. Panagiota Digkoglou and Papathanasiou (2018) use the PROMETHEE (*Preference ranking organization method for enrichment evaluation*) method to construct a ranking of the EU countries according to anenvironmental performance index. The same method is used by Zhang and Smelev (2019) to evaluate the sustainability of cities. Ardielli (2016) applies 3 different MCDM ranking methods for evaluation of eGovernment development. Murgante et al. (2017) applied the ELECTRE (*ELimination Et Choix Traduisant la REalité*) method to assess sustainability against the risks of earthquakes. Various MCDM methods are also being used for quality of living and living standard assessment and comparison (Kuszewski, Sielska 2010; Dinçer 2011; Sahin, Yapici Pehlivan 2017).

Multicriteria methods are also used for the analysis in health sector; however, it seems that most of their applications are for specific issues, not overall comparisons. Glaize et al. (2019) provide a detailed overview of their applications in healthcare decision-making.

TOPSIS

In the paper, the TOPSIS method (*Technique for Order Preference by Similarity to Ideal Solution*, Hwang, Yoon 1981) was used to construct rankings. It was chosen due to its easy interpretation and the relative independence of the additional parameters which are used in MCDM to describe a decision-maker's preferences. In the first step, TOPSIS requires each alternative (in this case – each country) to be assessed based on all of the criteria. The evaluation of an *i-th* alternative with respect to a *j-th* criterion is denoted in the formulas below by $f_j(a_i)$. In the second step, the values of those functions are standardized to compensate for the effect of scale. There are a few standardization formulas that may be used (Roszkowska 2011). In the paper, we used a formula by which the standardized values are defined as follows:

$$r_{ij} = \frac{f_j(a_i)}{\sqrt{\sum_{i=1}^{n} f_j(a_i)^2}}$$

Next, standardized values are weighted. In MCDM, the decision-maker's preferences are introduced in this step. In the multidimensional analysis, weights are supposed to reflect the relative importance of criteria and stem from the data.

The fourth step of ranking requires the construction of two reference points, based on which a decision is made. The ideal solution, i.e. the hypothetical best alternative, is defined using the values of criteria for the alternatives taken into account. Therefore, that point is not chosen arbitrarily but is data-driven. It is defined as:

$$T^+ = \left(t_1^+, t_2^+, \ldots, t_n^+\right)$$

where:

$$t_j^+ = \begin{cases} \max\limits_i t_{ij} & \text{if high values of } j\text{-th criterion are considered better (benefit criterion)} \\ \min\limits_i t_{ij} & \text{if low values of } j\text{-th criterion are considered better (cost criterion)} \end{cases}$$

$\boldsymbol{t_{ij}}$ denotes value of $\boldsymbol{f_j(a_i)}$ weighted using weight $\boldsymbol{w_j}$

Similarly, the negative ideal can be interpreted as the hypothetical worst alternative. Once again, the choice of the point does not depend on the analyst or decision-maker; its coordinates stem from the data. The negative ideal is defined as

$$T^- = \left(t_1^-, t_2^-, \ldots, t_n^-\right)$$

where:

$$t_j^- = \begin{cases} \min\limits_i t_{ij} & \text{if high values of } j\text{-th criterion are considered better (benefit criterion)} \\ \max\limits_i t_{ij} & \text{if low values of } j\text{-th criterion are considered better (cost criterion)} \end{cases}$$

If the problem is not a trivial one, both points mentioned above do not exist in the dataset but are artificial constructs. It can be seen that the best alternative should be as close to the ideal solution and/or as distant from the negative ideal solution as possible. This reasoning follows intuition and makes a foundation for the easy interpretation of the TOPSIS method.

The distance between each alternative and reference points can be measured in any metric, but it seems that Euclidean distance is most commonly used. It also provides more intuitive interpretation. Therefore, the final evaluation of alternatives is based on the relative distance, defined as:

$$D_p(a_i) = \frac{d_p^-(a_i)}{d_p^-(a_i) + d_p^+(a_i)}$$

where

$d_p^-(a_i) = \left(\sum_{j=1}^{n}\left(t_{ij} - t_j^-\right)^2\right)^2$ denotes the distance of alternative a_i from the negative ideal and

$d_p^+(a_i) = \left(\sum_{j=1}^{n}\left(t_{ij} - t_j^+\right)^2\right)^2$ denotes the distance of alternative a_i from the ideal.

The intuition behind that formula is that the optimal solution should simultaneously be as close to the ideal solution and as distant from the negative ideal solution as possible.

Because of the uncomplicated calculations and straightforward interpretation, the TOPSIS method is one of the MCDM methods most often used for multidimensional analysis, usually for comparisons among countries (see e.g. Baer-Nawrocka, Markiewicz 2010; Ertman 2011; Balcerzak, Pietrzak 2016; Eyüboğlu 2016; Masca 2017) or smaller territorial units (cf. Roszkowska, Karwowska 2014; Demirel, Tüzün 2016; Matel, Marcinkiewicz 2017; Roznerski, Brudniak 2019; Stecyk 2019). It is also used in specific applications in the health sector (Sheykholeslami et al. 2015; Rađenović, Veselinović 2017; Hosseini et al. 2019).

Criteria and Data

In the analysis, two sets of criteria are used, which are further referred to as input and output criteria. Data on the following categories have been collected from the Eurostat database (https://ec.europa.eu/eurostat):

– Output criteria:
- Healthy life years at age 65 – females (*hyf*)
- Healthy life years at age 65 – males (*hym*)
- Life expectancy at age 65 – females (*lef*)
- Life expectancy at age 65 – males (*lem*)
- Healthy life years in absolute value at birth – females (*hyfb*)
- Healthy life years in absolute value at birth – males (*hymb*)
- Life expectancy in absolute value at birth – females (*lefb*)
- Life expectancy in absolute value at birth – males (*lemb*)
- Share of people with good or very good perceived health (*good*)
- Infant mortality rate (*infmort*)

– Input criteria
- Practising dentists per hundred thousand inhabitants (*dent*)
- Practising physicians per hundred thousand inhabitants (*doctors*)
- Hospital beds per hundred thousand inhabitants (*beds*)
- Health care expenditure (million euro) per hundred thousand inhabitants (*money*)
- Computed Tomography Scanners in hospitals per hundred thousand inhabitants (*E1*)
- Magnetic Resonance Imaging Units in hospitals per hundred thousand inhabitants (*E2*)
- Gamma cameras in hospitals per hundred thousand inhabitants (*E3*).

All the criteria except for infant mortality rate (*infmort*) were considered to be benefit criteria, for which higher values are considered better. *Infmort* was considered to be a cost criterion.

Although the criteria presented above are divided into output units, which describe the health status of the population (*hyf* – *infmort*), and input criteria (*dent* – *E3*) we do not use DEA, because one of the assumptions for this method is that decision making units (in this case countries) are homogeneous (Gajdzik 2015). It is doubtful if this condition is satisfied in the face of different healthcare systems, needs and backgrounds in the analysed countries.

In the analysis, we focused on the period 2013–2016 and the following countries: Belgium, Bulgaria, Czechia, Denmark, Germany, Estonia, Ireland, Greece, Spain, France, Croatia, Italy, Cyprus, Latvia, Lithuania, Luxembourg, Hungary, Malta, Netherlands, Austria, Poland, Portugal, Romania, Slovenia, Slovakia, Finland, Sweden, United Kingdom, Iceland, Liechtenstein, Norway and Switzerland.

Four rankings were constructed. The first ranking, further denoted by R1, was based only on the output criteria (*hyf*, *hym*, *lef*, *lem*, *hyfb, hymb*, *lefb*, *lemb*, *good* and *infmort*), which reflect the overall health care performance and efficiency. Rankings were constructed for 2013–2016. Cases in which data were missing were relatively rare (comprising less than 5% of observations). If an individual value was missing in a time series, it was replaced by the mean of values from the previous and following years. In cases where the first or last values in the series were missing, EU means for each criterion in the given year were used instead.

Due to missing data, in order to compare rankings based on input criteria to other orderings, the rest of the analysis was performed for a shorter period (2014–2016) and on a smaller set of countries (Belgium, Bulgaria, Germany, Estonia, France, Croatia, Italy, Cyprus, Latvia, Lithuania, Luxembourg, Netherlands, Austria, Poland, Romania and Switzerland). After that change, the rate of missing data was very low (5 cases). In those few cases, the means of values from the previous and following years were used. Rankings R2–R4 were built based on this smaller dataset.

The second ranking (further denoted by R2) was built based on a set of output criteria (*hyf* – *infmort*). The third ranking (R3) was built using the whole set of criteria (both input and output criteria). The last, fourth ranking (R4) used only input criteria, i.e. *dent, doctors, beds*, *money, E1, E2,* and *E3*.

Criteria Weights

In the absence of the decision-maker, the alternatives need to be compared based on their differences and similarities, not subjective preferences. Weights were assigned to criteria based on the information each criterion provides. Criteria that do not contribute towards differences between alternatives are given lower weights. In the hypothetical case of zero differences between the alternatives, with respect to a given criterion, the alternatives would be considered equal based on that criterion. Therefore, such a criterion contributes no information and can be omitted in the ranking process (cf. Deng et al. 2000; Rao 2013; Wang, Luo 2010; Xu 2004). Similarly, if a criterion is not correlated with other criteria, the information it transmits can be considered unique and its removal may have a great impact on the ranking. Therefore, the criterion in question should be assigned a relatively higher weight.

The weighting procedure is as follows. Firstly, characteristics that were the most differentiated among the alternatives were ascribed higher weights:

$$w_i^1 = \frac{|v_i|}{\sum_{i=1}^{k}|v_i|}$$

$$v_i = \frac{s_i}{\overline{f_i}}$$

where: s_i denotes standard deviation of criterion *i*, $\overline{f_i}$ denotes the arithmetic mean of criterion *i* and *i = 1, ..., N*.

Secondly, those characteristics that were highly correlated with the other were assigned lower weights, to avoid repeated information:

$$w_i^2 = \frac{\sum_{j=1}^{N}|r^P{}_{ij}|}{\sum_{i=1}^{N}\sum_{j=1}^{N}|r^P{}_{ij}|}$$

where: r^P_{ij} denotes a correlation coefficient of criteria *i* and *j*.
The final formula according to which the weights have been calculated is the following:

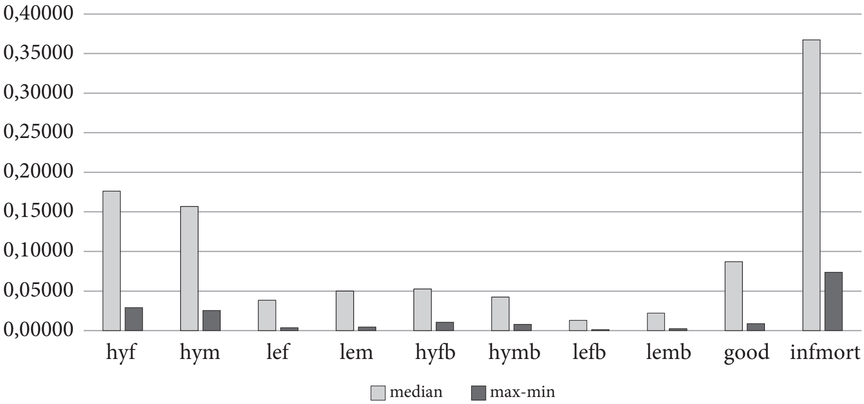
$$w_i = \frac{\frac{w_i^1}{w_i^2}}{\sum_{i=1}^{N}\frac{w_i^1}{w_i^2}}$$

The procedure, similar to the CRITIC method (*Criteria Importance Through Inter-criteria Correlation*) proposed by Diakoulaki et al. (1995) and the CCSD method (*Correlation Coefficient and Standard Deviation Integrated Approach*) proposed by Wang and Luo (2010) comes from Sielska (2010).

Results

Output criteria ranking R1 is based on criteria *hyf – infmort* which reflect the health quality in the population. The criteria weights differed each year, but the changes were rather small, as shown in Figure 1. 'Max-min' denotes the difference between the maximum and minimum weight. The highest importance overall was assigned to infant mortality rate (*infmort*), while the lowest was given to life expectancy in absolute value at birth for females (*lefb*). Healthy life years at age 65 for females (*hyf*) and males (*hym*) and the share of people with good or very good perceived health (*good*) were considered to be relatively important in the assessment.

Figure 1. Summary of weights for rankings R1



Source: The author's elaboration based on Eurostat data.

The whole R1 ranking is presented in Table 1. 'Max-min' in all the following tables denotes the difference between the maximum and minimum rank. The highest position was not stable; changes at the top of the ranking could be seen. Despite this, we can comment on the overall results. Firstly, the median rank was the highest in the case of Iceland, Norway and Sweden. In the case of those countries, rankings are rather stable, with the difference between highest and lowest ranking being no greater than 2. The quality of health is also high in Finland, Spain, Cyprus and Slovenia. On the other hand, there are countries which achieved low rankings each year. This group includes Bulgaria, Slovakia, Romania and Malta. Their positions (except Malta's) are very stable. Poland was ranked on average at 24th position. It is worth noticing that this position was rather stable, as the difference between maximum and minimum ranking was equal to 3.

Table 1. Rankings of countries based on output criteria (R1)

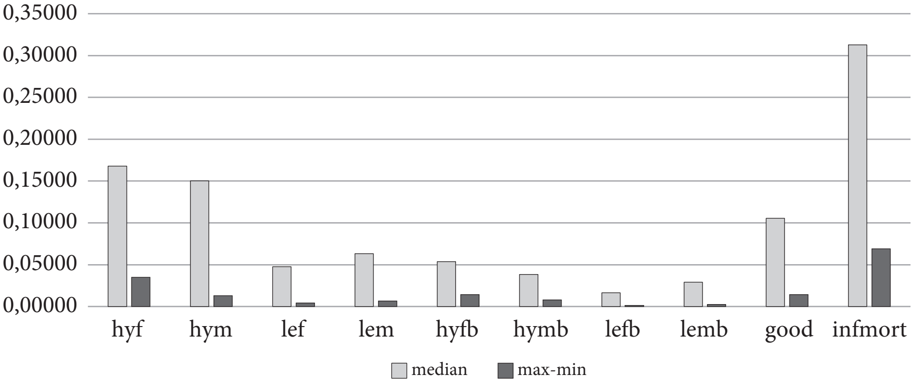
Year / Country	2013	2014	2015	2016	Medianrank	Max-min
Belgium	12	11	11	14	11.5	3
Bulgaria	31	31	31	31	31.0	0
Czechia	7	9	10	13	9.5	6

Year Country	2013	2014	2015	2016	Medianrank	Max-min
Denmark	9	15	14	12	13.0	6
Germany	17	22	6	15	16.0	16
Estonia	11	20	18	11	14.5	9
Ireland	13	6	9	9	9.0	7
Greece	22	23	23	26	23.0	4
Spain	6	10	7	8	7.5	4
France	16	13	17	20	16.5	7
Croatia	24	29	25	27	26.0	5
Italy	15	16	15	10	15.0	6
Cyprus	4	5	12	6	5.5	8
Latvia	26	27	28	24	26.5	4
Lithuania	23	26	26	28	26.0	5
Luxembourg	18	4	8	22	13.0	18
Hungary	27	28	24	25	26.0	4
Malta	30	24	27	30	28.5	6
Netherlands	21	14	13	17	15.5	8
Austria	10	17	16	16	16.0	7
Poland	25	25	22	23	24.0	3
Portugal	8	21	19	18	18.5	13
Romania	32	32	32	32	32.0	0
Slovenia	14	8	5	5	6.5	9
Slovakia	28	30	29	29	29.0	2
Finland	5	7	4	4	4.5	3
Sweden	3	1	3	3	3.0	2
United Kingdom	19	18	20	21	19.5	3
Iceland	1	2	1	1	1.0	1
Liechtenstein	29	12	30	7	20.5	23
Norway	2	3	2	2	2.0	1
Switzerland	20	19	21	19	19.5	2

Source: The author's elaboration based on Eurostat data.

A summary of weights for ranking R2 (second ranking based on output criteria) is presented in Figure 2. Overall, the most important criterion in the analysis is infant mortality (*infmort*). Healthy life years at age 65 for females (*hyf*) and males (*hym*) and the share of people with good or very good perceived health (*good*) were also assigned relatively high weights. Other criteria had negligible impact. It should also be noted that the criteria weights were relatively stable. The differences between the highest and lowest weights are no greater than 0.05. The only exception is the *infmort* criterion, where the difference is greater though still not big, compared to the median weight.

Figure 2. Summary of weights for rankings R2



Source: The author's elaboration based on Eurostat data.

Several observations can be made based on the R2 rankings for 2014–2016, presented in Table 2. Germany and Luxembourg have the least stable positions, while the most stable rankings were achieved by Bulgaria and Romania. The last two countries were ranked lowest with medians of 15 and 16, respectively. In 2014, the best quality of healthcare performance ecould be found in Cyprus, Luxembourg and Italy. In the following year, Germany, Belgium and Luxembourg were ranked in the first 3 positions. In the last year of the analysis, 2016, the top 3 positions were those of Cyprus, Italy and Belgium. Poland was systematically assessed as characterized by a rather low quality of healthcare performance (ranked at 12 in 2014, and 11 in 2015–2016).

Table 2. Rankings of countries based on the output criteria (R2)

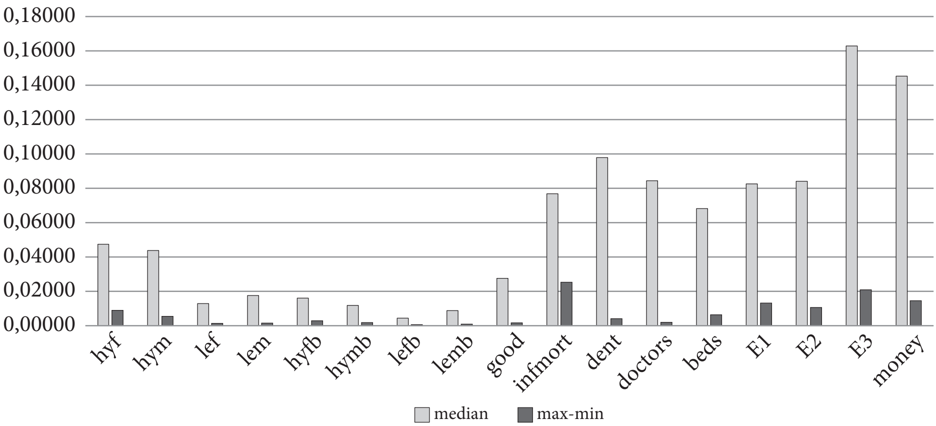
Year / Country	2014	2015	2016	Medianrank	Max-min
Belgium	5	2	3	3	3
Bulgaria	15	15	15	15	0
Germany	9	1	4	4	8
Estonia	8	10	9	9	2
France	6	5	7	6	2

Year / Country	2014	2015	2016	Medianrank	Max-min
Croatia	14	13	13	13	1
Italy	3	7	2	3	5
Cyprus	1	6	1	1	5
Latvia	13	14	12	13	2
Lithuania	11	12	14	12	3
Luxembourg	2	3	10	3	8
Netherlands	7	4	5	5	3
Austria	4	8	8	8	4
Poland	12	11	11	11	1
Romania	16	16	16	16	0
Switzerland	10	9	6	9	4

Source: Own elaboration based on Eurostat data.

In the next step, criteria which reflect the resources of healthcare sector have been added and rankings based on both input and output criteria (R3) were constructed. According to Figure 3, shown below, the most important criteria used for the R3 ranking were input criteria, mainly health care expenditures (*money*) and gamma cameras in hospitals (*E3*). Infant mortality (*infmort*), the number of practising dentists (*dent*) and doctors (*doctors*), the number of beds (*beds*) and other equipment (Computed Tomography Scanners – *E1*, and Magnetic Resonance Imaging Units – *E2*) were also considered to be important. Output criteria, connected to the results of healthcare performance, such as life expectancy, were assigned lower weights.

Figure 3. Summary of weights for rankings R3



Source: The author's elaboration based on Eurostat data.

Table 3 presents the ranking results for 2014–2016. All the positions are very stable; the greatest difference between the highest and the lowest ranking is 2 (for Bulgaria and Estonia). Belgium is ranked first each year. Switzerland and Luxembourg were assigned the second and third position, respectively. Poland is ranked 15th, and that position does not change.

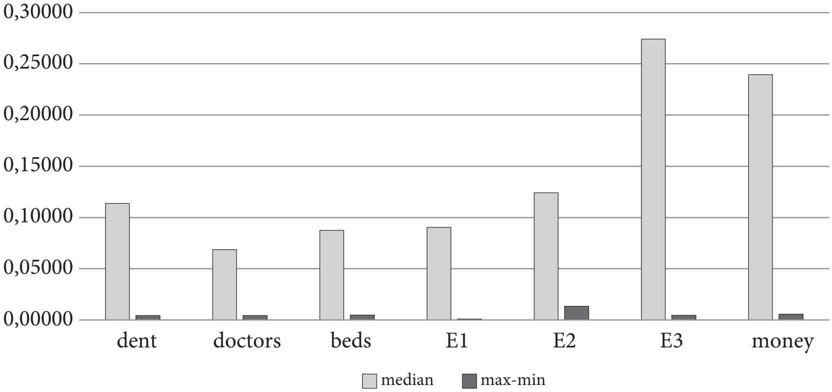
Table 3. Rankings of countries based on the whole set of criteria (R3)

Country / Year	2014	2015	2016	Medianrank	Max-min
Belgium	1	1	1	1	0
Bulgaria	13	11	11	11	2
Germany	7	6	7	7	1
Estonia	10	12	12	12	2
France	9	9	9	9	0
Croatia	11	10	10	10	1
Italy	8	8	8	8	0
Cyprus	4	4	4	4	0
Latvia	14	14	14	14	0
Lithuania	12	13	13	13	1
Luxembourg	2	3	3	3	1
Netherlands	6	7	6	6	1
Austria	5	5	5	5	0
Poland	15	15	15	15	0
Romania	16	16	16	16	0
Switzerland	3	2	2	2	1

Source: The author's elaboration based on Eurostat data.

The summary of weights used for the construction of the R4 ranking (based on input criteria only, such as dentists, doctors, hospital beds and equipment) is presented in Figure 4. The highest importance was assigned to gamma cameras in hospitals (*E3*) and health care expenditures (*money*). Weights assigned to other criteria were on similar levels.

Figure 4. Summary of weights for rankings R4



Source: The author's elaboration based on Eurostat data.

According to Table 4, which presents the results for 2014–2016, Belgium ranked first each year. Similarly, stable and high positions were achieved by Switzerland and Luxembourg. Very stable rankings were assigned to Germany, Bulgaria, Poland, Croatia, Latvia and Romania, however most of those countries (all apart from Germany) were ranked low. Poland ranked 15th (second to last) each year.

Table 4. Rankings of countries based on input criteria (R4)

Year / Country	2014	2015	2016	Medianrank	Max-min
Belgium	1	1	1	1	0
Bulgaria	11	11	11	11	0
Germany	7	7	7	7	0
Estonia	12	13	13	13	1
France	9	9	8	9	1
Croatia	10	10	10	10	0
Italy	8	8	9	8	1
Cyprus	5	4	4	4	1
Latvia	14	14	14	14	0
Lithuania	13	12	12	12	1
Luxembourg	2	3	3	3	1
Netherlands	4	6	6	6	2
Austria	6	5	5	5	1
Poland	15	15	15	15	0
Romania	16	16	16	16	0
Switzerland	3	2	2	2	1

Source: The author's elaboration based on Eurostat data.

Rankings Comparison

The goal of the final step of the analysis is to compare rankings, based on three different sets of criteria, and to analyse the changes over time.

Firstly, the evolution of rankings year after year is assessed based on Spearman's rank correlation coefficient (ρ). The results presented in Table 5 below show that with the exception of the R1 ranking, in which the values of Spearman's ρ fluctuate, the results become more similar over the years. In general, the values of correlation coefficients are high, which reflect a high similarity – the highest in those rankings which are based on input criteria (R4), and the lowest among rankings built based on output criteria only (R2). The similarities are also statistically significant.

Table 5. Similarity of rankings in period 2013–2016 measured by Spearman's rank correlation coefficient

Ranking / Years compared	R1		R2		R3		R4	
	Spearman's ρ	p-value	Spearman's ρ	p-value	Spearman's ρ	p-value	Spearman's ρ	p-value
2013–2014	0.7984	<0.0001	–	–	–	–	–	–
2013–2015	0.8651	<0.0001	–	–	–	–	–	–
2013–2016	0.8314	<0.0001	–	–	–	–	–	–
2014–2015	0.8578	<0.0001	0.7794	0.0006	0.9794	<0.0001	0.9853	<0.0001
2014–2016	0.8644	<0.0001	0.7882	0.0004	0.9824	<0.0001	0.9824	<0.0001
2015–2016	0.8141	<0.0001	0.8059	0.0002	0.9971	<0.0001	0.9971	<0.0001

Note: '–' no rankings constructed for at least one of two compared years.

Source: The author's elaboration based on Eurostat data.

Rankings allow for easier and more intuitive comparisons, though they also entail some information loss. The relative distance measures which are $D_p(a_i)$ calculated in the TOPSIS procedure can be used for more detailed similarity assessments. Using those values, we may focus on the final order while also considering the differences between alternatives (countries). In the study, to compare values of $D_p(a_i)$ for different rankings, the concept of distance was used once again. Each ranking was considered to be a point from n-dimensional space, where n=32 (in case of R1) or n=16 (for R2–R4) is the number of countries analysed, and the coordinates are represented by the values of the $D_{p(ai)}$ function. The distance between such points (rankings) was then measured using the Euclidean metric. The results are presented in Table 6, below. As is always the case with distance measures, the lower values reflect greater similarity.

Table 6. Euclidean distances between rankings in 2013–2016

Ranking / Years compared	R1	R2	R3	R4
2013–2014	0.5212	–	–	–
2013–2015	0.5268	–	–	–
2013–2016	0.7172	–	–	–
2014–2015	0.4914	0.3417	0.1400	0.0994
2014–2016	0.5902	0.4152	0.2292	0.1378
2015–2016	0.5470	0.3107	0.1022	0.0612

Note: '–' no rankings constructed for at least one of two compared years.

Source: The author's elaboration based on Eurostat data.

In the output criteria R1 ranking, the differences between rankings are fluctuating. The dissimilarities between the ranking for 2013 and the rankings constructed for other years are growing with time. The same conclusion can be made regarding rankings built for 2014. That means that there were changes in the relative quality of healthcare performance. In the case of R2, the rankings were most similar in 2015-2016. In 2014-2015, the difference was slightly greater. For R3, the evaluations are becoming more similar with time. In R4, similar evaluations were made for 2014-2015 and 2015-2016. The lowest values of the distance function were found among rankings based only upon input criteria (R4). That leads one to the conclusion that the relative situation in this sphere is the least prone to change. This confirms the previous conclusions, drawn from the values of Spearman's rank correlation coefficient. In the second step, rankings built based on different sets of criteria were compared. Rankings R2-R4 were included.

The results of a comparison based on Spearman's rank correlation coefficient ρ are presented in Table 7. The similarity is definitely the highest between rankings R3 and R4 – that is, those rankings that include input criteria. The similarities between other pairs or rankings (R2–R3 and R2–R4) are slightly less substantial but still clear. All Spearman's rank correlation coefficients are significant, at $\alpha=0.01$.

Table 7. Similaritiesm among rankings R2, R3 and R4, measured by Spearman’s rank correlation coefficient

Year / Ranking	2014		2015		2016	
	Spearman’s ρ	p-value	Spearman’s ρ	p-value	Spearman’s ρ	p-value
R2–R3	0.7912	0.0004	0.7647	0.0009	0.7176	0.0024
R2–R4	0.7147	0.0026	0.7500	0.0012	0.6882	0.0042
R3–R4	0.9765	<0.0001	0.9941	<0.0001	0.9941	<0.0001

Source: The author’s elaboration based on Eurostat data.

The values of the distance function presented in Table 8 confirm those results. It can be seen that in all 3 years the most similar rankings were R3 and R4, that is rankings built considering input criteria. It leads to the conclusion, that that determinants of healthcare quality are crucial in countries assessment.

Table 8. Euclidean distances between rankings between rankings R2, R3 and R4, in given years

Year / Ranking	2014	2015	2016
R2–R3	1.4080	1.3857	1.4006
R2–R4	1.6216	1.5350	1.4947
R3–R4	0.2325	0.1703	0.1104

Source: The author’s elaboration based on Eurostat data.

Summary

In the paper, we attempt to assess the situation in healthcare systems, based on criteria reflecting both the health situation in the population (such as life expectancy at birth, or the share of inhabitants who perceive their health status as ‘good’), which reflects healthcare system performance, and the characteristics of resources like diagnostic equipment and health care expenditures. Based on the TOPSIS method, four rankings were constructed. The results show that input criteria are crucial in the assessment. It is also found that the relative situation of countries, with respect to resources, is the least prone to change. This conclusion may seem surprising, because the change of inputs may occur at a faster rate than changes in the health situation of the population which is considered an output; however, it should be highlighted that the resources depend on the socio-economic situation and political determinants, which in most cases evolve slowly.

The results show that countries that were ranked high usually had stable positions, as well. Poland was systematically assessed as being characterized by rather low quality of healthcare performance, as well as low resources.

In general, the similarities among rankings in a given field is growing; therefore, we may conclude that the situation and relative differences between respective countries are stabilizing.

References

Ardielli, E. (2016), *Comparison of Multiple Criteria Decision Making Approaches: Evaluating eGovernment Development*. Littera Scripta [online]. České Budějovice: The Institute of Technology and Business in České Budějovice, 9(2): 10–19, http://journals.vstecb.cz/wp-content/uploads/2017/02/Ardielli-pro-tisk-2.pdf, accessed: 2019.12.20

Asandului, L., Roman, M., Fatulescu, P. (2014), 'The efficiency of healthcare systems in Europe: A Data Envelopment Analysis Approach', *Procedia Economics and Finance* 10: 261–268, https://doi.org/10.1016/s2212-5671(14)00301-3

Baer-Nawrocka, A., Markiewicz, N. (2010), 'Zróżnicowanie przestrzenne potencjału produkcyjnego rolnictwa w krajach Unii Europejskiej', *Roczniki Nauk Rolniczych*, Seria G, 97(4): 9–15, http://sj.wne.sggw.pl/pdf/RNR_2010_n4_s9.pdf, accessed: 9.12.2019

Balcerzak, A.P., Pietrzak, M.P. (2016), 'Application of TOPSIS Method for Analysis of Sustainable Development in European Union Countries', in: T. Loster, T. Pavelka (eds.), *The 10th International Days of Statistics and Economics. Conference Proceedings. September 8–10, 2016*, Prague: Institute of Economic Research: 82–92, http://econpapers.repec.org/bookchap/pesecchap/19.htm, accessed: 9.12.2019

Dąbrowa, M. (2011), 'Study in standard of living – methodology of structure of selected indicators'. The Małopolska School of Economics in Tarnów, *Research Papers Collection*1: 67–82, https://www.google.pl/url?sa=t&rct=j&q=&esrc=s& source=web&cd=3&cad=rja&uact=8&ved=2ahUKEwiKgJe8hL_mAhVx_SoKHepOC_kQFjACegQIBRAC&url=http%3A%2F%2Fyadda.icm.edu.pl%2Fyadda%2Felement%2Fbwmeta1.element.desklight-96b5b665-2e85-4d3b-bdb9-5f0a1be8db38%2Fc%2F67-82.pdf&usg=AOvVaw0N4TwZ1yLkJEIxGNVH2Sun, accessed: 11.12.2019

Dasgupta, P. (2000), *Valuation and Evaluation: Measuring the Quality of Life and Evaluating Policy* (Resources for the Future Discussion Paper 00–24), https://pdfs.semantic scholar.org/0c00/45ccaca901247a67b2536013d4bd695c9e41.pdf, accessed: 11.12.2019

Demirel, T., Tüzün, S. (2016), 'Analyzing and Comparing the Districts of Istanbul Using TOPSIS', *International Journal of Arts & Sciences* 09(01): 339–348, http://www.universitypublications.net/ijas/0901/pdf/DE5C492.pdf, accessed: 9.12.2019

Deng, H., Yeh, Ch.-S., Willis R.J. (2000), 'Inter-company comparison using modified TOPSIS with objective weights', *Computers & Operations Research* 27: 963–973, https://doi.org/10.1016/s0305-0548(99)00069-6

Diakoulaki, D., Mavrotas, G., Papayannakis, L. (1995), 'Determining objective weights in multiple criteria problems: The CRITIC method', *Computers & Operations Research*, *22*(7): 763–770, https://doi.org/10.1016/0305-0548(94)00059-h

Dinçer, S.E. (2011), 'Multi-criteria Analysis of Economic Activity for European Union Member States and Candidate Countries: TOPSIS and WSA Applications', *European Journal of Social Sciences* 21(4): 563–572, http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.465.4762&rep=rep1&type=pdf, accessed: 9.12.2019

Ertman, A. (2011), 'Zróżnicowanie elastyczności rynków pracy w wybranych krajach europejskich oraz USA w świetle metody TOPSIS', *Oeconomia Copernicana* 2(3): 43–64, https://doi.org/10.12775/oec.2011.012

Eyüboğlu, K. (2016), 'Comparison of Developing Countries' Macro Performances with AHP and TOPSIS Methods', *Journal of The Faculty of Economics and Administrative Sciences* 6(1): 131–146, https://dergipark.org.tr/en/download/article-file/382409, accessed: 9.12.2019

Gajdzik, B. (2015), 'Próba implementacji metody DEA do oceny efektywności procesu restrukturyzacji sektora hutniczego w Polsce', w: R. Knosala (red.), *Innowacje w zarządzaniu i inżynierii produkcji*, T. 1, Opole: Oficyna Wydaw. Polskiego Towarzystwa Zarządzania Produkcją: 424–435

Girginer, N. (2013), 'A Comparison of the Healthcare Indicators of Turkey and The European Union Members Countries Using Multidimensional Scaling Analysis and Cluster Analysis', *Iktisat Isletme ve Finans, Bilgesel Yayincilik* 28(323): 55–72, https://doi.org/10.3848/iif.2013.323.3387

Glaize, A., Duenas, A., Di Martinelly, Ch., Fagnot, I. (2019), 'Healthcare decision-making applications using multicriteria decision analysis: A scoping review', *Journal of Multi-Criteria Decision Analysis. Optimization, Learning and Decision Support* 26(1–2): 62–83, https://doi.org/10.1002/mcda.1659

Grausová, M., Hužvár M., Štrangfeldová, J. (2014), 'Healthcare systems efficiency in the Visegrád Group', 17th AMSE, Applications of Mathematics in Economics. International

Scientific Conference: Poland, 27–31 August, 2014. Conference Proceedings Full Text Papers: 104–113, https://doi.org/10.15611/amse.2014.17.11

Health care (n.d.), Retrieved from at http://www.numbeo.com/health-care/indices_explained.jsp (11.12.2019)

Hosseini, S.M., Bahadori, M., Raadabadi, M., Ravangard, R. (2019), 'Ranking Hospitals Based on the Disasters Preparedness Using the TOPSIS Technique in Western Iran', *Hospital Topics* 97(1): 23–31. DOI: 10.1080/00185868.2018.1556571

https://ec.europa.eu/eurostat, accessed: 3.12.2019

Hwang C.L., Yoon K. (1981), *Multiple Attribute Decision Making: Methods and Applications*. New York: Springer-Verlag

Kriksciuniene, D., Sakalauskas, V. (2017), 'DEA Based Algorithm for EU Healthcare Efficiency Evaluation', in: W. Abramowicz (ed.), *Business Information Systems Workshops. BIS 2017. Lecture Notes in Business Information Processing*, Vol. 303: 97–109, Cham: Springer

Kringos, D., Boerma, W., Bourgueil, Y., Cartier, T., Dedeu, T., Hasvold, T., Hutchinson, A., Lember, M., Oleszczyk, M., RotarPavlic, D., Svab, I., Tedeschi, P., Wilm, S., Wilson, A., Windak, A., Van der Zee, J., Groenewegen, P. (2013), 'The strength of primary care in Europe: an international comparative study', *British Journal of General Practice,* November 63 (616); e742-e750, https://doi.org/10.3399/bjgp13x674422

Kujawska, J. (2015), 'Wydatki na opiekę zdrowotną a efekty zdrowotne – analiza porównawcza krajów europejskich metodą DEA', *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu* 412: 156–165, https://doi.org/10.15611/pn.2015.412.13

Kujawska, J. (2018), 'Efficiency of healthcare systems in European countries – the DEA network approach', *Metody Ilościowe w Badaniach Ekonomicznych XIX* (1): 60–70, https://doi.org/10.22630/mibe.2018.19.1.6

Kuszewski, T., Sielska, A. (2010), 'Użyteczna sztuka rankingów ekonomiczno-społecznych', *Współczesna Ekonomia* 1(13): 143–162, http://yadda.icm.edu.pl/yadda/element/bwmeta1.element.ekon-element-000171289207, accessed: 9.12.2019

Majka, A. (2015), 'Przestrzenne zróżnicowanie poziomu życia ludności w Polsce w ujęciu dynamicznym' , *Wiadomości Statystyczne* 5: 27–41, http://cejsh.icm.edu.pl/cejsh/element/bwmeta1.element.ekon-element-000171388639, accessed: 16.12.2019

Masca, M. (2017), 'Economic performance evaluation of European Union countries by TOPSIS method', *North Economic Review* 1: 83–94, https://pdfs. semanticscholar.org/efb6/35ef12fe22243d76e050a1d061e0e478b48b.pdf?_ga=2.74830162 .1758802572.1576485176-364601084.1576485176, accessed: 9.12.2019

Matel, A., Marcinkiewicz, J. (2017), 'Analiza dostępności mieszkań w miastach wojewódzkich Polski z wykorzystaniem metody TOPSIS', *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu* 475: 170–182, doi:10.15611/pn.2017.475.15

Miszczyńska, K. (2013), 'Multivariate analysis of healthcare systems in selected European Union countries. Cluster analysis', *Quantitative methods in economic* XIV(2): 99–107, http://cejsh.icm.edu.pl/cejsh/element/bwmeta1.element.desklight -4dac21df-9540-47fc-bf1a-79c92c2c45d7, accessed: 16.12.2019

Murgante, B., Salmani, M., Molaei Qelichi, M., Hajilo, M. (2017), 'A Multiple Criteria Decision-Making Approach to Evaluate the Sustainability Indicators in the Villagers' Lives in Iran with Emphasis on Earthquake Hazard: A Case Study', *Sustainability* 9: 1491, https://doi.org/10.20944/preprints201703.0154.v2

Nixon, J. (2000*), How does the UK NHS compare with European standards? A review of EU health care systems using hierarchical cluster analysis* (The University of York discussion paper no 182), https://cirano.qc.ca/icirano/public/pdf/webevents201103_british_york_1.pdf, accessed: 10.12.2019

Önen, Z., Sayın, S. (2018), 'Evaluating Healthcare System Efficiency of OECD Countries: A DEA-Based Study', in: C. Kahraman, Y. Topcu (eds.), *Operations Research Applications in Health Care Management. International Series in Operations Research & Management Science* 262: 141–158, https://doi.org/10.1007/978-3-319-65455-3_6

Panagiota Digkoglou, P., Papathanasiou, J. (2018), 'Ranking the EU countries according to the environmental performance index using PROMETHEE', *International Journal of Sustainable Agricultural Management and Informatics* 4(3–4): 290–305, https://doi.org/10.1504/ijsami.2018.10020709

Pelone, F., de Belvis, A.G., Volpe, M., Ricciardi, W., (2008), 'Is there a relationship between health care models and their performance assessment? The results of an extensive review', *Italian Journal of Public Health* 5(2): 102–106, https://pdfs.semanticscholar.org/db83/2e26f911627e57d26746d6bf2717ce766cc3.pdf, accessed: 10.12.2019

Pezer, M. (2018), 'Maternity support policies: a cluster analysis of 22 European Union countries', in: *Proceedings of the ISCCRO – International Statistical Conference in Croatia ISCCRO'18,* 10–11 May 2018, Opatija, Croatia, https://www.google.pl/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=2ahUKEwjwxd-D3bnmAhWSrIsKHWLcCHsQFjAAegQIARAJ&url=https%3A%2F%2Fwww.bib.irb.hr%2F937900%2Fdownload%2F937900.Pezer_M._-_ISCCRO_18.pdf&usg=AOvVaw2qmDTDq_53P8MUJegTs7kZ, accessed: 16.12.2019

Przybyła, K. (2015), 'Poziom życia w wybranych miastach województwa śląskiego', *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu* 391: 183–192, https://doi.org/10.15611/pn.2015.391.19

Rađenović, Ž., Veselinović, I. (2017), 'Integrated AHP-TOPSIS Method for the Assessment of Health Management Information Systems Efficiency', *Economic Themes* 55(1): 121–142, https://doi.org/10.1515/ethemes-2017-0008

Rao, R.V. (2013), 'A Novel Subjective and Objective Integrated Multiple Attribute Decision Making Method', in: R.V. Rao, *Decision Making in Manufacturing Environment Using Graph Theory and Fuzzy Multiple Attribute Decision Making Methods*. Springer Series in Advanced Manufacturing, London: Springer: 137–157, https://doi.org/10.1007/978-1-4471-4375-8_4

Roszkowska, E. (2011), 'Multi-criteria decision making models by applying the TOPSIS method to crisp and interval data', in: T. Trzaskalik, T. Wachowicz (eds.), *Multiple Criteria Decision Making,* Vol. 6, Katowice: The University of Economics: 200–230, http://yadda.icm.edu.pl/yadda/element/bwmeta1.element.ekon-element-000171231687, accessed: 9.12.2019

Roszkowska, E. (2013), 'Rank ordering criteria weighting methods – a comparative overview', *Optimum. Studia Ekonomiczne 5*(65): 14–33, https://doi.org/10.15290/ose.2013.05.65.02

Roszkowska, E., Karwowska, R. (2014), 'Analiza porównawcza województw Polski ze względu na poziom zrównoważonego rozwoju w roku 2010', *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu. Taksonomia* 23(328): 30–40, https://www.dbc.wroc.pl/publication/27850, accessed: 16.12.2019

Roszkowska, E., Filipowicz-Chomko, M. (2019), 'Measuring sustainable development in the education area using multi-criteria methods: a case study', *Central European Journal of Operations Research:* 1–23, https://doi.org/10.1007/s10100-019-00641-0

Roznerski, M., Brudniak, G., Szymanowski, W. (2019), 'Analiza rozbieżności przestrzennych województw Polski w obszarze systemu bodźców gospodarczych i instytucjonalnych (2009–2016)', *Roczniki Kolegium Analiz Ekonomicznych, Szkoła Główna Handlowa w Warszawie*, nr 54: 211–220, http://rocznikikae.sgh.waw.pl/p/roczniki_kae_z54_16.pdf, accessed: 9.12.2019

Sahin, A., Yapici Pehlivan, N. (2017), 'Evaluation of life quality by integrated method of AHP and TOPSIS based on interval type-2 fuzzy sets', *Hacettepe Journal of Mathematics and Statistics* 46(3): 511–523, https://doi.org/10.15672/hjms.2017.420

Sheykholeslami, S.J.S., Rahmani, S., Asghar Yousefi, A. (2015), 'Assessment of the access of the cities of Alborz province to healthcare infrastructure using TOPSIS and MORIS models', *Management Issues in Healthcare System* 1: 38–48, https://doi.org/10.33844/mihs.2015.60347

Sielska, A. (2010), 'O pewnej propozycji zastosowania rankingów wielokryterialnych w analizie portfelowej', *Prace Naukowe, Akademia Ekonomiczna w Katowicach*, *Modelowanie preferencji a ryzyko* '09: 85–99

Slottje, D.J. (1991), 'Measuring the Quality of Life Across Countries', *The Review of Economics and Statistics* 73(4): 684–693, https://doi.org/10.2307/2109407

Stecyk, A. (2019), 'The AHP-TOPSIS Model in the Analysis of the Counties Sustainable Development in the West Pomeranian Province in 2010 and 2017', *Journal of Ecological Engineering* 20(7): 233–244, https://doi.org/10.12911/22998993/109870

van Zanden, J., et al. (eds.) (2014), *How Was Life?: Global Well-being since 1820*, Paris: OECD Publishing

Walczak, R., Piekut, M., Kludacz-Alessandri, M., Sloka, B., Šimanskiene, L., Paas, T. (2018), 'Health care spending structures in Poland, Latvia, Lithuania and Estonia over the years as compared to other EU countries', *Foundations of Management* 10: 45–58, https://doi.org/10.2478/fman-2018-0005

Wang, Y.-M., Luo, Y. (2010), 'Integration of correlations with standard deviations for determining attribute weights in multiple attribute decision making', *Mathematical and Computer Modelling* 51: 1–12, https://doi.org/10.1016/j.mcm.2009.07.016

Xu, X. (2004), 'A note on the subjective and objective integrated approach to determine attribute weights', *European Journal of Operational Research* 156: 530–532, https://doi.org/10.1016/s0377-2217(03)00146-2

Yilmaz Işikhan, S., Güleç, D. (2018), 'The Clustering of World Countries Regarding Causes of Death and Health Risk Factors', *Iranian Journal of Public Health* 47(10): 1520–1528, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6277727/pdf/IJPH-47-1520.pdf, accessed: 10.12.2019

Zhang, B., Shmelev, S.E. (2019), 'Multidimensional sustainability assessment for the cities of the Global South: the PROMETHEE approach', in: S.E. Shmelev (ed.), *Sustainable Cities Reimagined: Multidimensional Assessment and Smart Solutions*, Routledge: 115–172

Žmuk, B. (2015), 'Quality of life indicators in selected European countries: Statistical Hierarchical cluster analysis approach', *Croatian Review of Economic, Business and Social Statistics (CREBSS)* 1(1–2): 42–54, https://doi.org/10.1515/crebss-2016-0004