EKONOMETRIA ECONOMETRICS 3(57) • 2017

ISSN 1507-3866
e-ISSN 2449-9994

Marek Walesiak
Wrocław University of Economics
e-mail: marek.walesiak@ue.wroc.pl

THE APPLICATION OF MULTIDIMENSIONAL SCALING TO MEASURE AND ASSESS CHANGES IN THE LEVEL OF SOCIAL COHESION OF THE LOWER SILESIA REGION IN THE PERIOD 2005-2015

ZASTOSOWANIE SKALOWANIA WIELOWYMIAROWEGO W POMIARZE I OCENIE ZMIAN POZIOMU SPÓJNOŚCI SPOŁECZNEJ WOJEWÓDZTWA DOLNOŚLĄSKIEGO W LATACH 2005-2015

DOI: 10.15611/ekt.2017.3.01
JEL Classification: C38, C43, C88

Summary: Social cohesion is the ability of territorial communities to ensure the welfare of all its members, to reduce social stratification and avoid polarization (EU/EC/CE 2004, p. 3). The assessment of social cohesion of the Lower Silesia region in the cross-section of counties in the period 2005-2015, was performed based on the variables from the following areas: income and economic activity of the population, living conditions of the population and the availability of services and public space. Multidimensional scaling combined with linear ordering and Theil decomposition were applied to measure and assess changes in the level of social cohesion of the Lower Silesia region in the cross-section of counties in the period 2005-2015. The level of social cohesion of all the counties (although in varying degrees) was increased and the degree of differentiation of social cohesion was slightly decreased during this period. The calculations were made with scripts prepared in R environment.

Keywords: social cohesion, multidimensional scaling, linear ordering, GDM1 distance, Theil decomposition, R program.

Streszczenie: Spójność społeczna to zdolność terytorialnych społeczności do zapewnienia dobrobytu wszystkim swoim członkom, zmniejszania rozwarstwień społecznych i unikania polaryzacji (EU/EC/CE 2004, s. 3). W ocenie spójności społecznej regionu dolnośląskiego w przekroju powiatów w latach 2005-2015 uwzględniono zmienne z następujących dziedzin: dochody i aktywność ekonomiczna ludności, warunki mieszkaniowe ludności, dostępność usług i przestrzeni publicznej. Do pomiaru i oceny zmian poziomu spójności społecznej województwa dolnośląskiego w przekroju powiatów w latach 2005-2015 zastosowano skalowanie wielowymiarowe w połączeniu z porządkowaniem liniowym oraz dekompozycję Theila.

Słowa kluczowe: spójność społeczna, skalowanie wielowymiarowe, porządkowanie liniowe, odległość GDM1, dekompozycja Theila, program R.

1. Introduction[1]

Due to the fact that social cohesion is a multifaceted problem, its unambiguous identification and measurement offers diverse possibilities (cf. e.g. [Chan, To, Chan 2006; Jenson 2010; Ryszkiewicz 2013]). The definition of social cohesion used by the Council of Europe was adopted in the article. The Council of Europe guide defines social cohesion as "the capacity of society to ensure long-term prosperity for all its members, including ensuring equal access to resources, respect for human dignity and diversity, personal and collective autonomy and responsible participation" [EU/EC/CE 2005, p. 23]. This definition reflects four main aspects of prosperity: fair and equal access to resources, individual and collective dignity, individual autonomy and participation in social life.

Social cohesion can be analysed and assessed in relation to more or less complex territorial arrangements, including e.g. regions, counties, municipalities and even individual cities and rural locations.

The purpose of the study is to measure and assess changes in the level and the degree of differentiation in the social cohesion of Lower Silesia counties in the period 2005-2015.

The measurement and assessment of changes in the level and the degree of differentiation in social cohesion requires a multivariate approach. To solve the empirical problem we have applied in research methodology a hybrid approach combining the results of multidimensional scaling and linear ordering. Such an approach, with the visualization of its results, enriches the interpretation. Additionally measure (6), with Theil's decomposition, was used for the overall assessment of changes in the level and degree of social cohesion differentiation of the Lower Silesia region in the cross-section of counties in the years 2005-2015.

The subject literature provides studies presenting the application of multivariate statistical analysis methods in the study of social cohesion in the cross-section of territorial units at a different scale. Polish literature, e.g. [Balcerzak 2015], presents the analysis of social cohesion in European Union countries using the taxonomic measure of development by Z. Hellwig [Hellwig 1968; 1972]. Multidimensional scaling, structural equation modelling (SEM) and composite index were applied to measure social cohesion in, respectively, 47 and 33 European countries in various studies [Dickes, Valentova 2013; Dickes, Valentova, Borsenberger 2010]. The au-

[1] Based on [Walesiak, Obrębalski 2017].

thors of the study [Rajulton, Ravanera, Beaujot 2007], based on the results of factor analysis and standardization, developed a composite index to measure social cohesion for 49 metropolitan areas of Canada (Census Metropolitan Areas).

2. Research methodology

Multidimensional scaling along with linear ordering and the Theil measure were applied to measure and assess changes in the level of social cohesion of the Lower Silesia region in the cross-section of counties in the period 2005-2015.

A two-step research procedure, allowing the visualisation of linear ordering results, presented in the study by Walesiak [2016b], was used in the article to order the analysed objects in terms of the level of social cohesion in the years 2010-2015. First, as a result of multidimensional scaling, the visualization of the objects' arrangement in two-dimensional space is developed. Next, the linear ordering of objects is carried out using composite measure based on Euclidean distance from the development pattern.

The research procedure, allowing the visualization of linear ordering results of the set of objects, covers the following steps (see [Walesiak 2016b]):

1) The choice of a complex phenomenon, which is not directly measurable (social cohesion level).

2) Defining the set of objects and the set of variables substantively related to the analysed complex phenomenon. The variables used to describe objects are measured on metric scales (ratio, interval scale). Preference variables[2] (stimulants, destimulants and nominants) are included among the variables.

3) Due to the fact that the data refer to two periods t and q ($t > q$) the procedure should:

a) change nominants into stimulants,

b) determine the joint pattern and anti-pattern of development[3] based on the data matrix $[x_{ij}]$ covering data from t (matrix $[x_{ij}^t]$) and q (matrix $[x_{ij}^q]$) periods. Therefore, $[x_{ij}]$ matrix has $(2n+2)\times m$ dimensions, where $i=1,...,n$ is the object's number and $j=1,...,m$ stands for the variable number,

c) perform the normalization of variable values for the joint data matrix from t and q periods, i.e. for $[x_{ij}]$ matrix. 18 normalization methods of variable values were presented in the article [Walesiak 2014].

4) The distance between objects is calculated and presented as $[\delta_{ik}]$ distance matrix. The following distance measures can be applied here: city-block, Euclidean,

[2] The definitions of a stimulant and a destimulant were presented in [Hellwig 1981, p. 48] and of a nominant in [Borys 1984, p. 118]. These definitions are available in e.g. [Walesiak 2016a, p. 18].

[3] The pattern (upper pole) includes the most favourable variable values, whereas the anti-pattern (lower pole) the least favourable values of the preference variables.

squared Euclidean, Chebyshev, GDM1 (see [Walesiak 2016a, pp. 27, 43]). When calculating distance either equal or different weights can be adopted.

Multidimensional scaling is performed: $f: \delta_{ik} \rightarrow d_{ik}$ for all pairs (i, k). Multidimensional scaling is the method for representing distance matrix between objects in m-dimensional space $[\delta_{ik}]$ as distance matrix between objects in q-dimensional space $(q < m)$ $[d_{ik}]$ for the purposes of graphical presentation (visualisation) and interpretation of the relationships occurring between the analysed objects.

The distances d_{ik} are always unknowns. That is, MDS must find a configuration of predetermined dimensions q on which the distances are computed.

The particular choice of f specifies the type of multidimensional scaling model (MDS). For metric data in particular:

- $f(\delta_{ik}) = b \cdot \delta_{ik} = d_{ik}$ – f stands for ratio MDS, (1)
- $f(\delta_{ik}) = a + b \cdot \delta_{ik} = d_{ik}$ – f stands for interval MDS, (2)
- $f(\delta_{ik}) = a + b \cdot \delta_{ik} + c \cdot \delta_{ik}^2 = d_{ik}$ – f stands for spline MDS. (3)

In practice $f(\delta_{ik})$ equals roughly d_{ik} and therefore $f(\delta_{ik}) \approx d_{ik}$. In multidimensional scaling $f(\delta_{ik}) = \hat{d}_{ik}$ ($\hat{d}_{ik}$ means: d-hats, disparities, pseudo distances – see [Borg, Groenen 2005, p. 199]).

Dimensions (q) are not directly observable. They represent latent variables, which allow for explaining similarities and differences between objects. Due to the possibility of the graphic presentation of linear ordering, result q equals 2. Iterative procedure in `smacof` algorithm is presented in the study [Borg, Groenen 2005, pp. 204-205].

The solution allowing the choice of an optimal multidimensional scaling procedure was used in the article due to the application of the variables normalization method, distance measure and scaling models, according to the procedure presented in [Walesiak, Dudek 2017c]. The procedure available in the `mdsOpt` package [Walesiak, Dudek 2017b] of R program applies the `smacofSym` function of the `smacof` package [Mair et al. 2017].

Finally, as a result of the optimal multidimensional scaling procedure, the application the data matrix in two-dimensional space $[v_{ij}]_{(2n+2)x2}$ is developed.

5) Depending on the position of pattern and anti-pattern in two-dimensional scaling space $[v_{ij}]_{(2n+2)x2}$ the rotation of coordinate system is required by φ angle in line with the following formula [Bronsztejn et al. 2004, p. 206]:

(4)

where: $[v'_{ij}]_{(2n+2)x2}$ – data matrix in two-dimensional scaling space after the rotation of coordinate system by φ angle,

$$D = \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix} \text{ – rotation matrix.}$$

6) Graphic presentation and interpretation of the results in two-dimensional space (multidimensional scaling results) and one-dimensional space (linear ordering results):

- two points, standing for anti-pattern and pattern, are joined by a straight line into the so-called set axis on the picture in two-dimensional space (multidimensional scaling results). Isoquants of development (curves of equal development) are determined from the pattern point. The objects between isoquants present the similar level of development. The same development level can be achieved by the objects located in different points along the same isoquant of development (due to a different configuration of variable values). Owing to such presentation of the results, the interpretation of linear ordering results is more extensive.
- d_i^+ composite measure values are calculated following the formula below (cf. [Hellwig 1981, p. 62]):

$$d_i^+ = 1 - \frac{\sqrt{\sum_{j=1}^{2}\left(v_{ij} - v_{+j}\right)^2}}{\sqrt{\sum_{j=1}^{2}\left(v_{+j} - v_{-j}\right)^2}}, \ d_i^+ \in [0; 1], \tag{5}$$

where: $\sqrt{\sum_{j=1}^{2}\left(v_{ij} - v_{+j}\right)^2}$ – Euclidean distance of *i*-th object to the pattern object (upper pole of development),

$\sqrt{\sum_{j=1}^{2}\left(v_{+j} - v_{-j}\right)^2}$ – Euclidean distance of the pattern object (upper pole of development) from the anti-pattern object (lower pole of development).

The higher d_i^+ value, the higher the social cohesion level of the analysed objects. The analysed objects are ordered by the descending values of composite measure (5).

The Theil measure was used for the assessment of changes in the level and degree of social cohesion differentiation of the Lower Silesia region in the cross-section of counties in the years 2005-2015.

The Theil measure was calculated for the comparable values of composite measures in formula (6) from the years 2005 (d_{iq}^+) and 2015 (d_{it}^+), which measures not only the range of standard deviations from the values of the comparable d_{it}^+ and d_{iq}^+ composite measures, but also the range of deviations resulting from [Walesiak 1993; 2016a, pp. 89-90]:

1) the difference between the mean values of d_{it}^{+} and d_{iq}^{+}, composite measures,

2) the difference in the dispersion of d_{it}^{+} and d_{iq}^{+} composite measure values,

3) the inconsistency in the direction of changes of d_{it}^{+} and d_{iq}^{+} composite measure values.

The Theil measure takes the following form:

$$W_{tq}^{2} = \frac{1}{n}\sum_{i=1}^{n}\left(d_{it}^{+} - d_{iq}^{+}\right)^{2}. \tag{6}$$

W_{tq}^{2} measure takes 0 value when there are no differences in the values of d_{it}^{+} and d_{iq}^{+} composite measures. The square root of formula (6) informs about the average range of deviations for the comparable values of d_{it}^{+} and d_{iq}^{+} composite measures.

The value expressed by formula (6) can be divided into the sum of three components:

$$W_{tq}^{2} = W_{1}^{2} + W_{2}^{2} + W_{3}^{2}, \tag{7}$$

allowing to define more specifically the "range" and "nature" of differences in the values of d_{it}^{+} and d_{iq}^{+} composite measures.

Partial measures W_{1}^{2}, W_{2}^{2} and W_{3}^{2} (carrying information listed in points 1), 2) and 3)) are presented by the following formulas:

$$W_{1}^{2} = \left(\bar{d}_{\bullet t}^{+} - \bar{d}_{\bullet q}^{+}\right)^{2}, \tag{8}$$

$$W_{2}^{2} = \left(S_{t} - S_{q}\right)^{2}, \tag{9}$$

$$W_{3}^{2} = 2S_{t}S_{q}(1 - r), \tag{10}$$

where: $\bar{d}_{\bullet t}^{+}$, $S_{t}\left(\bar{d}_{\bullet q}^{+}, S_{q}\right)$ stands, respectively, for the arithmetic mean and standard deviation of t-th (q-th) value of d_{it}^{+} and d_{iq}^{+} composite measure; r – Pearson's linear correlation coefficient between $\underline{d}_{\bullet t}^{+} = \left(d_{1t}^{+}, \ldots, d_{nt}^{+}\right)$ and $\underline{d}_{\bullet q}^{+} = \left(d_{1q}^{+}, \ldots, d_{nq}^{+}\right)$.

The decomposition of formula (7) into three components was taken from Theil's decomposition of the *MSE* [Theil 1961][4].

[4] In Polish subject literature see e.g. [Pawłowski 1973, p. 119; Zeliaś 1984, p. 184].

3. Empirical research results

According to Bernard (see e.g. [Dickes, Valentova 2013, p. 829]) social cohesion is analysed in three spheres: economic, political and socio-cultural. The conducted study did not cover the political sphere due to unavailability of data.

Three areas crucial for social cohesion were taken into account in the analysis: income and economic activity of the population, living conditions of the population and the availability of services and public space. The variables reflect the main aspects of prosperity mentioned in the definition of social cohesion of the Council of Europe.

The assessment of social cohesion in the Lower Silesia region in the cross-section of Lower Silesia counties was carried out using 28 metric variables measured using a ratio scale [Walesiak, Obrębalski 2017]:

I. Income and economic activity of the population:

x1 – Average gross monthly salary (PLN) – stimulant, data for 2005 and 2015.

x2 – People in households (below income criterion) benefiting from social assistance per 1000 population – destimulant, data for 2009 and 2015.

x3 – Demographic burden index (non-productive population per 100 working-age population) – destimulant, data for 2005 and 2015.

x4 – % share of women in working population – nominant (50% nominal value), data for 2005 and 2014.

x5 – % total unemployment rate – destimulant, data for 2005 and 2015.

x6 – % share of young people (aged up to 25) among the registered unemployed – destimulant, data for 2005 and 2015.

x7 – % share of the long-term unemployed (i.e. longer than 12 months) among the registered unemployed – destimulant, data for 2005 and 2015.

x8 – Job offers per 1000 registered unemployed – stimulant, data for 2005 and 2015.

II. Living conditions of the population:

x9 – Average usable floor space per capita in m^2 – stimulant, data for 2005 and 2015.

x10 – Average number of people per room – destimulant, data for 2005 and 2015.

x11 – Percentage of total dwellings equipped with water supply installations – stimulant, data for 2005 and 2014.

x12 – Percentage of total dwellings equipped with a bathroom – stimulant, data for 2005 and 2014.

x13 – Percentage of total dwellings equipped with central heating – stimulant, data for 2005 and 2014.

III. Availability of services and public space:

x14 – Doctors and dentists per 10 thous. population – stimulant, data for 2006 and 2014.

x15 – Outpatient clinics per 10 thous. population – stimulant, data for 2005 and 2015.

x16 – Population per public pharmacy – destimulant, data for 2005 and 2015.

x17 – Places in stationary social welfare per 10 thous. population – stimulant, data for 2005 and 2015.

x18 – Children in pre-school facilities per 1000 children aged 3-5 – stimulant, data for 2005 and 2015.

x19 – Students in elementary schools for children and adolescents per division (class) – destimulant, data for 2005 and 2015.

x20 – Students in middle schools for children and adolescents per division (class) – destimulant, data for 2005 and 2015.

x21 – Students as compulsory English learners in primary and middle schools for children and adolescents (% of total students) – stimulant, data for 2008 and 2015.

x22 – Students in secondary schools for adolescents per division (class) – destimulant, data for 2005 and 2015.

x23 – People practising in sports clubs per 1000 population – stimulant, data for 2006 and 2014.

x24 – Public library books collection per 1000 population (vol.) – stimulant, data for 2005 and 2015.

x25 – Participants of cultural events (organized by centres, homes, culture centres, clubs and community centres) per 1000 population – stimulant, data for 2007 and 2015.

x26 – Area of public green spaces (parks, greenery and residential green areas) per 10 thous. population (ha) – stimulant, data for 2005 and 2015.

x27 – Length of municipal and country roads with hard improved surface per 10 thous. population (km) – stimulant, data for 2005 and 2014.

x28 – Sewage treatment plant users – stimulant, data for 2005 and 2015.

The statistical data for the majority of variables were collected for the years 2005 and 2015 from the Local Data Bank (LDB) of the Central Statistical Office in Poland (GUS). Due to the absence of statistical data for ten variables the data available for the proximate years were used. In the period 2002-2012 Wałbrzych did not have the status of a city with county rights as it was one of the urban municipalities of Wałbrzych county. The Local Data Bank and the County Employment Office in Wałbrzych provided data for the majority of the analysed variables in 2005 for Wałbrzych and Wałbrzych county. There are gaps in data for the following variables: x1, x5, x14 and x28. In such a situation the data adopted for Wałbrzych were the same as the ones for Wałbrzych county.

In accordance with the research methodology presented in point 2 for $[x_{ij}]$ data matrix covering data matrices from t (matrix $[x_{ij}^t]$) and q (matrix $[x_{ij}^q]$) periods:

a) x4 nominant was changed into a stimulant in accordance with the differential formula [Walesiak 2016a, p. 19]:

$$x_{ij} = -\left|x_{ij}^N - nom_j\right|, \tag{11}$$

where: x_{ij}^N – the value of j-th nominant observed in i-th object; nom_j – nominal level of j-th variable;

b) the joint pattern and anti-pattern of development was defined based on the matrix covering data from t and q periods, i.e. matrix $[x_{ij}]$,

c) when calculating distances between objects equal weights were adopted for sub-criteria (domains), but differentiated for the variables presented in Table 1.

Table 1. Weights for sub-criteria (domains) describing the level of social cohesion in the Lower Silesia region

Specification	Sub-criterion		
	income and economic activity of the population	living conditions of the population	availability of services and public space
Weights for domains	1/3	1/3	1/3
Number of variables	8	5	15
Weight for 1 variable within one domain	1/24	1/15	1/45

Source: [Walesiak, Obrębalski 2017].

The article uses the `mdsOpt` package of the R program [Walesiak, Dudek 2017b] allowing the choice of optimal multidimensional scaling procedure in accordance with the procedure presented in the study [Walesiak, Dudek 2017c].

The monograph authors [Borg, Groenen, Mair 2013] indicated in Chapter 7 that *Stress*-1 goodness-of-fit measure cannot remain the only choice criterion as it shows pooled error for all studied objects only. Apart from that, also the percentage shares of objects based on the value of *Stress*-1 (stress per point) goodness-of-fit measure and the interpretability of multidimensional scaling results should be taken into consideration.

Ten normalization methods (n1, n2, n3, n5, n5a, n8, n9, n9a, n11, n12a)[5], five distance measures (city-block, Euclidean, squared Euclidean, Chebyshev, GDM1) and two scaling models (ratio and interval MDS) were taken into account in the choice of an optimal scaling procedure.

Ten methods of variable values normalization, five distance measures and two scaling models applied in the analysis produce 100 multidimensional scaling procedures. Multidimensional scaling is carried out for each procedure separately. Next,

[5] Due to the fact that some normalization methods produce identical results (see [Walesiak 2016c]), 10 out of 18 methods were finally included in the study.

the procedures are arranged by the ascending values of *Stress*-1 goodness-of-fit measure (cf. e.g. [Borg, Groenen 2005, p. 42]):

$$Stress\text{-}1_p = \sqrt{\frac{\sum_{i,k}\left(\hat{d}_{ik} - d_{ik}\right)^2}{\sum_{i,k} d_{ik}^2}}, \tag{12}$$

where: $p = 1,...,100$ – the number of multidimensional scaling procedure.

Based on the percentage shares of objects in the value of *Stress*-1 (*spp* – stress per point) goodness-of-fit measure, the Hirschman-Herfindahl index is calculated [Herfindahl 1950; Hirschman 1964]:

$$HHI_p = \sum_{i=1}^{n} spp_{pi}^2, \tag{13}$$

where: $i = 1,..., n$ – object's number.

HHI_p index takes values in the interval $\left[\frac{10{,}000}{n}; 10{,}000\right]$. The value $\frac{10{,}000}{n}$ means that error distribution for individual objects is uniform ($\forall_i\, spp_i = \frac{100}{n}$). Maximum value occurs when the pooled error (*Stress*-1) results from an error for one object only. For other objects the error will equal zero. From the perspective of multidimensional scaling the lowest value of HHI_p index is desirable.

A graph is developed (see Figure 1) on which along the axis of abscissae the value of $Stress\text{-}1_p$ fit function is presented, while on the axis of ordinates the values of HHI_p index for $p = 100$ multidimensional scaling procedures are marked.

Among the acceptable multidimensional scaling procedures (for which $Stress\text{-}1_p \leq \text{median}(Stress\text{-}1)$) the one was selected for which $\min_p\{HHI_p\}$ occurs. This is the 40 procedure: n5 normalization method (normalization in the interval [–1; 1]), scaling model (interval), distance measure (GDM1).

Figure 2 (left panel) presents d_{ik} and $\hat{d}_{ik}$ residual plot ($R^2 = 0.9303$). Figure 2 (right panel) shows the Shepard diagram which confirms the correctness of the selected scaling model (Pearson's linear correlation coefficient $r = 0.9645$).

Finally, as the result of optimal multidimensional scaling procedure application the data matrix in two-dimensional space $[v_{ij}]_{(2n+2)x2}$ is developed.

The results of multidimensional scaling of 62 objects (30 Lower Silesia counties in 2005, 30 Lower Silesia counties in 2015, pattern and anti-pattern), in terms of social cohesion level, are presented on Figure 3. The rotation of coordinate system by $\varphi = \pi / 8$ was performed.

The anti-pattern (object 62) and pattern (object 61) were connected by a straight line into a so-called set axis. Six isoquants[6] of development were defined by dividing the set axis into six equal parts.

Next, the values of composite measure (5) were calculated. Table 2 presents the ordering of 30 counties in the years 2005 and 2015, in terms of social cohesion level, by descending values of (5) measure for 2015. The calculations were performed using the `clusterSim` package [Walesiak, Dudek 2017a] of the R program [R Core Team 2017].

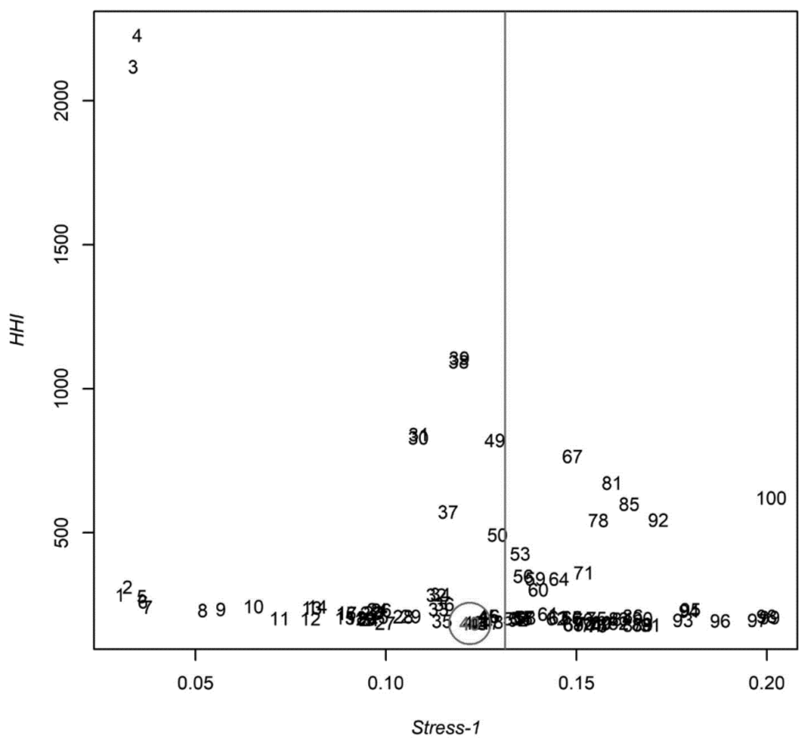
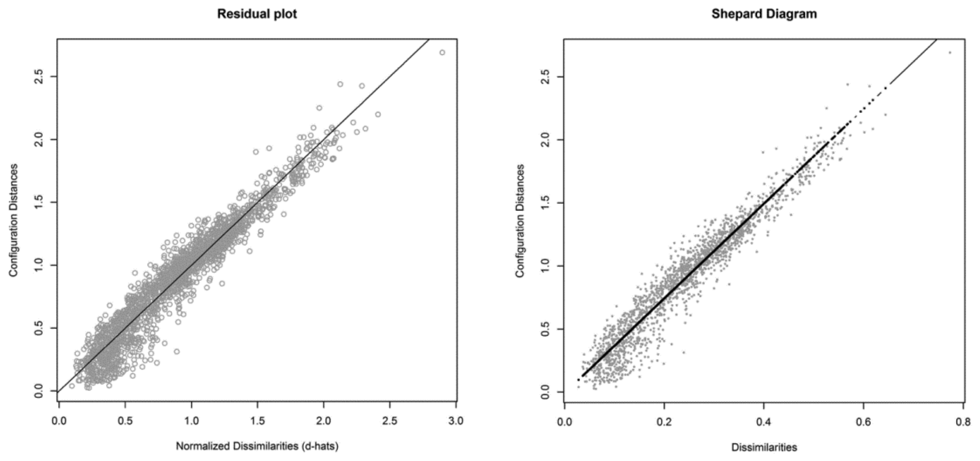


Fig. 1. $Stress\text{-}1_p$ fit function values and HHI_p index values for p multidimensional scaling procedures

Source: author's compilation using the R program.

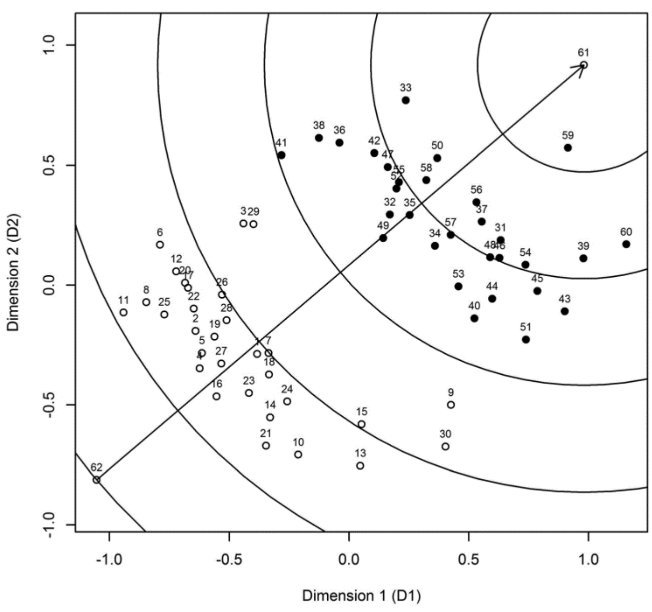
[6] The course of isoquants of development was visualized using `draw.circle` function of the `plotrix` package [Lemon et al. 2017].



(d_{ik} – Configuration Distances, $\hat{d}_{ik}$ – Normalized Dissimilarities (d-hats), δ_{ik} – Dissimilarities)

Fig. 2. Residual plot and Shepard diagram

Source: author's compilation using the R program.



Description: 61 – pattern, 62 – anti-pattern, 1-30 – numbers of objects in 2005, 31-60 numbers of objects in 2015, numbers of objects are explained in Table 2.

Fig. 3. Multidimensional scaling results of 62 objects (30 Lower Silesia counties in 2005, 30 Lower Silesia counties in 2015, pattern and anti-pattern) in terms of social cohesion level

Source: author's compilation using the R program.

Table 2. Ordering of Lower Silesia counties by the level of social cohesion (d_i^+ measure values) in 2005 and 2015

No. 2005/2015	County	2015		2005		$\Delta d_i^+ = d_{it}^+ - d_{iq}^+$	Position change
		d_{it}^+	No.	d_{iq}^+	No.		
29 / 59	Wrocławski	0.8683	1	0.4269	2	**0.4414**	1
20 / 50	Ząbkowicki	0.7286	2	0.2897	13	**0.4390**	**11**
26 /56	Średzki	0.7279	3	0.3302	7	0.3977	4
3 / 33	Jeleniogórski	0.7160	4	0.4130	3	0.3030	–1
30 /60	Wrocław	0.7124	5	0.3661	4	0.3463	–1
7 / 37	Zgorzelecki	0.7080	6	0.3326	6	0.3753	0
9 / 39	Jelenia Góra	0.6981	7	0.4300	1	0.2682	–6
1 /31	Bolesławiecki	0.6972	8	0.3183	8	0.3790	0
28 / 58	Wołowski	0.6950	9	0.3136	9	0.3814	0
24 / 54	Oławski	0.6751	10	0.2990	11	0.3761	1
16 / 46	Dzierżoniowski	0.6711	11	0.2268	27	**0.4443**	**16**
18 / 48	Świdnicki	0.6659	12	0.3102	10	0.3557	–2
27 / 57	Trzebnicki	0.6630	13	0.2661	20	0.3969	7
25 / 55	Strzeliński	0.6578	14	0.2371	25	**0.4207**	**11**
17 / 47	Kłodzki	0.6546	15	0.2899	12	0.3647	–3
22 / 52	Milicki	0.6492	16	0.2811	17	0.3681	1
12 / 42	Legnicki	0.6447	17	0.2856	14	0.3591	–3
5 / 35	Lubański	0.6407	18	0.2525	23	0.3882	5
15 / 45	Legnica	0.6398	19	0.3402	5	0.2996	–14
4 / 34	Kamiennogórski	0.6341	20	0.2349	26	0.3992	6
2 / 32	Jaworski	0.6171	21	0.2646	21	0.3526	0
13 / 43	Lubiński	0.6142	22	0.2831	16	0.3311	–6
14 / 44	Polkowicki	0.6081	23	0.2627	22	0.3453	–1
23 /53	Oleśnicki	0.6025	24	0.2674	19	0.3351	–5
6 / 36	Lwówecki	0.5988	25	0.2801	18	0.3187	–7
19 /49	Wałbrzyski	0.5858	26	0.2835	15	0.3023	–11
8 / 38	Złotoryjski	0.5702	27	0.2219	29	0.3483	2
10 / 40	Głogowski	0.5691	28	0.2450	24	0.3241	–4
21 / 51	Wałbrzych	0.5618	29	0.2250	28	0.3368	–1
11 / 41	Górowski	0.5067	30	0.1830	30	0.3237	0
Parameter		2015		2005		Increase	
Mean		0.6527		0.2920		0.3607	
Standard deviation		0.0669		0.0587		0.0082	
Median		0.6519		0.2833		0.3686	
Median absolute deviation		0.0661		0.0488		0.0173	

Source: author's calculations using `clusterSim` package of the R program.

By applying the script of the R program the results of the Theil measure decomposition were obtained, as presented in Table 3 (q stands for 2005 and t for 2015).

The average deviation range of d_i^+ comparable composite measure values in 2005 and 2015 (W Theil measure) was 0.3633. This primarily resulted from the mean value increase in d_i^+ composite measure ($W_1^2 = 0.13013$ partial measure), thus the significant increase in social cohesion level (the increase of mean composite measure value by 0.3607). There was a slight increase in the variation of d_i^+ composite measure value showing the increase in the degree of disproportions between counties in terms of social cohesion ($W_2^2 = 0.0000678$ for $S_t = 0.06689$ and $S_q = 0.05866$)[7]. Figure 4 presents the relationships between the level and degree of social cohesion disproportions for Lower Silesia counties in the years 2005-2015.

Table 3. The results of the Theil measure decomposition

Specification	$t = 2015$	$q = 2005$
Composite measure mean values	0.65274	0.29201
Standard deviations from composite measure values	0.06689478	0.05866117
Pearson correlation coefficient between composite measure values in 2005 and 2015	0.7719240	
W Theil measure value	0.3632959	
W^2 Theil measure value	0.1319839 (100.00%)	
W_1^2 Theil partial measure value	0.1301261 (98.59%)	
W_2^2 Theil partial measure value	0.00006779226 (0.05%)	
W_3^2 Theil partial measure value	0.001789998 (1.36%)	

Source: author's calculations using the R program.

The relatively high consistency in the direction of d_i^+ composite measure value changes against the comparable periods ($W_3^2 = 0.00179$ partial measure for $r = 0.7719$). The geometric interpretation of correlation coefficient is illustrated in Figure 5.

The d_i^+ measure values for the particular Lower Silesia counties in the analysed period were subject to diverse changes (see column $\Delta d_i^+ = d_{it}^+ - d_{iq}^+$ in Table 2). The relatively most extensive range of changes in the analysed areas of social cohesion, i.e. in terms of income and economic activity of the population, living conditions and the availability of services in public space refers to the following counties (bold-printed values in Table 2): Dzierżoniowski, Wrocławski, Ząbkowicki and Strzeliński, whereas the smallest range to the following cities with county rights (underlined values in Table 2): Jelenia Góra and Legnica as well as Wałbrzyski county.

[7] Similar indications are observed using median absolute deviation (see Table 2).

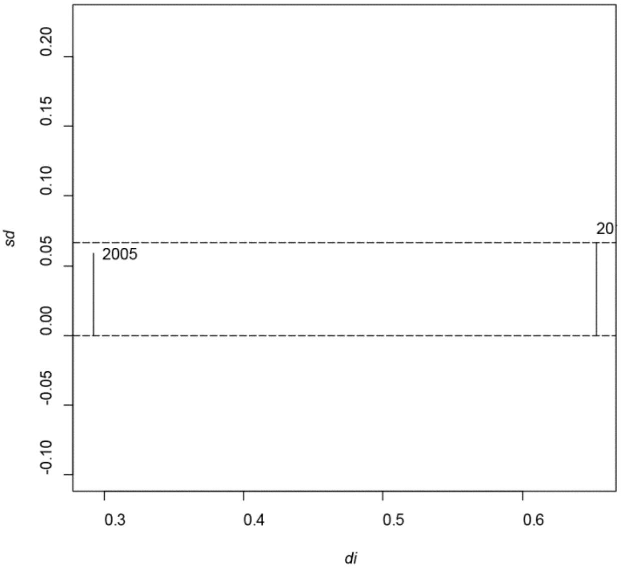


Fig. 4. Level (*di*) vs. degree of differentiation (*sd*) in social cohesion for Lower Silesia counties in the years 2005-2015

Source: author's compilation using the R program.

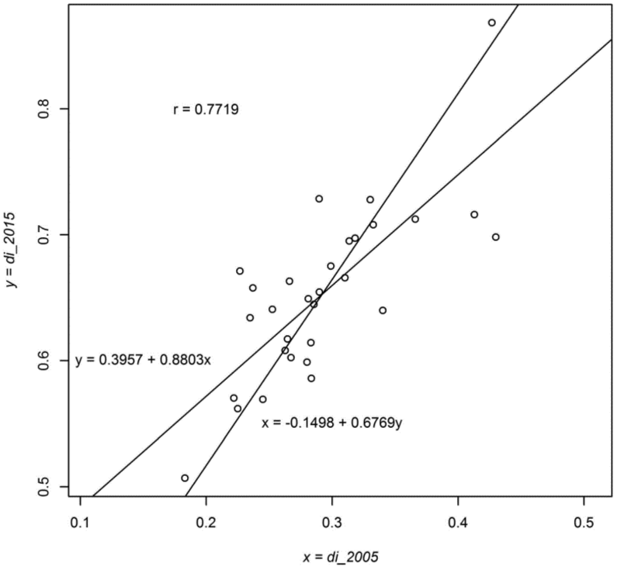


Fig. 5. Geometric interpretation of the correlation coefficient ($d^{+}_{i2015} = x = di\_2015$; $d^{+}_{i2005} = x = di\_2005$)

Source: author's compilation using the R program.

4. Final remarks

The article discusses the methodology allowing the measurement and assessment of changes in the level and degree of social cohesion differentiation in regions.

The article proposes a hybrid approach combining multidimensional scaling with linear ordering to measure and assess changes in the level of social cohesion of the Lower Silesia region in the period 2005-2015.

The main advantage of this two-step approach (first step – multidimensional scaling, second step – linear ordering), enriches the interpretation by the visualization of the results of linear ordering carried out based on composite measure (5).

On the other hand the main limitation of the presented approach is the use of multidimensional scaling in the first phase of the study. The presentation of the results in two-dimensional space causes partial loss of information. At the start point objects were described by 28 variables.

The overall assessment of social cohesion differentiation degree in Lower Silesia counties, in the period 2005-2015, was performed based on measure (6) with Theil's decomposition.

The level of social cohesion increased in all the counties in the years 2005-2015 (even though its degree varied) and the degree of disproportions between counties went up in terms of social cohesion.

The author's own scripts, prepared in R environment, were applied in the calculations.

Bibliography

Balcerzak A.P., 2015, *Wielowymiarowa analiza spójności społecznej w krajach Unii Europejskiej w kontekście strategii Europa 2020*, [in:] B. Bartniczak, K. Trzeciak (eds.), *Aktualne trendy w zarządzaniu środowiskiem*, AD REM Publishers, Jelenia Góra, pp. 343-352.

Borg I., Groenen P.J.F., 2005, *Modern Multidimensional Scaling. Theory and Applications*, 2nd Edition, Springer Science+Business Media, New York, http://www.springeronline.com/0-387-25150-2.

Borg I., Groenen P.J.F., Mair P., 2013, *Applied Multidimensional Scaling*, Springer, Heidelberg, New York, Dordrecht, London, URL http://dx.doi.org/10.1007/978-3-642-31848-1.

Borys T., 1984, *Kategoria jakości w statystycznej analizie porównawczej*, Prace Naukowe Akademii Ekonomicznej we Wrocławiu nr 284, Seria: Monografie i Opracowania nr 23, Wydawnictwo AE, Wrocław.

Bronsztejn I.N., Siemiendiajew K.A., Musiol G., Mühlig H., 2004, *Nowoczesne kompendium matematyki*, Wydawnictwo Naukowe PWN, Warszawa.

Chan J., To H-P., Chan E., 2006, *Reconsidering social cohesion: Developing a definition and analytical framework for empirical research*, Social Indicators Research, vol. 75, issue 2, pp. 273-302, https://doi.org/10.1007/s11205-005-2118-1.

Dickes P., Valentova M., 2013, *Construction, validation and application of the measurement of social cohesion in 47 European countries and regions*, Social Indicators Research, vol. 113, issue 3, pp. 827-846, https://doi.org/10.1007/s11205-012-0116-7.

Dickes P., Valentova M., Borsenberger M., 2010, *Construct validation and application of a common measure of social cohesion in 33 European countries*, Social Indicators Research, vol. 98, issue 3, pp. 451-473, https://dx.doi.org/10.1007/s11205-009-9551-5.

EU/EC/CE, 2004, *A New Strategy for Social Cohesion*, European Committee for Social Cohesion, Strasburg.
EU/EC/CE, 2005, *Concerted Development of Social Cohesion Indicators*, Methodological Guide, Council of Europe Publishing, Strasbourg.
Hellwig Z., 1968, *Zastosowanie metody taksonomicznej do typologicznego podziału krajów ze względu na poziom ich rozwoju i strukturę wykwalifikowanych kadr*, Przegląd Statystyczny, vol. 15, issue 4, pp. 307-327.
Hellwig Z., 1972, *Procedure of Evaluating High-Level Manpower Data and Typology of Countries by Means of the Taxonomic Method*, [in:] Gostkowski Z. (ed.), *Towards a system of Human Resources Indicators for Less Developed Countries*, Papers Prepared for UNESCO Research Project, Ossolineum, The Polish Academy of Sciences Press, Wrocław, pp. 115-134.
Hellwig Z., 1981, *Wielowymiarowa analiza porównawcza i jej zastosowanie w badaniach wielocechowych obiektów gospodarczych*, [in:] W. Welfe (ed.), *Metody i modele ekonomiczno-matematyczne w doskonaleniu zarządzania gospodarką socjalistyczną*, PWE, Warszawa, pp. 46-68.
Herfindahl O.C., 1950, *Concentration in the steel industry. Doctoral thesis*, Columbia University.
Hirschman A.O., 1964, *The Paternity of an Index*, The American Economic Review, vol. 54, no. 5, 761-762, URL http://www.jstor.org/stable/1818582.
Jenson J., 2010, *Defining and Measuring Social Cohesion*, Commonwealth Secretariat, London.
Lemon J. et al., 2017, *plotrix: Various Plotting Functions. R package version 3.6-5*, URL http://CRAN.R-project.org/package=plotrix.
Mair P., De Leeuw J., Borg I., Groenen P.J.F., 2017, *smacof: Multidimensional Scaling. R package version 1.9-6*, URL http://CRAN.R-project.org/package=smacof.
Pawłowski Z., 1973, *Prognozy ekonometryczne*, PWN, Warszawa.
R Core Team, 2017, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, URL http://www.R-project.org.
Rajulton F., Ravanera Z.R., Beaujot R., 2007, *Measuring social cohesion: An experiment using the Canadian national survey of giving, volunteering, and participating*, Social Indicators Research, 80, pp. 461-492, http://dx.doi.org/10.1007/s11205-006-0011-1.
Ryszkiewicz A., 2013, *Od konwergencji do spójności i efektywności – podstawy teoretyczne polityki spójności gospodarczej, społecznej i terytorialnej Unii Europejskiej*, Oficyna Wydawnicza SGH, Warszawa.
Theil H., 1961, *Economic Forecasts and Policy*, North Holland, Amsterdam.
Walesiak M., 1993, *Zagadnienie oceny podobieństwa zbioru obiektów w czasie w syntetycznych badaniach porównawczych*, Przegląd Statystyczny, vol. 40, issue 1, pp. 95-102.
Walesiak M., 2014, *Przegląd formuł normalizacji wartości zmiennych oraz ich własności w statystycznej analizie wielowymiarowej*, Przegląd Statystyczny, vol. 61, issue 4, pp. 363-372.
Walesiak M., 2016a, *Uogólniona miara odległości GDM w statystycznej analizie wielowymiarowej z wykorzystaniem programu R*, Wrocław University of Economics Press, Wrocław.
Walesiak M., 2016b, *Visualization of linear ordering results for metric data with the application of multidimensional scaling*, Ekonometria 2(52), pp. 9-21, http://dx.doi.org/10.15611/ekt.2016.2.01.
Walesiak M., 2016c, *Wybór grup metod normalizacji wartości zmiennych w skalowaniu wielowymiarowym*, Przegląd Statystyczny, vol. 63, issue 1, pp. 7-18.
Walesiak M., Dudek A., 2017a, *clusterSim: Searching for Optimal Clustering Procedure for a Data Set. R package version 0.46-2*, URL http://CRAN.R-project.org/package=clusterSim.
Walesiak M., Dudek A., 2017b, *mdsOpt: Searching for Optimal MDS Procedure for Metric Data. R package version 0.1-4*, URL http://CRAN.R-project.org/package=mdsOpt.
Walesiak M., Dudek A., 2017c, *Selecting the Optimal Multidimensional Scaling Procedure for Metric Data with R Environment*, Statistics in Transition – new series, vol. 18, no. 3, pp. 521-540. http://stat.gov.pl/en/sit-en/issues-and-articles-sit/.
Walesiak M., Obrębalski M., 2017, *Spójność społeczna województwa dolnośląskiego w latach 2005--2015 – pomiar i ocena zmian*, Przegląd Statystyczny, vol. 64, issue 4, pp. 437-455.
Zeliaś A., 1984, *Teoria prognozy*, PWE, Warszawa.