

Article

The Measurement of Social Cohesion at Province Level in Poland Using Metric and Interval-Valued Data

Marek Walesiak ¹ and Grażyna Dehnel ^{2,*}

¹ Department of Econometrics and Computer Science, Wrocław University of Economics and Business, 53-345 Wrocław, Poland; marek.walesiak@ue.wroc.pl

² Department of Statistics, Poznań University of Economics and Business, 61-875 Poznań, Poland

* Correspondence: grazyna.dehnel@ue.poznan.pl

Received: 14 July 2020; Accepted: 13 September 2020; Published: 16 September 2020



Abstract: The notion of social cohesion is increasingly used in the political, economic and academic debate. Due to its multidimensional, the assessment of social cohesion is not easy, especially if it is conducted at a lower than national level of aggregation. The aim of the study is to assess social cohesion in provinces of Poland in 2018 using the hybrid approach involving multidimensional scaling and linear ordering based on an aggregate measure. This type of study is usually conducted using classic metric data. However, the traditional approach does not account for the variation between lower level units (i.e., districts). The authors propose a methodology which makes this possible. Additionally the results of assessment of the multidimensional phenomenon can be presented in a two-dimensional space. Classic metric data and symbolic interval-valued data (three data types: min-max, 1st decile and 9th decile, 2nd decile and 8th decile) are jointly represented in a single diagram. The consistency of the research method ensures comparability of results of linear ordering. Two criteria were used in the comparative analysis of four rankings of social cohesion. The results of the study clearly showed that the current level of social cohesion at the provinces level is geographically and historically dependent.

Keywords: social cohesion; interval-valued data; symbolic data; multidimensional scaling; composite indicators

1. Introduction and Motivation

The purpose of the study described in the article was to assess social cohesion of the provinces of Poland in 2018. The study was based on classic metric data and symbolic interval-valued data. Interval-valued variables describe objects of interest more precisely than classical metric variables. For classical metric data an observation on each variable in a data matrix is expressed as one real number (atomic approach). In contrast, for symbolic interval-valued data, observations on each variable are expressed as intervals $[x_{ij}^l, x_{ij}^u]$ ($x_{ij}^l \leq x_{ij}^u$), where x_{ij}^l denotes the lower bound and x_{ij}^u the upper bound of the interval. Studies by [1,2] provide different examples of data that in real life are of interval type.

Social cohesion is often measured by means of various composite indicators. Duhaime et al. [3] make an attempt to measure the level of social cohesion in the Canadian Arctic using six sets of indices: presence of social capital, demographic stability, social and economic inclusion, community quality of life, individual quality of life. Based on the definition of social cohesion by Bernard [4] and Chan et al. [5], the VALCOS (VALeurs et COhéSion Sociale) index of social cohesion elaborated for European countries was developed [6]. It covers the political and socio-cultural domains of life in their

formal and substantial relations. Langer et al. [7] developed two social cohesion indices: a national average SCI and a Social Cohesion Index Variance-Adjusted (SCIVA) to measure the national-level of social cohesion for 19 African countries in 2005, 2008 and 2012. The indices represent social cohesion as a triangle composed of the three components of societal relationships and attitudes: inequalities, trust, and identities. In the Polish literature, e.g., [8] analysed social cohesion in EU countries using a synthetic measure of development put forward by Z. Hellwig [9]. There are a number of other publications that have proposed composite indicators of social cohesion (i.e., [10,11]).

In some studies the index-based analysis is extended by the inclusion of additional statistical methods. Janmaat [12] considered 14 indicators describing 8 components of social cohesion and used exploratory factor analysis (EFA) to map 41 countries in the world in a two-dimensional space showing the relationship between two factors (solidarity, participation). In another study, Bottoni [13] used 24 indicators describing 7 dimensions of social cohesion for 29 European countries to build a multilevel CFA (confirmatory factor analysis) model of social cohesion. Dickes and Valentova [14] used multidimensional scaling (MDS) and a CFA model of social cohesion to map 47 European countries a two-dimensional MDS space. The results served as the basis for comparing levels of social cohesion in six broader geographical regions in Europe. An analogical study involving the same methods (MDS, CFA) for 33 European countries was described in Dickes et al. [15]. Using factor analysis and standardisation, Rajulton et al. [16] created an overall index of social cohesion across 49 Census Metropolitan Areas of Canada based on three dimensions: political (voting and volunteering), economic (occupation, income, labour force participation) and social (social interactions, informal volunteering). Lafuente et al. [17] make an attempt to assess the sustainability of social cohesion in the EU using nonlinear time-varying factor model and analyzing the level of convergence across EU countries.

In Dehnel et al. [18] the level of social cohesion was assessed using a hybrid approach combining multidimensional scaling and linear ordering, applied separately to classic metric data and symbolic interval-valued data (1st and 3rd quartile); the results were then compared.

The novelty of the study presented in this article consists in jointly mapping, by means of multidimensional scaling, classic metric data and symbolic interval-valued data (three data types: min-max, 1st decile and 9th decile, 2nd decile and 8th decile) in one chart. In the next step, all objects (districts) were ordered according to the level of social cohesion determined by means of an aggregate measure (composite indicator) based on the Euclidean distance from the pattern object. The application of a consistent research method ensures comparability of rankings of Poland's provinces in terms of social cohesion. In the following step assessments of social cohesion based on 4 different types of data were compared using two criteria: results of cluster analysis involving a distance based on two correlation coefficients (Spearman's ρ and Kendall's τ) and the analysis of the degree of compatibility between rankings of districts based on individual variables and the overall ranking based on the aggregate measure.

It is worth noting that existing studies aimed at assessing the level of social cohesion are usually based on primary data, from a surveys. The use of these sources are associated with a certain limitation. Namely primary data do not allow for the inclusion for assessing the level of social cohesion in higher level territorial units (e.g., provinces) data for lower-level units (districts). Only secondary sources can provide this possibility by introducing interval-valued data into the analysis. The empirical study described in this article was based on secondary data from official statistical sources. They were obtained from the Local Data Bank (BDL) using the bdl R package [19] and API interface (Application Programming Interface). BDL is the Polish acronym of the Local Data Bank (Bank Danych Lokalnych).

2. Overview of Social Cohesion Concepts

Social cohesion is a concept which is frequently mentioned in various projects and analyses, both in research and in government policies. The measurement of social cohesion and comparative analyses of its level in different territorial units are far from easy. This is because no clear definition or conceptualization of social cohesion has been proposed so far. There is still no consensus on

how to define values and factors related to the construct of social cohesion. The complexity of the concept of social cohesion can be illustrated by the variety of approaches that can be found in the literature [4,5,7,15,20–27]. All of them refer, to a varying degree, to six dimensions: social relations, identification, orientation towards the common good, shared values, quality of life, and (in)equality, though it is often indicated that the last three are antecedents or consequences of social cohesion rather than its core dimensions [28]. Although various approaches involve different areas of social cohesion, refer to different political views, are informed by different ideologies or concerns of policy makers, the majority of them overlap, covering similar dimensions [29] (p. 23). Nonetheless, each of the approaches proposed in the literature can be viewed as representing one of two discourses: the academic or the policy discourse [5].

The first category academic discourse is closely connected with social sciences, such as sociology and psychology. Studies referring to the academic discourse focus on processes of social integration and stability and social exclusion, while ignoring dilemmas associated with the definition of social cohesion [30–32].

The second category policy discourse refers to policies undertaken by governments and various national and international institutions (the European Union, Council of Europe, World Bank or OECD). In this case, social cohesion is viewed as a prerequisite of economic well-being. This goal can be achieved by overcoming numerous economic and social problems resulting from unequal income distribution, employment, housing issues, limited access to health care and education, participation in political and public life. Thus policy discourse can be described as problem-driven [5,15]. The list of major problems includes unemployment, poverty and social exclusion.

Policy-oriented studies are initiated and conducted by many national and international socio-political entities (governments, think tanks, foundations, organizations). Policy-makers usually address problems of social cohesion with a focus on their own concerns and from particular policy fields. Sometimes they even try to use the social cohesion term to promote their own agendas. As a result, policy-orientated analyses need to be treated with caution. For one thing, such studies often include too many social indicators, many of which do not properly capture components of social cohesion. Moreover, because territorial units face a complex mixture of socio-economic problems, every entity conducting research tends to create its own definition of social cohesion. Another problem is associated with the lack of distinction between factors affecting social cohesion and its actual components [5,13,15,28].

Both types of discourse on social cohesion described above are multi-faceted, but in each case the concept of social cohesion is different. The academic discourse is focused on a conceptual and analytic understanding of social cohesion. The policy discourse is rather social and economic problem-oriented. This is the main cause of existing discrepancies in measurement, and consequently in the assessment of the level of social cohesion. An additional difficulty of the academic approach is that it is largely based on subjective, qualitative assessments. The need to standardize the measurement method used to monitor the level and changes in social cohesion across time and societies is not new. However, the development of universal, assessable indicators would require the establishment of a single definitional framework of social cohesion. It should be based on the core elements of existing approaches to social cohesion presented and systematically mentioned in the literature of the past decades. Most approaches proposed so far refer to the policy discourse, which focuses on all types of social challenges faced by society [5,15,32]. A synthetic review of key approaches representing both the policy and academic discourse is presented below.

One of the first and most frequently cited concepts of social cohesion was formulated by Jane Jenson and was based on results of her own studies [20]. It represents the perspective found in the policy discourse. Rather than being a single definition, it consists of five dimensions:

1. Belonging vs. isolation—which refers to the existence or lack of shared values and a sense of identity,
2. Inclusion vs. exclusion—which refers to equal opportunities and citizens' access to economic institutions and the market,
3. Participation vs. non-involvement—which refers to political and social participation at various levels of government, especially the local level,
4. Recognition vs. rejection—which refers to respect and tolerance of diversity in a diverse society,
5. Legitimacy vs. illegitimacy—which addresses the question of respect for existing social norms and laws, the legitimacy of the main political and social institutions, especially the state, as mediators between different stakeholders.

Bernard [4] developed Jenson's definition by adding another dimension—equality vs. inequality—as an important element of the economic sphere. He introduced a typology based on two aspects. The first one comprised spheres of activity undertaken by society (economic, political and socio-cultural, while the second referred to social relations in the strict sense [6]).

In 2002, Beauvais and Jenson [22], drawing on Jenson's pluralistic approach [20], put forward a definition of social cohesion consisting of five elements: 1. Common values and a civic culture, 2. Social order and social control, 3. Social solidarity and reduction in wealth disparities, 4. Social networks and social capital, 5. Place attachment and identity [5]. The authors indicate that the way social cohesion is defined should depend on the main aspects investigated by the researcher, on a specific policy and should also address the most urgent social problems faced by territorial units, such as unemployment, poverty, discrimination, exclusion or any other problems that the researcher deems relevant [5].

In the literature a lot of attention has been paid to concepts of social cohesion are defined in terms of instruments enabling the achievement of cohesion, which is known as the means-end approach. In this case, the definition contains conditions that need to be met in order to ensure social cohesion [15]. Such concepts, representing the political approach, have been proposed, among others, by Berger-Schmitt [33,34] and Noll [35]. They demonstrated that social cohesion consists of two analytically different dimensions: inequality and social capital. The first refers to the question of promoting equal opportunities and reducing disparities and inequalities. The goal of the second one involves the strengthening of social relations, interactions and ties and comprises all aspects generally regarded as social capital [34]. The concepts have been criticized for being based on the definition of social cohesion defined in terms of conditions that can foster its development. An alternative proposal was advanced by Duhaime [5], who identified two components of social cohesion: access to formal economic and governmental institutions and access to family and community-based, face-to-face relations. Indicators identified in this concept (in line with the means-end approach), in many cases made reference to conditions favouring the development of social cohesion [3].

Another concept worth mentioning was formulated in 2006 by Chan et al. [5]. Following Bernard's definition, social cohesion was described as “a state of affairs concerning both vertical and the horizontal interactions among members of society as characterized by a set of attitudes and norms that includes trust, sense of belongingness and the willingness to participate and help as well as their behavioural manifestations” [5]. While retaining the political and sociocultural sphere of social cohesion, the authors decided to exclude the economic dimension, arguing for a minimalist definition which ignores all characteristics regarded as factors or determinants of social cohesions, such as equality of opportunities, equality and social integration [6]. It should be emphasized that both Bernard [4] and Chan et al. [5] assert that social cohesion is a property of a group or society, not an individual. This means that even if it is measured at individual level, ultimately such data are aggregated and social cohesion is described at the level of different groups, regions or communities [15].

Following the studies of Bernard [4] and Chan et al. [5], Dickes et al. [15] and Dickes and Valentova [14] proposed their own definitions of social cohesion, which also did not account for the economic dimension. Four elements of social cohesion were distinguished: institutional trust (i.e., legitimacy vs. illegitimacy), solidarity and concern for the common good (i.e., acceptance vs. rejection), political participation and socio-cultural participation.

When one analyses the conceptualisation of social cohesion, it is possible to track the direction of changes which reflect the increasing role of socio-cultural and political indicators and the omission of the economic sphere. In contrast, a similar review of social cohesion research reveals that the emphasis is shifting towards spatial analysis of social cohesion, which takes into account not only the national but also regional and local level [36–39]. Moreover, a given society's level of social cohesion can only be properly assessed when it can be compared across territorial units and over time. This can be achieved using methods of measuring social cohesion applied in the EU or proposed by OECD (Organisation for Economic Cooperation and Development). In the case of studies relating to countries and regions of the European Union, the EU regional Social Progress Index (EU-SPI) has been used since 2016 (https://ec.europa.eu/regional_policy/en/information/maps/social_progress). For purposes of the index, social progress is defined as “a society's capacity to meet the basic human needs of its citizens, to establish the basis for people and communities to improve and sustain their quality of life and to create the conditions for people to reach their full potential” [40] (p. 91). The EU-SPI is calculated on the basis of variables representing the socio-cultural and political dimension. The economic dimension is deliberately excluded, which facilitates the assessment of the level of social cohesion, as economic indicators make it difficult to distinguish between causes and effects of this level in the final analysis. The EU-SPI is consistent with the overall framework of the global Social Progress Index and is based on fifty indicators, primarily from Eurostat. It covers three dimensions of social progress [41]:

- basic human needs (nutrition and basic medical care, water and sanitation, shelter housing, personal safety),
- foundations of well-being (access to basic knowledge, access to information and communication, health and wellness, environmental quality),
- opportunity (personal rights, personal freedom and choice, tolerance and inclusion, access to advanced education).

Given the nature of statistical (symbolic interval-valued) data used in the analysis, the empirical research described in the article was based on the approach adopted in studies conducted by EU countries using the EU-SPI. In our study this concept was applied at a lower level of spatial aggregation, namely at province level. These territorial units (Pol. województwo) vary considerably in economic and social terms, reflecting different historical developments in three parts of Poland annexed by three neighbouring countries in the 18th century. As a result, western provinces are generally characterised by a higher level of economic development than the ones in the eastern part of the country. The purpose of the analysis was to determine in this differentiation is reflected in the assessment of the level of social cohesion.

3. Research Methodology

The assessment of social cohesion in the provinces of Poland was performed using four datasets: classic metric data and interval-valued data (three types: min-max, 1st and 9th deciles, 2nd and 8th deciles) by applying a hybrid approach involving multidimensional scaling and linear ordering. Multidimensional scaling made it possible to map 4 datasets describing 16 provinces onto a two-dimensional space; and then results of linear ordering were used to compare rankings of provinces in terms of social cohesion.

The research methodology is a modified approach proposed by Walesiak [42] and Dehnel et al. [18]. The research procedure, which makes it possible to present four types of data in one study, consists of the following steps:

1. Select a complex phenomenon that cannot be measured directly. In this study, it was the level of social cohesion.
2. Select a set of objects and a set of variables closely related with the complex phenomenon of interest. The study involves measuring characteristics of n objects ($i, k = 1, \dots, n; i, k$ —object number) described by means of m variables ($j = 1, \dots, m; j$ —variable number). Collected information

- comprises classic metric data and three types of interval-valued data (min-max, 1st and 9th deciles, 2nd and 8th deciles). Metric data converted into interval-valued data are arranged in the form of data table $\mathbf{X}_1 = \left[x_{ij}^l, x_{ij}^u \right]_{n \times m}$ ($x_{ij}^l = x_{ij}^u$; $i, k = 1, \dots, n$; $j = 1, \dots, m$). The three types of interval-valued data min-max, 1st and 9th deciles, 2nd and 8th deciles are arranged in data tables $\mathbf{X}_2 = \left[x_{ij}^l, x_{ij}^u \right]_{n \times m}$ ($x_{ij}^l \leq x_{ij}^u$; $i, k = n + 1, \dots, 2n$; $j = 1, \dots, m$), $\mathbf{X}_3 = \left[x_{ij}^l, x_{ij}^u \right]_{n \times m}$ ($x_{ij}^l \leq x_{ij}^u$; $i, k = 2n + 1, \dots, 3n$; $j = 1, \dots, m$), $\mathbf{X}_4 = \left[x_{ij}^l, x_{ij}^u \right]_{n \times m}$ ($x_{ij}^l \leq x_{ij}^u$; $i, k = 3n + 1, \dots, 4n$; $j = 1, \dots, m$).
- Combine the data in the form of a single data table $\mathbf{X}' = \left[x_{ij}^l, x_{ij}^u \right]_{4n \times m}$ ($x_{ij}^l \leq x_{ij}^u$) containing data tables $\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3$ and $\mathbf{X}_4$.
 - Add a pattern and anti-pattern object to the set of objects. Variables of interest can be divided into three types of preference variables: stimulants (where higher values are preferred), destimulants (where lower values are preferred), nominants (where the preferred value lies somewhere within the variable range). Formal definitions of stimulants, destimulants can be found in [43] (p. 48) while nominants are defined in [44] (p. 118). These definitions are also provided in [42]. Owing to the structure of the anti-pattern object nominants need to be converted into stimulants. Coordinates of the pattern object represent the most favourable values of preference variables (maximum values for stimulants and minimum values for destimulants). Coordinates of the anti-pattern object represent the least favourable values (minimum values for stimulants and maximum values for destimulants). In the case of symbolic interval-valued variables, coordinates area calculated separately for the lower and upper value of the interval. After including the pattern and anti-pattern object, the joint data table has the form $\mathbf{X} = \left[x_{ij}^l, x_{ij}^u \right]_{(4n+2) \times m}$ ($x_{ij}^l \leq x_{ij}^u$).
 - Normalise interval-valued variables and arrange the data in the form of a normalised data table $\mathbf{Z} = \left[z_{ij}^l, z_{ij}^u \right]_{(4n+2) \times m}$ ($z_{ij}^l \leq z_{ij}^u$; z_{ij} normalised observation) for symbolic interval-valued variables. The purpose of normalization is to ensure comparability of variables (cf. [45]). This is achieved by removing units from measurement results and standardizing their orders of magnitude. Symbolic interval-valued data require special normalization treatment. The lower and upper bound of the interval of the j -th variable for $4n + 2$ objects (n objects for 4 types of data, pattern and anti-pattern) are combined into one vector containing $2(4n + 2)$ observations. This approach makes it possible to apply normalization methods used for classic metric data. Metric data were normalized using the interval_normalization function from the clusterSim package implemented in the R program [46]. Normalisation methods can be represented by the following formula (cf. [47]):

$$z_{ij} = \frac{x_{ij} - A_j}{B_j} \quad (1)$$

where:

$x_{ij}(z_{ij})$ (normalised) value of j -th variable for i -th object,

A_j —shift parameter to arbitrary zero for j -th variable (see Table 1),

B_j —scaling parameter for the j -th variable (see Table 1).

Table 1. Selected normalisation methods.

Type	Method	Parameter	
		B_j	A_j
n1	Standardisation	s_j	$\bar{x}_j$
n2	Positional standardisation	mad_j	med_j
n3	Unitisation	r_j	$\bar{x}_j$
n3a	Positional unitisation	r_j	med_j
n5	Normalisation to $[-1; 1]$ range	$\max_i x_{ij} - \bar{x}_j $	$\bar{x}_j$
n5a	Positional normalisation to $[-1; 1]$ range	$\max_i x_{ij} - med_j $	med_j
n12a	Positional normalisation	$\sqrt{\sum_{i=1}^n (x_{ij} - med_j)^2}$	med_j

$\bar{x}_j$ —mean for j -th variable, s_j —standard deviation for j -th variable r_j —range for j -th variable, $med_j = med(x_{ij})$ —median for j -th variable, $mad_j = mad(x_{ij})$ —median absolute deviation for j -th variable.

6. Select a measure of distance for symbolic interval-valued data (see Table 2), calculate distances and arrange into a distance matrix $\delta = [\delta_{ik}(\mathbf{Z})]_{(4n+2) \times (4n+2)}$.

Table 2. Distance measures for symbolic interval-valued data.

Symbol	Name	Distance Measure $\delta_{ik}(\mathbf{Z})$
U_2_q1	Ichino-Yaguchi $q = 1, \gamma = 0.5$	$\sum_{j=1}^m \varphi(z_{ij}, z_{kj})$
U_2_q2	Euclidean Ichino-Yaguchi $q = 2, \gamma = 0.5$	$\sqrt{\sum_{j=1}^m \varphi(z_{ij}, z_{kj})^2}$
H_q1	Hausdorff $q = 1$	$\sum_{j=1}^m \left[\max(z_{ij}^l - z_{kj}^l , z_{ij}^u - z_{kj}^u) \right]$
H_q2	Euclidean Hausdorff $q = 2$	$\left\{ \sum_{j=1}^m \left[\max(z_{ij}^l - z_{kj}^l , z_{ij}^u - z_{kj}^u) \right]^2 \right\}^{1/2}$

$z_{ij} = [z_{ij}^l, z_{ij}^u]$; $\varphi(z_{ij}, z_{kj}) = |z_{ij} \oplus z_{kj}| - |z_{ij} \otimes z_{kj}| + \gamma(2 \cdot |z_{ij} \otimes z_{kj}| - |z_{ij}| - |z_{kj}|)$; $| \cdot |$ —interval length; $z_{ij} \oplus z_{kj} = z_{ij} \cup z_{kj}$; $z_{ij} \otimes z_{kj} = z_{ij} \cap z_{kj}$. Source: [48,49].

7. Conduct multidimensional scaling (MDS): $f : \delta_{ik}(\mathbf{Z}) \rightarrow d_{ik}(\mathbf{V})$ for all pairs (i, k) , where f denotes distance mapping from m -dimensional space $\delta_{ik}(\mathbf{Z})$ into corresponding distances $d_{ik}(\mathbf{V})$ in q -dimensional space ($q < m$). To enable graphic presentation of results q is set to 2. Distances $d_{ik}(\mathbf{V})$ are unknown. The iterative procedure, implemented in the smacof algorithm, used to find configuration $\mathbf{V}$ (given q dimensions) and calculate distance matrix $d_{ik}(\mathbf{V})$, is presented in [50] (pp. 204–205).

The solution used in the study makes it possible to select an optimal procedure of multidimensional scaling (MDS) for a given normalization method (Table 1), distance measure for symbolic interval-valued data (Table 2) and scaling models (ratio, interval, spline—polynomial function of second and third degree), available in the mdsOpt R package [46], which use the smacofSym function from the smacof package [51]. Two criteria were used to choose the optimal MDS procedure: the value of Kruskal's *STRESS*–1 goodness-of-fit function and Hirschman–Herfindahl Index (*HHI*), calculated for percentage shares of objects in the value of the *STRESS*–1 function (*stress per point*). Out of MDS procedures for which $STRESS-1_p \leq s$ (s —acceptable value of the goodness-of-fit measure), we select one for which $\min_p \{HHI_p\}$ (p —number of the MDS procedure). More information about the selection of the optimal MDS procedure can be found in the mdsOpt package vignette.

8. In the end, as a result of applying multidimensional scaling, we obtain a two-dimensional data matrix $\mathbf{V} = [v_{ij}]_{(4n+2) \times q}$ ($q = 2$). Depending on the location of the pattern and anti-pattern object in the two-dimensional scaling space the coordinate system needs to be rotated by an angle of φ according to the formula:

$$[v'_{ij}]_{(4n+2) \times 2} = [v_{ij}]_{(4n+2) \times 2} \times D \quad (2)$$

where:

$[v'_{ij}]_{(4n+2) \times 2}$ —data matrix in a two-dimensional scaling space after rotating the coordinate system by an angle of φ ,

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

The rotation does not change the arrangement of objects relative to one another but makes it possible to position the set axis connecting the pattern and anti-pattern along the identity line, which improves the visualization of results.

9. Visualise and interpret the results (of multidimensional scaling) in a two-dimensional space. This is done by first joining two points, representing the anti-pattern and pattern, by a straight line to form the so-called set axis in the diagram. Then isoquants of development (curves of equal development) are drawn from the pattern point. Objects located between the isoquants represent a similar level of development. The same level can be achieved by objects located at different points along the same isoquant of development (due to a different configuration of variable values).
10. Order objects according to the value of the aggregate measure d_i based on the Euclidean distance from the pattern object [43]:

$$d_i = 1 - \sqrt{\sum_{j=1}^2 (v_{ij} - v_{+j})^2} / \sqrt{\sum_{j=1}^2 (v_{+j} - v_{-j})^2} \quad (3)$$

where: v_{ij} — j -th coordinate for i -th object in the two-dimensional MDS space, v_{+j} (v_{-j})— j -th coordinate for the pattern (anti-pattern) object in the two-dimensional MDS space.

Values of the aggregate measure d_i belong to the interval $[0; 1]$. The higher the value of d_i , the higher the social cohesion of the objects. The objects are arranged according to descending values of the aggregate measure (3).

4. Data

Following the approach adopted in studies of social cohesion using the regional EU-SPI, three dimensions are considered: basic human needs, foundations of well-being and opportunities. Under this approach, the assessment of social cohesion in the provinces of Poland was carried out on the basis of 25 metric variables (see Table 3).

Variables $x_1, x_3, x_5, x_6, x_9, x_{17}, x_{23}$ and x_{24} represent stimulants (where higher values are more preferred), variables $x_2, x_4, x_7, x_8, x_{18}, x_{19}, x_{21}, x_{22}$ and x_{25} take the form of destimulants (where lower values are more preferred), and x_{20} is a nominant (with the nominal value of 50%). Statistical data for the 25 variables for 2018 come from the Local Data Bank maintained by Statistics Poland. The nominant variable x_{20}^N was converted into a stimulant according to the difference formula: $x_{20} = |x_{20}^N - 50|$.

Table 3. Variables applied in the assessment of social cohesion in the provinces of Poland.

Dimensions	Variables
Basic human needs	x1—mean gross monthly wage in PLN (Polish new zloty)
	x2—total unemployment rate in %
	x3—mean useful floor area of a dwelling per inhabitant in m ²
	x4—average number of persons per room
	x5—length of the sewerage network in relation to the length of the water supply network in %
	x6—number of doctors and dentists per 10,000 population
	x7—crimes reported (criminal offenses, against life and health, against property) per 10,000 population
	x8—road accidents per 100,000 population
Foundations of well-being	x9—users of water treatment services (% of total population)
	x10—percentage of all dwellings equipped with central heating
	x11—children enrolled in day-care centres per 1000 children up to the age of 3
	x12—children enrolled in nursery schools per 1000 children aged 3–5
	x13—students taking obligatory classes of English in primary and intermediate schools (% of all students)
	x14—members of sports clubs per 1,000 population
	x15—people participating in cultural events (organised by cultural centres and clubs) per 1000 population
	x16—area of public greenspace (parks, residential greenspace) per 10,000 population (in ha)
Opportunities	x17—length of municipal and district improved hard surface roads per 10,000 population (in km)
	x18—death rate among persons below the age of 60
	x19—dependency ratio (ratio of the dependent and elderly population per 100 working age population)
	x20 ^N —percentage share of women in the labour force
	x21—percentage share of young adults (up to the age of 25) among registered unemployed
	x22—percentage share of long-term unemployed (over 12 months) in the population of registered unemployed in %
	x23—places in stationary social welfare facilities per 10,000 population
	x24—voter turnout local elections (for municipal authorities and town councils with district rights) in 2018 in %
	x25—beneficiaries of social assistance at the place of residence (below the means test threshold) per 1000 population

5. Results of the Empirical Study

5.1. Results for Metric and Interval-Valued Data

In line with the procedure described in Section 3, four types of data—classic metric data and three types of symbolic interval-valued data (min-max, 1st and 9th deciles, 2nd and 8th deciles)—were mapped into a two-dimensional space and then rankings of provinces in Poland were compared in terms of social cohesion.

Data on social cohesion in the 16 provinces of Poland, described by 25 variables, were arranged in a data matrix $\mathbf{X}'_1 = [x_{ij}]_{16 \times 25}$. Because the data were to be combined with symbolic interval-valued data, they had to be put in a data table $\mathbf{X}_1 = [x^l_{ij}, x^u_{ij}]_{16 \times 25}$ ($x^l_{ij} = x^u_{ij}; i, k = 1, \dots, 16$).

Poland has a three-tier system of administrative division, consisting of 16 provinces (Pol. województwo), 380 districts (Pol. powiat) and 2477 communes (Pol. gmina). In order to obtain symbolic interval-valued data, classic metric data on social cohesion in 380 districts described by 25 variables were aggregated at province level. The lower and upper bound of the interval for each variable in each province was obtained by calculating the minimum and maximum, 1st and 9th deciles, 2nd and 8th deciles, using district-level data. Interval-valued data (min-max, 1st and 9th deciles, 2nd and 8th deciles) were arranged in three data tables $\mathbf{X}_2 = [x^l_{ij}, x^u_{ij}]_{16 \times 25}$ ($x^l_{ij} \leq x^u_{ij}; i, k = 17, \dots, 32$), $\mathbf{X}_3 = [x^l_{ij}, x^u_{ij}]_{16 \times 25}$ ($x^l_{ij} \leq x^u_{ij}; i, k = 33, \dots, 48$), $\mathbf{X}_4 = [x^l_{ij}, x^u_{ij}]_{16 \times 25}$ ($x^l_{ij} \leq x^u_{ij}; i, k = 49, \dots, 64$).

The data tables X_1 , X_2 , X_3 and X_4 were combined into one data table $X' = \begin{bmatrix} x_{ij}^l, x_{ij}^u \end{bmatrix}_{64 \times 25}$. After adding the pattern and anti-pattern object the final dataset was a data table $X = \begin{bmatrix} x_{ij}^l, x_{ij}^u \end{bmatrix}_{66 \times 25}$ ($x_{ij}^l \leq x_{ij}^u; i, k = 1, \dots, 66$).

The optimal scaling procedure was selected after testing combinations of seven normalization methods (n1, n2, n3, n3a, n5, n5a, n12a—see Table 1), four distance measures for interval-valued data (Ichino-Yaguchi, Euclidean Ichino-Yaguchi, Hausdorff, Euclidean Hausdorff—see Table 2) and four MDS models (ratio, interval, polynomial function of second and third degree), yielding a total of 112 MDS procedures. Values of Kruskal's $STRESS-1$ belong to the interval $[0.097572; 0.183817]$. Of the MDS procedures for which $STRESS - 1_p \leq 0.120478$ (acceptable value of the goodness-of-fit measure calculated as a median), we selected the combination (using the `optSmacofSymInterval` function from the `mdsOpt` R package) for which $\min_p \{HHI_p\} = 174.2$. This procedure involves normalisation n5 (normalisation to $[-1; 1]$ range), the scaling model based on polynomial function of the 3rd degree and the Euclidean Hausdorff distance. For this MDS procedure $STRESS - 1 = 0.120065$.

The Shepard diagram (Figure 1a) and the Stress Plot (Figure 1b) confirm the correctness of the selected multidimensional scaling procedure.

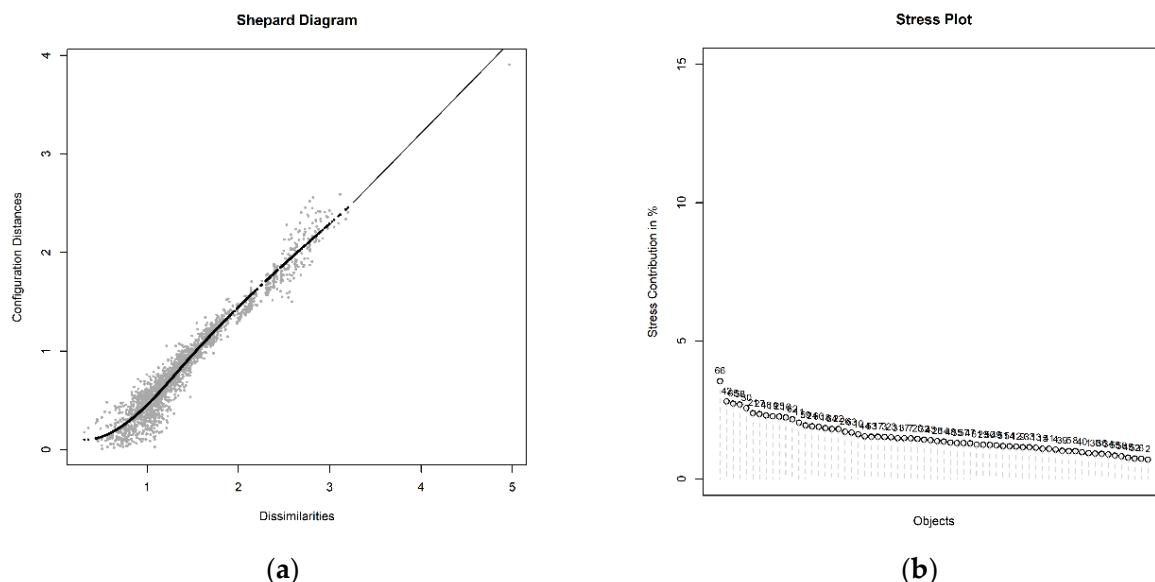


Figure 1. Shepard Diagram (a) and Stress Plot (b) of the optimal multidimensional scaling of 66 objects according to the level of social cohesion. Source: calculation and plot produced using the R program.

Figure 2 shows results of multidimensional scaling of 16 provinces of Poland for four types of data according to the level of social cohesion in 2018. In the diagram the anti-pattern (AP) object and the pattern (P) object are connected by a straight line, known as the set axis. Six isoquants of development (curves of equal development) were arbitrarily identified, which divided the set axis into six equal parts. Isoquants located further away from the pattern object represent a lower level of social cohesion.

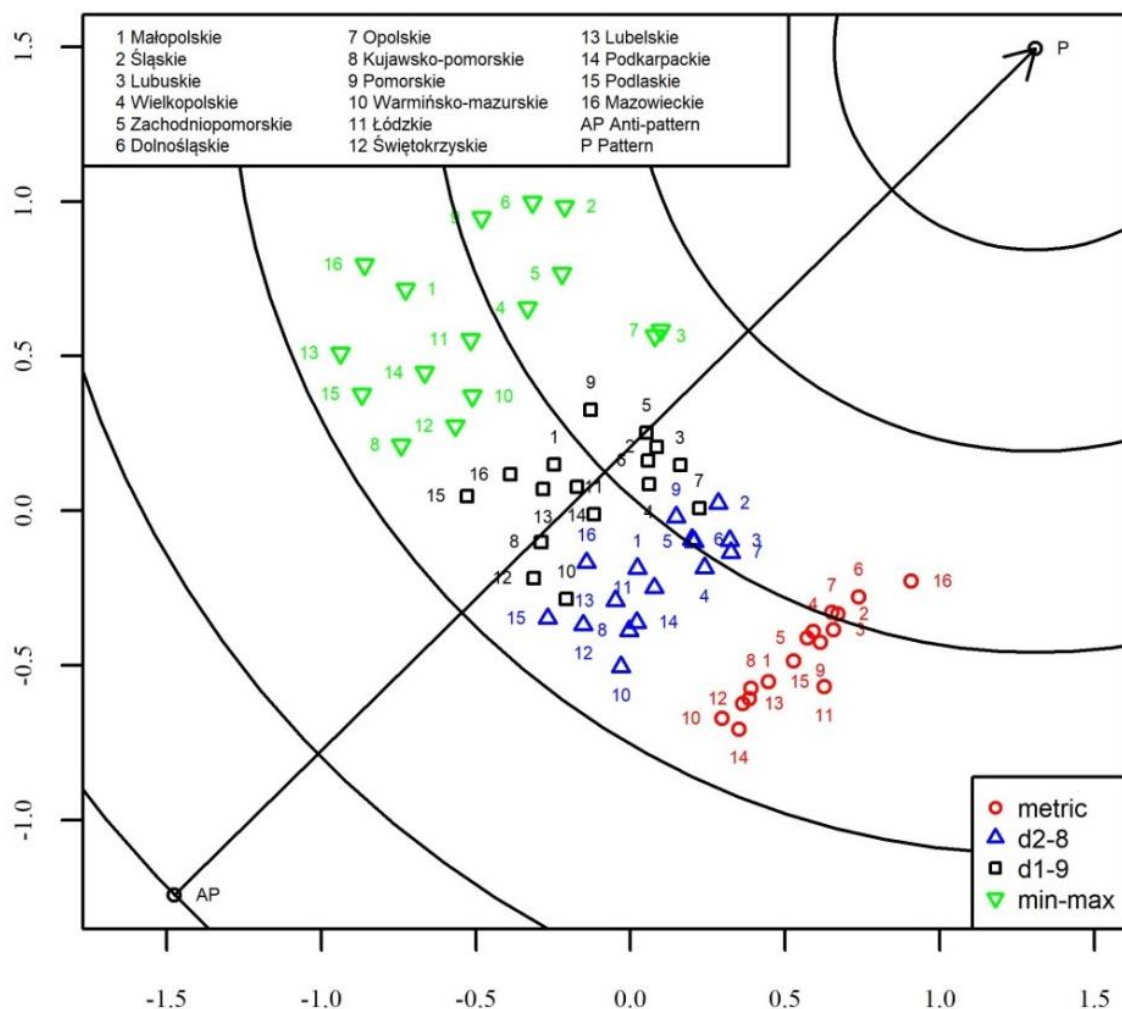


Figure 2. Results of multidimensional scaling of 16 provinces of Poland for 4 types of data according to the level of social cohesion in 2018. Sources: calculations and plot produced using the R program.

The results made it possible to assess the level of social cohesion in the provinces using four types of data simultaneously. One thing worth noting in Figure 2 is the arrangement of provinces in relation to the set axis AP – P. Sets of provinces based on symbolic interval-valued data are located increasingly further away from the set based on metric data ($x_{ij}^l = x_{ij}^u$), as the width of the interval for each set increases (2nd and 8th deciles, 1st and 9th deciles, minimum and maximum).

Table 4 shows a ranking of 16 provinces for four types of data according to the level of social cohesion in 2018. Calculations were made using the clusterSim package.

Moreover, the dispersion of provinces, measured by the standard deviation and median absolute deviation (Table 4), increases as one moves from results based on metric data to those based on symbolic interval-valued data (with increasing interval width). The width of each set of provinces with respect to the set axis AP – P increasing as the its interval width increases (it is the smallest for metric data, where $x_{ij}^l = x_{ij}^u$, and largest for symbolic interval-valued data comprising 100% observations, from minimum to maximum).

Table 4. Rankings of 16 provinces of Poland for four types of data according to the level of social cohesion in 2018 (value of measure d_i).

Province	no	Metric	Rank	d2-8	Rank	d1-9	Rank	Min-Max	Rank
Mazowieckie	16	0.5472	1	0.4348	10	0.4400	12	0.4161	13
Dolnośląskie	6	0.5229	2	0.5040	5	0.5316	4	0.5641	5
Śląskie	2	0.5042	3	0.5410	1	0.5450	3	0.5887	3
Opolskie	7	0.5035	4	0.5125	3	0.5283	5	0.6116	1
Lubuskie	3	0.4904	5	0.5204	2	0.5465	2	0.6048	2
Wielkopolskie	4	0.4832	6	0.4901	7	0.5177	7	0.5273	6
Pomorskie	9	0.4772	7	0.5113	4	0.5254	6	0.5203	7
Zachodniopomorskie	5	0.4765	8	0.5034	6	0.5473	1	0.5657	4
Małopolskie	1	0.4546	9	0.4578	8	0.4732	9	0.4413	10
Łódzkie	11	0.4431	10	0.4531	9	0.4684	10	0.4736	8
Podlaskie	15	0.4311	11	0.3786	16	0.4009	15	0.3727	15
Kujawsko-Pomorskie	8	0.4202	12	0.4121	13	0.4215	13	0.3805	14
Świętokrzyskie	12	0.4117	13	0.3933	14	0.3959	16	0.4268	12
Lubelskie	13	0.4058	14	0.4255	11	0.4529	11	0.3714	16
Warmińsko-Mazurskie	10	0.3874	15	0.3835	15	0.4010	14	0.4517	9
Podkarpackie	14	0.3849	16	0.4214	12	0.4746	8	0.4274	11
Parameters									
Mean		0.4590		0.4589		0.4794		0.4840	
Standard deviation		0.0477		0.0518		0.0543		0.0821	
Median		0.4656		0.4555		0.4739		0.4627	
Median absolute deviation		0.0568		0.0715		0.0792		0.1088	

metric—metric data; d2-8—symbolic interval-valued data comprising 2nd and 8th decile (60% of observations), d1-9—symbolic interval-valued data comprising 1st and 9th decile (80% of observations); min-max—symbolic interval-valued data comprising minimum and maximum (100% of observations). Source: calculations made using R.

5.2. Comparative Analysis of the Results in the Assessment of Social Cohesion

The rankings of provinces of Poland according to the level of social cohesion were compared on the basis of the aggregate measure d_i for four types of data (see Table 4): metric data, symbolic interval-valued data comprising 2nd and 8th decile (60% of observations), 1st and 9th decile (80% of observations) as well as minimum and maximum (100% of observations).

The results of the assessment of social cohesion were compared using two criteria. The first one was based on coefficients of correlation (Spearman's ρ and Kendall's τ) between aggregate measures d_i for the four types of data in order to determine similarities and differences between different rankings of provinces according to social cohesion. The second criterion was the degree of compatibility between rankings of provinces based on individual variables and that based on the aggregate measure d_i for the four types of data. The results of this comparison were used to choose the ranking providing the best reflection of the level of social cohesion in the provinces of Poland. This reason why this second criterion was used is that the overall ranking is the result of rankings obtained for individual variables.

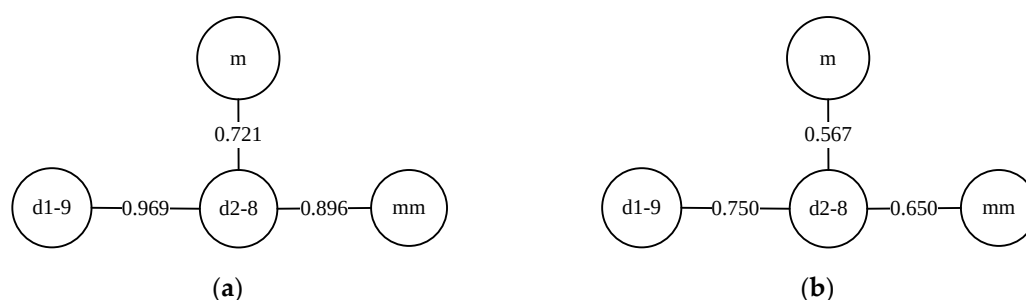
Spearman's ρ and Kendall's τ coefficients of correlation between aggregate measures d_i calculated for the four types of data are shown in Table 5.

For the purpose of cluster analysis, correlation coefficients were converted into distances: $d = \frac{1}{2}(1 - \rho)$ and $d = \frac{1}{2}(1 - \tau)$. Cluster analysis was used to identify similarities and differences in the rankings of provinces according to the level of social cohesion based on the aggregate measure d_i . The results were presented in the form of dendrites (Figure 3), following the Wrocław taxonomic method [52].

Table 5. Spearman's ρ and Kendall's τ coefficients of correlation between aggregate measures d_i for 4 types of data.

Data	m	d2-8	d1-9	mm
m	1.000/1.000			
d2-8	0.721/0.567	1.000/1.000		
d1-9	0.629/0.417	0.969/0.750	1.000/1.000	
mm	0.625/0.450	0.896/0.650	0.877/0.633	1.000/1.000

m—metric data; d2-8—symbolic interval-valued data comprising 2nd and 8th decile (60% of observations), d1-9—symbolic interval-valued data comprising 1st and 9th decile (80% of observations); mm—symbolic interval-valued data comprising minimum and maximum (100% of observations). Source: the authors' tabulation.

**Figure 3.** Dendrite diagrams showing results of cluster analysis the degree of similarity between aggregation measures d_i for 4 types of data: (a) Spearman's ρ correlation coefficient; (b) Kendall's τ correlation coefficient. Source: own presentation.

Results of the symbolic interval-valued approach differ considerably from the approach based on metric data. Assessments of social cohesion obtained using metric data are closest to results obtained for symbolic interval-valued data comprising 2nd and 8th decile.

The next step involved analysing the degree of compatibility between rankings of provinces based on individual variables and that based on the aggregate measure d_i for the four types of data, using the following procedure:

1. 16 provinces of Poland for 4 datasets (metric, min-max, 1st and 9th decile, 2nd and 8th decile) are linearly ordered according to a set of m variables to produce 4 rankings based on aggregate measures d_i (see Table 4).
2. For each variable ($j = 1, \dots, m$) a distance between each object and the pattern object is calculated according to the formula (the Ichino-Yaguchi distance for one variable):

$$d_i^j = 1 - \left| \varphi(x_{ij}, x_{+j}) \right| / \varphi(x_{+j}, x_{-j}) \quad (4)$$

where: $x_{ij} = [x_{ij}^l, x_{ij}^u]$ ($x_{ij}^l \leq x_{ij}^u$) interval ($x_{ij}^l = x_{ij}^u$, min-max, 1st and 9th decile, 2nd and 8th decile); $\varphi(x_{ij}, x_{+j}) = |x_{ij} \oplus x_{+j}| - |x_{ij} \otimes x_{+j}| + 0.5(2 \cdot |x_{ij} \otimes x_{+j}| - |x_{ij}| - |x_{+j}|)$; $\|$ interval length; $x_{ij} \oplus x_{+j} = x_{ij} \cup x_{+j}$; $x_{ij} \otimes x_{+j} = x_{ij} \cap x_{+j}$; x_{+j} (x_{-j}) interval of the pattern (anti-pattern) object for j -th variable. This yields m values of measures $d_i^1, d_i^2, \dots, d_i^m$.

3. The general rankings based on the aggregate measures d_i (step 1) are compared with individual rankings based on measures $d_i^1, d_i^2, \dots, d_i^m$ (step 2), separately for each data type using Spearman's ρ and Kendall's τ correlation coefficients.
4. For each data type, the median of results obtained in step 3 is calculated. A higher value of the median represents a higher degree of compatibility between rankings based on individual variables and the ranking based on the aggregate measure. The results are displayed in Table 6.

Table 6. The degree of compatibility between rankings of objects based on individual variables and the ranking based on the aggregate measure d_i .

No.	Types of Data	Median Value			
		Spearman's ρ	Rank	Kendall's τ	Rank
1	Metric data	0.4176	3	0.2667	3
2	Interval-valued (2nd and 8th deciles)	0.4371	1	0.3333	1
3	Interval-valued (1st and 9th deciles)	0.4297	2	0.2833	2
4	Interval-valued (min-max)	0.2676	4	0.1667	4

Source: the authors' tabulation.

The greatest degree of compatibility between rankings of objects based on individual variables and the ranking created using the aggregate measure d_i was obtained for intervals based on 2nd and 8th deciles. The use of deciles is an example of a robust approach (helps to eliminate the effect of outliers).

The ranking of provinces obtained on the basis of metric data does not account for the variation in the level of social cohesion between districts of a given province (lower level units). Mazowieckie (16) and Podlaskie (15) are characterised by a big disparity in the level of social cohesion between the leading district (Warsaw and Białystok, respectively) and the other ones. The leading districts have a strong impact on the overall level of variables at province level (metric data), and, consequently, on the more favourable assessment of social cohesion (see Table 4, columns 3 and 4). When metric data are replaced with symbolic interval-valued data, which account for the variation between districts with respect to the variables of interest, these two provinces are located much closer to the anti-pattern (see Figure 2). A switch from metric to interval-valued data comprising 2nd and 8th decile leads to the biggest decline in the values of measure d_i for Mazowieckie (−0.1124) and Podlaskie (−0.0525). This is also reflected by the ranking drops registered by these provinces: down 9 places (Mazowieckie) and down 5 places (Podlaskie).

The results of the study indicate that the differences in the rankings depend on the degree of variation between districts within a given province. The biggest changes in the ranking of provinces after switching from metric to interval-valued data were observed for those objects (provinces) which were characterised by the relatively biggest variation among districts in terms of the study variables (see Figure 4). The use of interval-valued data comprising 2nd and 8th decile made it possible to eliminate the impact of outliers. Provinces in which districts were not very different from one another, in both cases occupied the same or similar position in the ranking (see Figure 5).

Figure 6 shows the spatial distribution of Polish provinces in 2018 in terms of the level of social cohesion based on values of measure d_i for metric data and interval-valued data (the approach involving 2nd and 8th deciles). After applying another method of measurement the position of provinces has changed significantly. The provinces were not assessed merely on the basis of mean values, which can easily be affected by outliers, but also by accounting for how the variables varied across districts. In the case of provinces, districts classified as spatial poles of growth behave like extreme observations and can strongly influence the measurement for the entire province they belong to. This phenomenon is exemplified by measurements for the province of Mazowieckie, which are strongly affected by the district of Warsaw. The proposed modification of the method (using of interval-valued data) has made it possible to avoid this problem.

The assessment of the level of social cohesion across provinces on the basis of interval-valued data (the approach involving 2nd and 8th deciles) is clearly consistent with the assessment of the level of economic development. Provinces situated in Western Poland, characterised by a higher level of economic development, have also a higher level of social cohesion. It should be assumed that such a system of spatial differentiation results to a large extent from historical conditions, mainly related to the period of partitions (1795–1918) and changes after World War II.

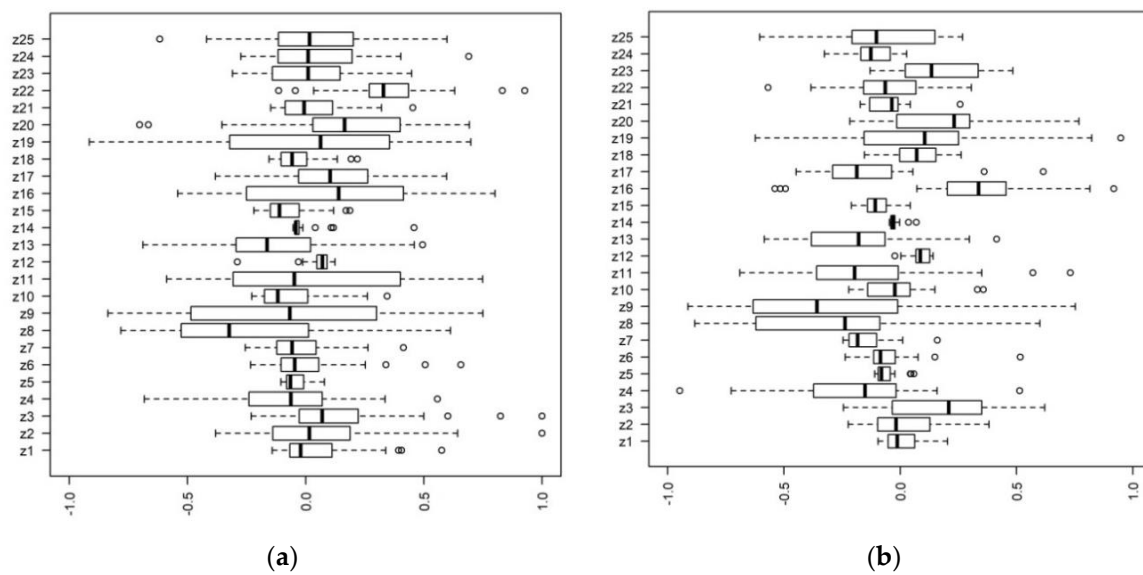


Figure 4. Distribution of normalised values of variables (boxplots) for Mazowieckie and Podlaskie across districts: (a) Mazowieckie (down 9 places); (b) Podlaskie (down 5 places). Source: calculations and plots made using R.

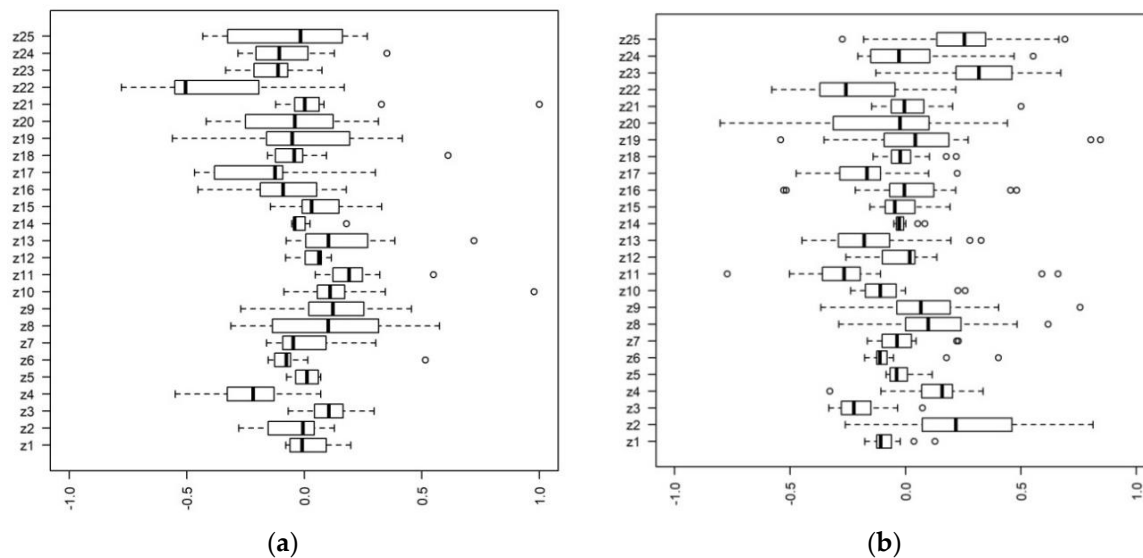


Figure 5. Distribution of normalised values of variables (boxplots) for Opolskie and Warmińsko-Mazurskie across districts: (a) Opolskie (up 1 place); (b) Warmińsko-Mazurskie (no change). Source: calculations and plots made using R.

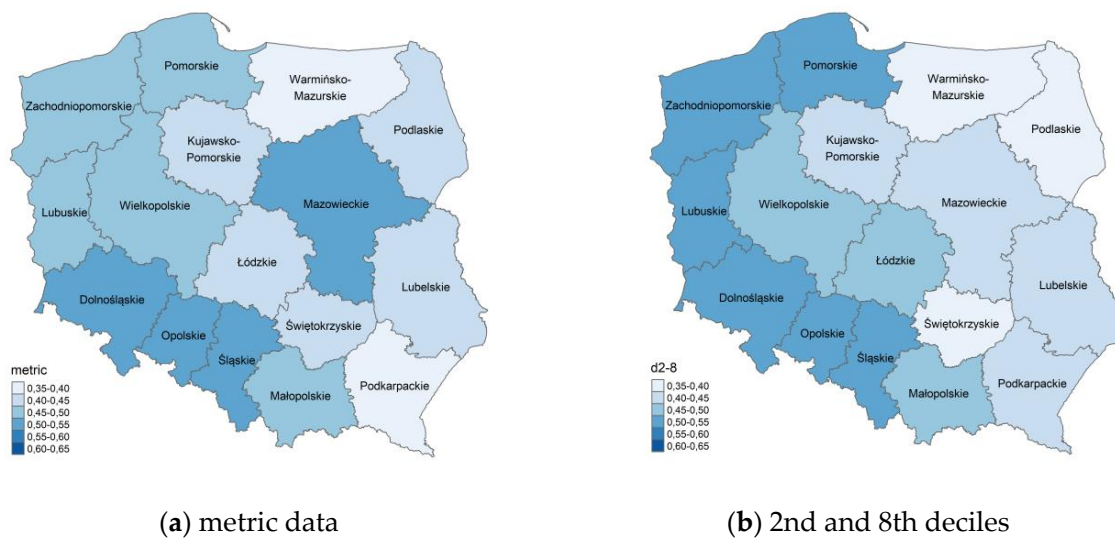


Figure 6. (a,b) The level of social cohesion in 2018 based on value of measure d_i for metric data and interval-valued data (2nd and 8th deciles approach). Source: calculations and plots made using R.

6. Discussion and Conclusions

As a result of economic changes there is a growing demand for assessment of social cohesion. It plays an important role both in the economic and social sphere not only for the country as a whole but, more importantly, at the regional level. This is because there is a need to develop a methodology dedicated to this area of interest.

In the study the level of social cohesion in provinces of Poland in 2018 was assessed using a hybrid approach combining multidimensional scaling and linear ordering. The traditional approach is conducted using classic metric data and it does not account for the variation between lower level units (i.e., districts). The authors propose a methodology which makes this possible. The dataset containing the classic metric data was extended to symbolic interval-valued data (three data types: min-max, 1st decile and 9th decile, 2nd decile and 8th decile) and used in the analysis. In addition to producing a ranking of provinces according to the level of social cohesion, the results of assessment were presented in a two-dimensional space.

The results of the symbolic interval-valued approach differ considerably from those obtained using metric data (see Table 4 and dendrite diagrams in Figure 3). The differences become more evident, the wider the intervals are. The biggest degree of similarity between rankings of objects based on individual variables and the ranking based on the aggregate measure was obtained for intervals based on 2nd decile and 8th decile comprising 60% of observations at district level. In this case, social cohesion in the provinces was not assessed on the basis of atomic variable values, but using information from 60% of districts (the interval between 2nd and 8th decile). This approach made it possible to take into account the degree of variation in social cohesion between lower level units (districts) within a given province. The use of 2nd decile and 8th decile is an example of a robust approach, which is used for eliminating the effect of outliers.

The proposed modification makes it possible to assess social cohesion in provinces not only on the basis of a single real number (metric data), but also by taking into account characteristics of particular districts within each province (interval-valued data). It was found that the assessment of social cohesion in provinces based on interval-valued data is strongly affected by lower level units (districts). The overall assessment of a given province does not depend on one or two districts but on all of them. Decision makers at province level should therefore ensure that all districts develop more or less uniformly.

The novelty of the study presented in this article consists in jointly mapping, by means of a hybrid approach combining multidimensional scaling and linear ordering, classic metric data and symbolic

interval-valued data (three data types: min-max, 1st decile and 9th decile, 2nd decile and 8th decile) in one chart. This approach makes it possible to take into account the degree of variation in social cohesion between lower level units (districts) within a given province.

It is worth adding that the method of assessing the level of social cohesion in territorial units proposed in the article does not require an additional survey, as it relies on secondary data sources. However, its application is limited by the availability of appropriate data. Not all phenomena associated with social cohesion are measured by official statistics. The situation looks different regarding the assessment of social cohesion based on primary data. In such cases, the researcher can include all dimensions of social cohesion when designing the survey questionnaire.

All calculations were performed using R scripts written by the authors.

Author Contributions: conceptualization, M.W., D.G.; methodology, M.W.; software, M.W., validation, M.W., D.G.; formal analysis, M.W., D.G.; investigation, M.W., D.G.; resources, M.W.; data curation, M.W.; writing—original draft preparation, M.W., D.G.; writing—review and editing, M.W., D.G.; visualization, M.W.; supervision, M.W., D.G. All authors have read and agreed to the published version of the manuscript.

Funding: The project is financed by the Polish National Science Centre (Narodowe Centrum Nauki) DEC-2015/17/B/HS4/00905.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Gioia, F.; Lauro, C.N. Principal component analysis on interval data. *Comput. Stat.* **2006**, *21*, 343–363. [CrossRef]
2. Brito, P.; Noirhomme-Fraiture, M.; Arroyo, J. Editorial for special issue on symbolic data analysis. *Adv. Data Anal. Classif.* **2015**, *9*, 1–4. [CrossRef]
3. Duhaime, G.; Searles, E.; Usher, P.J.; Myers, H.; Frechette, P. Social cohesion and living conditions in the Canadian Arctic: From theory to measurement. *Soc. Indic. Res.* **2004**, *66*, 295–318. [CrossRef]
4. Bernard, P. La cohésion sociale: Critique dialectique d'un quasi-concept. *Lien Soc. Polit.* **1999**, *41*, 47–59. [CrossRef]
5. Chan, J.; To, H.P.; Chan, E. Reconsidering social cohesion: Developing a definition and analytical framework for empirical research. *Soc. Indic. Res.* **2006**, *75*, 273–302. [CrossRef]
6. Acket, S.; Borsenberger, M.; Dickes, P.; Sarracino, F. Measuring and validating social cohesion: A bottom-up approach. In *International Conference on Social Cohesion and Development*; OECD, Development Center: Paris, France, 2011; pp. 20–21. Available online: <https://search.oecd.org/development/pgd/46839973.pdf> (accessed on 12 July 2020).
7. Langer, A.; Stewart, F.; Smedts, K.; Demarest, L. Conceptualising and measuring social cohesion in Africa: Towards a perceptions-based index. *Soc. Indic. Res.* **2017**, *131*, 321–343. [CrossRef]
8. Balcerzak, A.P. Wielowymiarowa analiza spójności społecznej w krajach Unii Europejskiej w kontekście strategii Europa 2020. In *Aktualne Trendy w Zarządzaniu Środowiskiem*; Bartniczak, B., Trzeciak, K., Eds.; Wydawnictwo AD REM: Jelenia Góra, Poland, 2015; pp. 343–352.
9. Hellwig, Z. Procedure of evaluating high-level manpower data and typology of countries by means of the taxonomic method. In *Towards a System of Human Resources Indicators for Less Developed Countries*; Gostkowski, Z., Ed.; The Polish Academy of Sciences Press: Wrocław, Poland, 1972; pp. 115–134.
10. Schaeffer, M. Can competing diversity indices inform us about why ethnic diversity erodes social cohesion? A test of five diversity indices in Germany. *Soc. Sci. Res.* **2013**, *42*, 755–774. [CrossRef]
11. Holgado, M.M.; Salinas, J.A.; Rodriguez, J.A. A synthetic indicator to measure the economic and social cohesion of the regions of Spain and Portugal. *Rev. De Econ. Mund.* **2015**, *39*, 223–240. Available online: <https://digibug.ugr.es/handle/10481/36997> (accessed on 12 July 2020).
12. Janmaat, J.G. Social cohesion as a real-life phenomenon: Assessing the explanatory power of the universalist and particularist perspectives. *Soc. Indic. Res.* **2011**, *100*, 61–83. [CrossRef]
13. Bottoni, G. A multilevel measurement model of social cohesion. *Soc. Indic. Res.* **2018**, *136*, 835–857. [CrossRef]
14. Dickes, P.; Valentova, M. Construction, validation and application of the measurement of social cohesion in 47 European countries and regions. *Soc. Indic. Res.* **2013**, *113*, 827–846. [CrossRef]

15. Dickes, P.; Valentova, M.; Borsenberger, M. Construct validation and application of a common measure of social cohesion in 33 European countries. *Soc. Indic. Res.* **2010**, *98*, 451–473. [\[CrossRef\]](#)
16. Rajulton, F.; Ravanera, Z.R.; Beaujot, R. Measuring social cohesion: An experiment using the Canadian national survey of giving, volunteering, and participating. *Soc. Indic. Res.* **2007**, *80*, 461–492. [\[CrossRef\]](#)
17. Lafuente, J.Á.; Marco, A.; Monfort, M.; Ordóñez, J. Social exclusion and convergence in the EU: An assessment of the Europe 2020 strategy. *Sustainability* **2020**, *12*, 1843. [\[CrossRef\]](#)
18. Dehnel, G.; Walesiak, M.; Obregalski, M. Comparative analysis of the ordering of Polish provinces in terms of social cohesion. *Argum. Oeconomica Crac.* **2019**, *1*, 71–85. [\[CrossRef\]](#)
19. Kania, K. bdl: Interface and tools for ‘BDL’ API. R package, version 1.0-0. 2019. Available online: <http://CRAN.R-project.org/package=bdl> (accessed on 15 February 2020).
20. Jenson, J. *Mapping Social Cohesion: The State of Canadian Research*; Canadian Policy Research Networks: Ottawa, ON, Canada, 1998; pp. 109–128. Available online: http://www.cccg.umontreal.ca/pdf/CPRN/CPRN_F03.pdf (accessed on 12 July 2020).
21. Jenson, J. *Defining and Measuring Social Cohesion*; UNRISD & Commonwealth Secretariat: London, UK, 2010.
22. Beauvais, C.; Jenson, J. *Social Cohesion: Updating the State of the Research*; Canadian Policy Research Networks: Ottawa, ON, Canada, 2002; Available online: http://oaresource.library.carleton.ca/cprn/12949_en.pdf (accessed on 12 July 2020).
23. Hulse, K.; Stone, W. Social cohesion, social capital and social exclusion: A cross cultural comparison. *Policy Stud.* **2007**, *28*, 109–128. [\[CrossRef\]](#)
24. Novy, A.; Swiatek, D.C.; Moulart, F. Social cohesion: A conceptual and political elucidation. *Urban. Stud.* **2012**, *49*, 1873–1889. [\[CrossRef\]](#)
25. Klein, C. Social capital or social cohesion: What matters for subjective well-being? *Soc. Indic. Res.* **2013**, *110*, 891–911. [\[CrossRef\]](#)
26. Ariely, G. Does diversity erode social cohesion? Conceptual and methodological issues. *Political Stud.* **2014**, *62*, 573–595. [\[CrossRef\]](#)
27. Fonseca, X.; Lukosch, S.; Brazier, F. Social cohesion revisited: A new definition and how to characterize it. *Innovation. Eur. J. Soc. Sci. Res.* **2019**, *32*, 231–253. [\[CrossRef\]](#)
28. Schiefer, D.; van der Noll, J. The essentials of social cohesion: A literature review. *Soc. Indic. Res.* **2017**, *132*, 579–603. [\[CrossRef\]](#)
29. Council of Europe. *Concerted Development of Social Cohesion Indicators: Methodological Guide*; Council of Europe Publishing: Strasbourg, France, 2005; Available online: https://www.coe.int/t/dg3/socialpolicies/socialcohesiondev/source/GUIDE_en.pdf (accessed on 12 July 2020).
30. Berger, P.R. (Ed.) *The Limits of Social Cohesion: Conflict and Mediation in Pluralist Societies*; Routledge: New York, NY, USA, 1998; Available online: <https://www.taylorfrancis.com/books/e/9780429496325> (accessed on 12 July 2020).
31. Gough, I.; Olofsson, G. (Eds.) *Capitalism and Social Cohesion: Essays on Exclusion and Integration*; Palgrave Macmillan: London, UK, 1999. [\[CrossRef\]](#)
32. Friedkin, N.E. Social cohesion. *Annu. Rev. Sociol.* **2004**, *30*, 409–425. [\[CrossRef\]](#)
33. Berger-Schmitt, R. *Social Cohesion as an Aspect of the Quality of Societies: Concept and Measurement*; Centre for Survey Research and Methodology (ZUMA): Mannheim, Germany, 2000; Available online: https://www.gesis.org/fileadmin/upload/dienstleistung/daten/soz_indikatoren/eusi/paper14.pdf (accessed on 12 July 2020).
34. Berger-Schmitt, R. Considering social cohesion in quality of life assessments: Concept and measurement. *Soc. Indic. Res.* **2002**, *58*, 403–428. [\[CrossRef\]](#)
35. Noll, H.H. Towards a European system of social indicators: Theoretical framework and system architecture. *Soc. Indic. Res.* **2002**, *58*, 47–87. [\[CrossRef\]](#)
36. Naranjo, J.M. Impacts on the social cohesion of mainland Spain’s future motorway and high-speed rail networks. *Sustainability* **2016**, *8*, 624. [\[CrossRef\]](#)
37. Rodríguez, J.A.; Martín, J.M.; Salinas, J.A.; Zermeno, K.; Añños, K.G. A spatial analysis of the achievements, in terms of regional development, accomplished by the initial EU-Member cohesion fund beneficiaries using a synthetic indicator. *Sustainability* **2019**, *11*, 2343. [\[CrossRef\]](#)

38. Darvas, Z.; Collin, A.M.; Mazza, J.; Midoes, C. Effectiveness of cohesion policy: Learning from the project characteristics that produce the best results, Study prepared for the European Parliament's Committee on Budgetary Control. European Union: Brussels, 2019. Available online: <https://www.bruegel.org/2019/06/effectiveness-of-cohesion-policy-learning-from-the-project-characteristics-that-produce-the-best-results-2/> (accessed on 11 September 2020).
39. Stiftung, B.; Eurofound. *Social Cohesion and Well-Being in Europe*; Publications Office of the European Union: Luxembourg, 2018; Available online: https://www.eurofound.europa.eu/sites/default/files/ef_publication/field_ef_document/ef18035en.pdf (accessed on 11 September 2020).
40. My Region, My Europe, Our Future. *Seventh Report on Economic, Social and Territorial Cohesion*; European Commission: Brussels, Belgium, 2017; Available online: http://ec.europa.eu/regional_policy/en/information/cohesion-report/ (accessed on 12 July 2020).
41. Annoni, P.; Dijkstra, L. *The EU Regional Social Progress Index: Methodological Note*; European Commission: Brussels, Belgium, 2016; Available online: https://ec.europa.eu/regional_policy/sources/information/maps/methodological_note_eu_spi_2016.pdf (accessed on 12 July 2020).
42. Walesiak, M. Visualization of linear ordering results for metric data with the application of multidimensional scaling. *Ekonometria* **2016**, *2*, 9–21. [CrossRef]
43. Hellwig, Z. Wielowymiarowa analiza porównawcza i jej zastosowanie w badaniach wielocechowych obiektów gospodarczych. In *Metody i Modele Ekonomiczno-Matematyczne w Doskonaleniu Zarządzania Gospodarką Socjalistyczną*; Welfe, W., Ed.; PWE: Warszawa, Poland, 1981; pp. 46–68.
44. Borys, T. *Kategoria Jakości w Statystycznej Analizie Porównawczej. Category of Quality in Statistical Comparative Analysis*; Wydawnictwo Akademii Ekonomicznej we Wrocławiu: Wrocław, Poland, 1984.
45. Milligan, G.W.; Cooper, M.C. A study of standardization of variables in cluster analysis. *J. Classif.* **1988**, *5*, 181–204. [CrossRef]
46. R Core Team. R: A language and environment for statistical computing. In *R Foundation for Statistical Computing*; Vienna, Austria, 2020; Available online: <http://www.R-project.org> (accessed on 15 June 2020).
47. Jajuga, K.; Walesiak, M. Standardisation of data set under different measurement scales. In *Classification and Information Processing at the Turn of the Millennium*; Decker, R., Gaul, W., Eds.; Springer: Berlin/Heidelberg, Germany, 2000; pp. 105–112. [CrossRef]
48. Billard, L.; Diday, E. *Symbolic Data Analysis: Conceptual Statistics and Data Mining*; John Wiley: Chichester, UK, 2006. [CrossRef]
49. Ichino, M.; Yaguchi, H. Generalized Minkowski metrics for mixed feature-type data analysis. *IEEE Trans. Syst. Man and Cybern.* **1994**, *24*, 698–708. [CrossRef]
50. Borg, I.; Groenen, P.J.F. Modern Multidimensional Scaling. In *Theory and Applications*; Springer Science+Business Media: New York, NY, USA, 2005. [CrossRef]
51. Mair, P.; De Leeuw, J.; Borg, I.; Groenen, P.J.F. Smacof: Multidimensional Scaling. R Package, Version 2.0-0. 2019. Available online: <http://CRAN.R-project.org/package=smacof> (accessed on 15 June 2020).
52. Florek, K.; Łukaszewicz, J.; Perkal, J.; Steinhaus, H.; Zubrzycki, S. Sur le liaison et la division des points d'un ensemble fini. *Colloquium Math.* **1951**, *2*, 282–285. Available online: <http://pldml.icm.edu.pl/pldml/element/bwmeta1.element.bwnjournal-article-cmv2i1p282bwm> (accessed on 12 July 2020).

