QUANTITATIVE METHODS IN ECONOMICS
Vol. XVI, No. 2, 2015, pp. 53 – 62

A COMPARISON OF THE METHODS OF RELATIVE TAXONOMY FOR THE ASSESSMENT OF INFRASTRUCTURAL DEVELOPMENT OF COUNTIES IN WIELKOPOLSKIE VOIVODESHIP

Jarosław Lira
Department of Finance and Accounting
University of Life Sciences in Poznan
e-mail: jlira@up.poznan.pl

Abstract: The study analyses the possibility to measure the scale of disproportion in the development of the synthetic feature between spatial objects over a period of time on the basis of relativised values of diagnostic features. The study also proposes the construction of taxonomically relative indices of development according to the approach based on spatial median and it compares this approach with the classic approach proposed by Wydymus (2013). Both approaches are illustrated with a numerical example referring to the economic infrastructure in rural areas of the counties in Wielkopolskie Voivodeship.

Keywords: relative taxonomy, economic infrastructure, rural areas

INTRODUCTION

Relative taxonomy considers relativised indices, which are defined as quotients of the values of individual diagnostic features describing the synthetic feature for each spatial object relative to the values of other objects. Methods used in relative taxonomy enable us to determine the position of a particular spatial object relative to other objects. In the dynamic approach, they allow us to determine the process of levelling developmental disproportions in terms of development of the synthetic feature. Wydymus [2013] proposed the construction of taxonomically relative indices of development in the dynamic approach based on mean values of indices of relativised diagnostic features for individual spatial objects. On the other hand, Lira et al. [2014] proposed the identification of relative typological classes based on relative indices of development and their analysis in the dynamic approach.

The study presents the construction of taxonomically relative indices of development according to the approach based on Weber's spatial median and it compares this approach with the classic approach proposed by Wydymus [2013]. Both approaches were applied to measure the position of a particular county relative to other counties in terms of rural inhabitants' access to infrastructural services in Wielkopolskie Voivodeship in 2013. The research material was based on data provided in the electronic form by the Central Statistical Office in Warsaw – Local Data Bank [2015].

RESEARCH METHODOLOGY

The synthetic feature is composed of simple features, i.e. diagnostic features, which can be expressed with values in the form of structure or intensity indices. Let us assume that simple features are stimulants. Otherwise, it is necessary to unify their character (by converting destimulants and nominants into stimulants). Next, the values of individual diagnostics for each object and each time period were relativised according to the formula:

$$d_{(b/c)jt} = \begin{cases} x_{bjt}/x_{cjt} & x_{cjt} \neq 0 \\ 0 & x_{cjt} = 0 \end{cases}$$

where: $b \neq c$, $b = 1, 2, \cdots, n$, $c = 1, 2, \cdots, n$,

x_{ijt} – denoted the observation in the *i-th* object ($i = 1, 2, \cdots, n$) of the *j-th* diagnostic feature ($j = 1, 2, \cdots, m$) in time period *t* ($t = 1, 2, \cdots, k$).

Thus relativised values of diagnostic features in an object relative to other spatial objects for diagnostic *j* and time period *t* could be presented in the following form:

$$\mathbf{D}_{jt} = \begin{bmatrix} 1 & d_{(2/1)jt} & \cdots & d_{(\mathrm{n}/1)jt} \\ d_{(1/2)jt} & 1 & & d_{(\mathrm{n}/2)jt} \\ \vdots & & \ddots & \vdots \\ d_{(1/\mathrm{n})jt} & d_{(2/\mathrm{n})jt} & \cdots & 1 \end{bmatrix}.$$

Matrices $\mathbf{D}_{jt}$ make the basis for the construction of taxonomically relative indices of development of the synthetic feature in the classic and positional approach based on Weber's spatial median.

Classic approach proposed by Wydymus [2013]

In order to classify the objects with respect to all diagnostic features simultaneously the subsequent matrices were calculated [Wydymus 2013]:

$$\mathbf{D}^*_{jt} = \mathbf{A} \cdot \mathbf{D}_{jt}$$

where the matrix **A** was defined as:

$$\mathbf{A} = \begin{bmatrix} 0 & \cdots & \frac{1}{(n-1)} \\ \vdots & \ddots & \vdots \\ \frac{1}{(n-1)} & \cdots & 0 \end{bmatrix}$$

The diagonal elements of $\mathbf{D}_{jt}^{*}$ matrices formed matrices $\mathbf{W}_t$ for each time period *t*:

$$\mathbf{W}_t = \begin{bmatrix} w_{11t} & w_{12t} & \cdots & w_{1mt} \\ w_{21t} & w_{22t} & \cdots & w_{2mt} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n1t} & w_{n2t} & \cdots & w_{nmt} \end{bmatrix}$$

Next, the $\mathbf{W}_t$ matrices were used to compute the S_{it} matrix of relative synthetic indices of development for given objects and time periods [Wydymus 2013]:

$$\mathrm{S}_{it} = \frac{1}{m}\sum_{j=1}^{\mathrm{m}}\frac{1}{w_{ijt}},$$

but if $w_{ijt} = 0$, we assume that $1/w_{ijt}$ is an arbitrarily determined value close to 0, e.g. 0.001.

Positional approach based on Weber's spatial median

Matrices $\mathbf{D}_{jt}$ make the basis for the construction of each spatial object *i* at period in time *t* of matrix $\mathbf{\Delta}_{it}$, which adopt the following form for individual objects:

$$\mathbf{\Delta}_{1t} = \begin{bmatrix} d_{(1/2)1t} & d_{(1/2)2t} & \cdots & d_{(1/2)mt} \\ d_{(1/3)1t} & d_{(1/3)2t} & & d_{(1/3)mt} \\ \vdots & & \ddots & \vdots \\ d_{(1/\mathrm{n})1t} & d_{(1/\mathrm{n})2t} & \cdots & d_{(1/\mathrm{n})mt} \end{bmatrix},$$

$$\mathbf{\Delta}_{2t} = \begin{bmatrix} d_{(2/1)1t} & d_{(2/1)2t} & \cdots & d_{(2/1)mt} \\ d_{(2/3)1t} & d_{(2/3)2t} & & d_{(2/3)mt} \\ \vdots & & \ddots & \vdots \\ d_{(2/\mathrm{n})1t} & d_{(2/\mathrm{n})2t} & \cdots & d_{(2/\mathrm{n})mt} \end{bmatrix},$$

$$\cdots$$

$$\mathbf{\Delta}_{nt} = \begin{bmatrix} d_{(\mathrm{n}/1)1t} & d_{(\mathrm{n}/1)2t} & \cdots & d_{(\mathrm{n}/1)mt} \\ d_{(\mathrm{n}/2)1t} & d_{(\mathrm{n}/2)2t} & & d_{(\mathrm{n}/2)mt} \\ \vdots & & \ddots & \vdots \\ d_{(\mathrm{n}/\mathrm{n}-1)1t} & d_{(\mathrm{n}/\mathrm{n}-1)2t} & \cdots & d_{(\mathrm{n}/\mathrm{n}-1)mt} \end{bmatrix}.$$

Next, Weber's spatial median

$$L_1\_med_{it} = (L_1\_med_{i1t}, L_1\_med_{i2t}, \cdots, L_1\_med_{imt})'$$

was calculated for each spatial object *i* at period in time *t* for the data gathered in matrices $\mathbf{\Delta}_{it}$, which can be treated as *n*-l observation vectors of *m*-feature objects. Matrix $\mathbf{\Omega}_t$ was used to classify spatial objects according to all diagnostic features.

$$\mathbf{\Omega}_{\mathrm{t}} = \begin{bmatrix} \omega_{11t} & \omega_{12t} & \cdots & \omega_{1mt} \\ \omega_{21t} & \omega_{22t} & \cdots & \omega_{2mt} \\ \vdots & \vdots & \ddots & \vdots \\ \omega_{n1t} & \omega_{n2t} & \cdots & \omega_{nmt} \end{bmatrix},$$

where:

for object $i = 1$:

$$\omega_{11t} = L_1\_med_{1t}\{d_{(1/2)1t}, d_{(1/3)1t}, \cdots, d_{(1/n)1t}\} = L_1\_med_{11t}$$
$$\omega_{12t} = L_1\_med_{1t}\{d_{(1/2)2t}, d_{(1/3)2t}, \cdots, d_{(1/n)2t}\} = L_1\_med_{12t}$$
$$\cdots$$
$$\omega_{1mt} = L_1\_med_{1t}\{d_{(1/2)mt}, d_{(1/3)mt}, \cdots, d_{(1/n)mt}\} = L_1\_med_{1mt}$$

for object $i = 2$:

$$\omega_{21t} = L_1\_med_{2t}\{d_{(2/1)1t}, d_{(2/3)1t}, \cdots, d_{(2/n)1t}\} = L_1\_med_{21t}$$
$$\omega_{22t} = L_1\_med_{2t}\{d_{(2/1)2t}, d_{(2/3)2t}, \cdots, d_{(2/n)2t}\} = L_1\_med_{22t}$$
$$\cdots$$
$$\omega_{2mt} = L_1\_med_{2t}\{d_{(2/1)mt}, d_{(2/3)mt}, \cdots, d_{(2/n)mt}\} = L_1\_med_{2mt}$$

for object $i = n$:

$$\omega_{n1t} = L_1\_med_{nt}\{d_{(n/1)1t}, d_{(n/2)1t}, \cdots, d_{(n/n-1)1t}\} = L_1\_med_{n1t}$$
$$\omega_{n2t} = L_1\_med_{nt}\{d_{(n/1)2t}, d_{(n/2)2t}, \cdots, d_{(n/n-1)2t}\} = L_1\_med_{n2t}$$
$$\cdots$$
$$\omega_{nmt} = L_1\_med_{nt}\{d_{(n/1)mt}, d_{(n/2)mt}, \cdots, d_{(n/n-1)mt}\} = L_1\_med_{nmt}.$$

Matrix $\mathbf{\Omega}_t$ was used to calculate the Φ_{it} matrix of relative indices of development for individual spatial objects at consecutive periods in time *t*:

$$\Phi_{it} = med\left\{\frac{1}{\omega_{i1t}}, \frac{1}{\omega_{i2t}}, \cdots, \frac{1}{\omega_{imt}}\right\},$$

but if $\omega_{ijt} = 0$, we assume $1/\omega_{ijt}$ is an arbitrarily determined value close to 0, e.g. 0.001.

The construction of taxonomically relative indices of development was extended for the case when a particular diagnostic feature assumes the value of 0 for a specific spatial objects. It can be observed especially with smaller objects, e.g. when rural inhabitants do not have access to a gas network in their communes.

This approach applies the median vector defined according to Weber's criterion. Therefore, it is called Weber's spatial median or it is defined as L_1 median or spatial median in reference publications [Lira 1999]. Let $\mathrm{K}_n^m = \{\mathbf{X}_1, \mathbf{X}_2, \cdots, \mathbf{X}_n\} \in \boldsymbol{\mathcal{R}}^m$ be a set of *n* vectors of observation of *m*-feature objects and let $\widehat{\mathbf{\Theta}} \in \boldsymbol{\mathcal{R}}^m$ be the vector solving the optimisation problem

$$\mathrm{T}\left(\widehat{\mathbf{\Theta}}, \mathrm{K}_n^m\right) = \min_{\mathbf{\Theta} \in \boldsymbol{\mathcal{R}}^m} \mathrm{T}(\mathbf{\Theta}, \mathrm{K}_n^m),$$

where the objective function of this problem assumes the following form:

$$\mathrm{T}(\boldsymbol{\Theta}, \mathrm{K}_n^m) = \sum_{i=1}^{n}\left[\sum_{j=1}^{m}\left(\mathrm{x}_{ij} - \theta_j\right)^2\right]^{1/2},$$

where $\mathbf{X}_i = (\mathrm{x}_{i1}, \mathrm{x}_{i2}, \cdots, \mathrm{x}_{im})'$, $i = 1, 2, \cdots, n$ and $\boldsymbol{\Theta} = (\theta_1, \theta_2, \cdots, \theta_m)'$.

The measurement of relative synthetic feature variations is based on the construction of taxonomically relative indices of development and it consists of the following four stages:

stage 1 – proposing a system of diagnostic features which can be determinants of development of the synthetic feature,

stage 2 – relativising the values of diagnostic features for each spatial object at a period in time,

stage 3 – constructing taxonomically relative indices of development of the synthetic feature,

stage 4 – identifying relative typological classes of spatial objects according to relative synthetic feature variations and describing them.

Stage 1. The selection of diagnostic features is based on substantive premises and statistical analysis of diagonal elements in the inverse matrix of correlation matrix **R** in order to avoid excessive correlation of features[1] [Lira, Wysocki 2004].

Stage 2. We calculate individual indices, assuming that individual spatial objects at a period in time under analysis are the basis for comparisons of each diagnostic feature. Next, we construct appropriate relative matrices for these features at a particular period in time.

Stage 3. It is possible to apply the classic approach proposed by Wydymus to construct taxonomically relative indices of development. When the set of diagnostic features includes strongly asymmetric features or features with outlying observations, we can apply the positional approach based on Weber's spatial median. The lesser the value of relative index is than 1, the greater the relative advantage of a particular spatial object is over all other spatial objects in terms of synthetic evaluation of the period in time under investigation.

Stage 4. The values of relative indices are used for linear ordering of spatial objects and identification of relative typological classes. Grouping spatial objects from a (very) low to a (very) high level of relative development can be based e.g. on analysis of differences in the relative indices. When we have ordered spatial objects according to decreasing values of relative indices, we can calculate differences between its values for neighbouring objects, i.e. for the first and second object, for the second and third object, etc. When we analyse consecutive differences, starting with the first difference (between the first and second object), and when we find that the value of this difference is much greater than the others,

[1] If a feature is excessively correlated with other features, diagonal elements of inverse matrix $\mathbf{R}^{-1}$ are much greater than 10, which is a symptom of wrong numerical conditioning of matrix **R** [Malina & Zeliaś 1997].

we can identify a relative typological class with low development. The other differences enable us to identify more classes.

RESEARCH RESULTS

We conducted an empirical analysis in four stages, as described in the research methodology. In order to measure relative variations in the economic infrastructure in rural areas of the counties in Wielkopolskie Voivodeship in 2013 we selected five continuous quantitative diagnostic features, whose values were expressed as stimulant structure indices (the intensity index, except roads):

- length of public communal roads of improved hard surface in km per 100 km^2 of county's rural areas (roads),
- users of water supply network as percentage of total rural population[2] (water supply),
- users of sewerage network as percentage of total rural population (sewerage),
- users of gas network as percentage of total rural population (gas),
- users serviced by sewage treatment facilities as percentage of total rural population (sewage treatment).

The calculations presented in Table 1 let us draw the following conclusions:

1. There is strong right-sided asymmetry in rural inhabitants' access to gas networks (1.47). This situation is strongly influenced by the outlying observation for Poznań County (66.6%), which is over 22 p.p. greater than in Grodzisk County, which is in the second position. Apart from that, this feature is characterised by relatively high dispersion (98.64% in the classic approach or 72.48% in the positional approach) among the features for which the coefficient of variation was calculated.
2. The coefficients of variation based on Weber's spatial median resulted in lower values for all the features than the classic coefficient of variation with a corresponding value.

Table 2 shows the results of research conducted at stages 3 and 4. In the classic approach the taxonomically relative indices of development resulted in excessively high values in Turek County (14.622) and Czarnków-Trzcianka County (2.425). Apart from that, we studied the similarity of orderings obtained by means of taxonomically relative indices based on the classic approach and on Weber's spatial median. We assumed Spearman's rank correlation coefficient as a measure of similarity between the orderings. Although the value of this coefficient was relatively high $r = 0.84; p = 0.000$, there were some differences between the positions of counties in terms of their economic infrastructure. The greatest difference was observed in Konin County, because in the positional

[2] taken as the number of actual inhabitants as of December 31, 2013

approach the position of this county was 10 positions worse than in the classic approach. Only two counties, i.e. Śrem and Turek, occupied the same positions in both approaches.

Table 1. Descriptive statistics characterising the components of economic infrastructure in rural areas of the counties in Wielkopolskie Voivodeship (as of 31 December 2013)

Characteristics	Roads	Water supply	Sewerage	Gas	Sewage treatment
minimum	10.24	75.57	16.60	0.03	19.19
lower quartile	22.13	85.76	27.21	3.64	27.02
marginal median	41.55	89.03	34.29	13.04	39.59
Weber's spatial median	38.90	88.59	35.69	14.64	38.58
upper quartile	52.03	91.50	43.04	22.05	51.41
maximum	75.23	94.13	53.68	66.64	72.47
coefficient of variation (%)	44.68	4.92	29.33	98.64	37.21
median absolute deviation[1)]	15.52	3.28	8.41	10.61	12.43
coefficient of variation based on Weber's spatial median[2)] (%)	39.91	3.70	23.57	72.48	32.23
kurtosis	-0.97	0.78	-0.81	2.32	-0.60
skewness	-0.01	-0.95	-0.13	1.47	0.42
diagonal elements of the inverse correlation matrix $\mathbf{R}^{-1}$	1.68	1.64	5.30	1.31	5.10

[1)] Median absolute deviation is defined by $mad_j = med_{ij}|x_{ij} - L_1\_med_j|$.

[2)] Coefficient of variation based on Weber's spatial median is given by $v_{pj} = \frac{mad_j}{L_1\_med_j} \cdot 100\%$.

Source: own calculations based on Local Data Bank, Central Statistical Office, Warsaw 2015

In both approaches we identified 4 relative typological classes (stage 4) and we made the following observations:

- 54.8% of the total number of counties belonged to the same relative typological class, whereas the other counties were grouped one class higher or lower,
- the class with the highest degree of relative infrastructural development (class I) was less numerous in the positional approach (4 counties) than in the classic approach (7 counties),
- class II in the positional approach (12 counties) was more numerous than in the classic approach (7 counties),
- there were similar numbers of counties in classes III and IV.

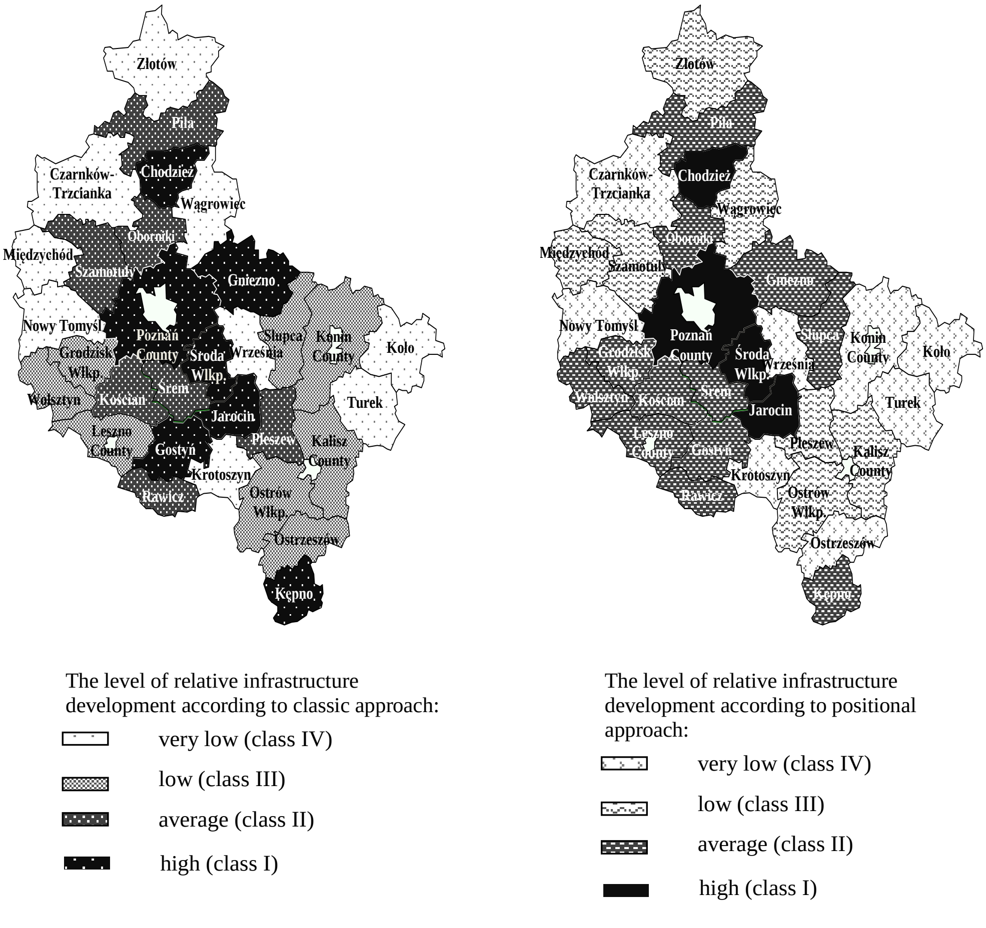
Table 2. The ordering and classification of rural areas of the counties in Wielkopolskie Voivodeship according to the economic infrastructure by means of the taxonomically relative indices of development

No.	Counties	Relative indices based on mean			Relative indices based on Weber's spatial median		
		ranks	indices	typological classes	ranks	indices	typological classes
1	Chodzież	3	0.571	I	1	0.664	I
2	Czarnków-Trzcianka	30	2.425	IV	25	1.307	IV
3	Gniezno	5	0.624	I	6	0.815	II
4	Gostyń	6	0.643	I	11	0.933	II
5	Grodzisk Wlkp.	22	0.889	III	16	0.984	II
6	Jarocin	1	0.542	I	2	0.668	I
7	Kalisz	15	0.808	III	21	1.084	III
8	Kępno	7	0.656	I	8	0.871	II
9	Koło	24	1.025	IV	30	1.730	IV
10	Konin	16	0.818	III	26	1.373	IV
11	Kościan	12	0.750	II	5	0.800	II
12	Krotoszyn	25	1.044	IV	29	1.618	IV
13	Leszno	19	0.836	III	13	0.950	II
14	Międzychód	27	1.067	IV	22	1.086	III
15	Nowy Tomyśl	29	1.132	IV	27	1.432	IV
16	Oborniki	13	0.782	II	9	0.879	II
17	Ostrów Wlkp.	21	0.889	III	17	1.006	III
18	Ostrzeszów	18	0.835	III	24	1.252	IV
19	Piła	8	0.693	II	7	0.863	II
20	Pleszew	11	0.730	II	20	1.078	III
21	Poznań	2	0.546	I	4	0.762	I
22	Rawicz	9	0.711	II	12	0.948	II
23	Słupca	20	0.853	III	15	0.979	II
24	Szamotuły	10	0.728	II	18	1.010	III
25	Środa Wlkp.	4	0.601	I	3	0.693	I
26	Śrem	14	0.784	II	14	0.954	II
27	Turek	31	14.622	IV	31	1.822	IV
28	Wągrowiec	26	1.046	IV	19	1.045	III
29	Wolsztyn	17	0.828	III	10	0.882	II
30	Września	23	1.001	IV	28	1.532	IV
31	Złotów	28	1.069	IV	23	1.088	III

Source: as in Table 1

The identification and analysis of the values of the taxonomically relative indices of development of the economic infrastructure according to the classic and positional approach revealed considerable spatial diversification in rural areas of the counties in Wielkopolskie Voivodeship. The diversification is illustrated in Figure 1, where the counties with the highest level of relative infrastructural development are marked with the darkest colour, whereas the counties with the lowest level of relative infrastructural development are marked with the lightest colour.

Figure 1. The delimitation of rural areas in Wielkopolskie Voivodeship according to the relative development of availability of infrastructural services in individual counties in 2013



Source: the Author's compilation based on the information in Table 2

CONCLUSIONS

The application of multidimensional methods of relative taxonomy enables analysis of relative developmental disproportions between individual counties and all the others in terms of the economic infrastructure in rural areas. The article presents two methods of construction of taxonomically relative indices of development according to the classic and positional approach in the dynamic aspect. The comparison of both approaches was based on one period of time that was arbitrarily selected for presentation.

The positional approach based on Weber's spatial median, which was applied for the construction of taxonomically relative indices of development, let us draw the following conclusions:

- the application of the positional approach is justified in a situation when the set of diagnostic features includes strongly asymmetric features or features with outlying observations,
- the taxonomically relative indices of development based on Weber's spatial median better reflects developmental disproportions between individual counties and all the others than the classic approach,
- the counties with sustainable development of economic infrastructure are positioned higher in the positional approach,
- the analysis of relative developmental disproportions can be conducted in a dynamic aspect, observing whether a particular county or relative typological class increases or decreases the developmental advantage over the others within the period of time under investigation.

REFERENCES

Lira J. (1998) Zastosowanie wybranych metod statystycznych opartych na medianie w doświadczalnictwie rolniczym, Praca doktorska, AR Poznań.

Lira J., Głowicka-Wołoszyn R., Wołoszyn A. (2014) The application of relative taxonomy methods to the study of technical infrastructure development in rural areas across the provinces of Poland, Quantitative Methods in Economics, Vol. XV, No. 2, pp. 330-338.

Lira J., Wysocki F. (2004) Zastosowanie pozycyjnego miernika rozwoju do pomiaru poziomu zagospodarowania infrastrukturalnego powiatów, Wiadomości Statystyczne, 9, GUS PTS, Warszawa, str. 39-49.

Local Data Bank, www.stat.gov.pl

Malina A., Zeliaś A. (1997) Taksonomiczna analiza przestrzennego zróżnicowania jakości życia ludności w Polsce w 1994 r., Przegląd Statystyczny 1, str. 11-27.

Wydymus S. (2013) Rozwój gospodarczy a poziom wynagrodzeń w krajach Unii Europejskiej – analiza taksonomiczna, Zeszyty Naukowe Uniwersytetu Szczecińskiego Nr 756, Finanse, Rynki Finansowe, Ubezpieczenia nr 57, str. 631-645.