QUANTITATIVE METHODS IN ECONOMICS
Vol. XV, No. 2, 2014, pp. 330 – 338

THE APPLICATION OF RELATIVE TAXONOMY METHODS TO THE STUDY OF TECHNICAL INFRASTRUCTURE DEVELOPMENT IN RURAL AREAS ACROSS THE PROVINCES OF POLAND

Jarosław Lira, Romana Głowicka-Wołoszyn
Department of Finance and Accouting, University of Life Sciences in Poznan
e-mails: jlira@up.poznan.pl; roma@up.poznan.pl
Andrzej Wołoszyn
Department of Informatics, University of Physical Education in Poznan
e-mail: woloszyn@awf.poznan.pl

Abstract: The article applies dynamic multi-dimensional methods of relative taxonomy in an attempt to evaluate disparities in development of technical infrastructure in rural areas between the provinces of Poland. The results show that although the indices that describe the level of infrastructure development have been rising between 2004 and 2012, regional inequalities have remained high. Moreover, some of the provinces least developed in 2004 have not taken advantage of the opportunities provided by the EU accession to develop technical infrastructure in rural areas and boost their attractiveness.

Keywords: relative taxonomy, technical infrastructure, rural areas

INTRODUCTION

Quality of life and economic growth in rural areas depend heavily on the level of development of their technical infrastructure. The access to its components: water supply, sewerage, gas, electricity, transportation, and communication systems, determines the investment attractiveness of these areas [Chudy 2011]. Due to high capital intensity, infrastructure development requires substantial investments and significant State participation in their financing. Polish accession to the European Union provided access to EU funds and opportunity to quicken the pace of rural infrastructure development. To be sure, the indices of this development were climbing steadily between 2004 and 2012. The pace, however, has not been uniform across the provinces.

Study of complex phenomena that are typically described by several diagnostic features often employs multi-dimensional comparative analysis to reduce the space of the features to one-dimensional synthetic index. Temporalized taxonomic methods, which are described i.a. in Grabiński [1984] or Zeliaś [2000], permit not only classification of objects in a given time period, but also analysis of evolution of the synthetic index. Hydzik [2011] suggests to study the progression simultaneously in two synthetic indices by what may be called *object development matrix*. A different method to study changes in synthetic indices between objects is proposed by Wydymus [2013]. His method consists of constructing the indices based on relativized diagnostic features.

The article aims to evaluate the scale of disparities in rural technical infrastructure development among the Polish provinces between 2004 and 2012. The dynamic approach to multidimensional methods of relative taxonomy not only allowed to compare provincial levels of this development but also to examine the process of levelling regional disparities following the accession of Poland to the European Union.

MATERIALS AND METHODS

For the evaluation of rural technical infrastructure development among the provinces of Poland in 2004-2012, five diagnostic features were selected[1] (all stimulants):

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

The method used for studying disparities in infrastructure development was devised by Wydymus [2013] and consists of constructing relative synthetic indices. The values of individual features for each object (province) and each time period (year) were relativized according to the formula:

$$d_{(b/c)jt} = x_{bjt} / x_{cjt}$$

where: $b \neq c$, $b=1,\ldots, n$, $c=1,\ldots, n$

[1] the selection process followed a thorough appraisal of their merits but also statistical analysis of diagonal elements of the inverse correlation matrix of the features, which helped to avoid excessive correlation in the diagnostic set [Lira, Wysocki 2004]

[2] taken as the number of actual inhabitants as of December, 31 of any given year

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

Thus transformed infrastructure indices of the *c-th* object relative to other objects for feature *j* and time period *t* could be presented in the following form:

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

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

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

where the matrix A was defined as:

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

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

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

The higher the value of w_{ijt} index, the greater was the advantage of the *i-th* object over remaining ones in the *j-th* feature and in the *t-th* period.

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:

$$S_{it} = \frac{1}{m} \sum_j \frac{1}{w_{ijt}}$$

The values of S_{it} smaller than 1 signified relative advantage of the *i-th* object over others in period *t.*

Research material was obtained from the Local Data Bank published by the Central Statistical Office in Warsaw. Calculations were performed using the R program; the script of the method's algorithm is available in the Appendix.

RESEARCH RESULTS

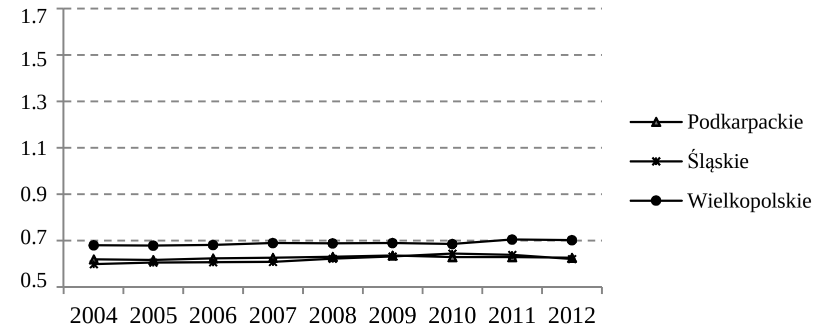
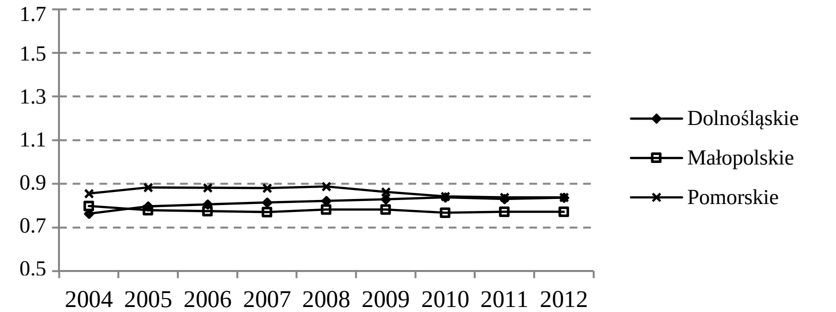
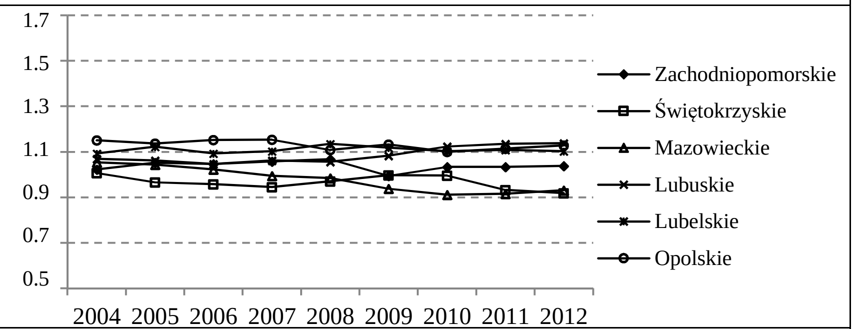
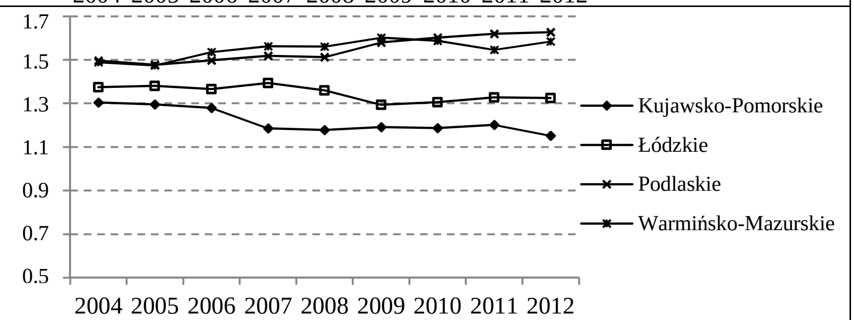
The values of rural technical infrastructure development relative synthetic index S_{it} calculated for 2004 period were used for linear ordering of the provinces: from the highest values of the index to the lowest. Next, the differences between adjacent provinces were computed and used to classify all the provinces into four typological classes. The decision to split a class was made when the differences came out relatively high. Class I of high relative development level comprised the provinces of Śląskie, Podkarpackie and Wielkopolskie, class II of medium high relative level: Dolnośląskie, Małopolskie and Pomorskie, class III of medium low relative level: Lubelskie, Lubuskie, Mazowieckie, Opolskie, Świętokrzyskie and Zachodniopomorskie, and class IV of low relative level: Kujawsko-Pomorskie, Łódzkie, Podlaskie and Warmińsko-Mazurskie.

During the first stage of the analysis the S_{it} values of each province were evaluated for the whole 2004-2012 period. Table 1 depicts them graphically in four 2004 typological classes. All four plots of Table 1 preserved the same scale of the vertical axis in order to facilitate comparison of S_{it} values.

In the whole 2004-2012 period the most noticeable improvement in relative estimates of technical infrastructure development was observed in four provinces: Kujawsko-Pomorskie, Mazowieckie, Świętokrzyskie and Łódzkie, which in 2004 belonged to the medium low (III) or low (IV) relative classes of development.

The maximum drop in the S_{it} index, from 1.30 in 2004 to 1.15 in 2012, which indicated marked improvement in development, occurred in Kujawsko-Pomorskie (class IV). Nonetheless, even with such a striking progress the relative estimate of rural infrastructure in the province remained low. Kujawsko-Pomorskie was followed by Mazowieckie (class III) with the second highest drop in the index, from 1.05 to 0.93. Less pronounced falls in the index, or in other words smaller progress in relative estimates of development, were noted in Świętokrzyskie (class III) and Łódzkie (class IV) and amounted to 0.09 and 0.05, respectively. Four provinces: Dolnośląskie (class II), Lubuskie (class III), Warmińsko-Mazurskie and Podlaskie (class IV) showed worsened estimates of relative development. In Dolnośląskie, despite the increase in the index from 0.76 to 0.84, the relative estimate of infrastructure was still better than in the third and fourth relative classes. Similarly, Lubuskie, where the index reached 1.14 in 2012, was still better off than the provinces of the fourth class. The most worrying changes were observed in Warmińsko-Mazurskie and Podlaskie. Not only had they been considered the provinces least developed in terms of rural infrastructure in 2004, their relative position further deteriorated in 2012. The remaining eight provinces (including the three that in 2004 counted among the first class) showed stable relative estimates: the changes of S_{it} index did not exceed 0.03 either way.

Table 1. Values of relative synthetic index S_{it} for all Polish provinces in 2004-2012

Relative class of development	Values of S_{it} index
I	
II	
III	
IV	

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

Analysis that was conducted in the second stage involved the relative classes created in 2004. Table 2 shows the basic characteristics of analyzed diagnostic features and relative synthetic indices in 2004 and 2012.

The first relative class, which comprised the provinces of Śląskie, Podkarpackie and Wielkopolskie, covered 18.4% of rural areas and close to 25% of population actually living in the rural areas between 2004 and 2012. This class stood out for the highest values of gas and sewerage network, and sewage treatment diagnostic features during the whole 2004-2012 period. The rate of technical infrastructure development in this class was better only than in the slowest second class (except in water supply network) and slower than in the third and fourth classes (except for sewerage network in class IV).

The second relative class with the provinces of Dolnośląskie, Małopolskie and Pomorskie covered 16.6% of rural areas and roughly 22% of rural population. This class was marked by increasing advantage in access to the sewerage network relative to classes III (by 1.1 p.p.) and IV (by 4.1 p.p.) and in access to sewage treatment facilities (by 2.6 p.p. and 4.3 p.p., respectively). Moreover, it maintained its advantage over the two lower classes in road network density (by 13 and 17 km/100 km^2, respectively) and in access to the gas network (by 16.7 p.p. and 27.2 p.p., respectively). A peculiarity of this class lay in its poor access to water supply network, poorer than in any of the remaining classes. As mentioned before, the rate of development of this class was lower than in class I (except in water supply network), class III, and class IV (except in water and sewerage systems).

The largest third class of six provinces covered 38.5% of rural areas and 34.2% of rural population. It had advantage over the fourth class in road network density and in access to gas network, and leveled out the fourth class' advantage in access to sewerage network and sewage treatment plants. Moreover, it had the fastest rate of development of all classes (except for road network density in class IV).

The fourth class with the provinces of Kujawsko-Pomorskie, Łódzkie, Podlaskie and Warmińsko-Mazurskie covered 26.5% of rural areas and 18.7% of rural population. It was the least developed of all four classes, but showed the fastest rate of development in road network density and only slightly slower than the third class' in gas network and sewer treatment access.

The individual diagnostic features in the classes that were characterized by notably faster annual growth rate in 2004-2012 showed positive values of relative annual growth rate (Table 2).

Figure 1 illustrates the values of relative synthetic index S_{it} calculated for the four relative classes. There was little of the levelling out in 2012 of the differences existing in 2004. Slight deterioration was observed in the second class of medium high development (S_{it} rose by 0.04) and equally slight improvement in the third class of medium low development (S_{it} dropped by 0.06).

Table 2. Relative classification of provinces and inter-class disparity in rural technical infrastructure development in 2004 and 2012

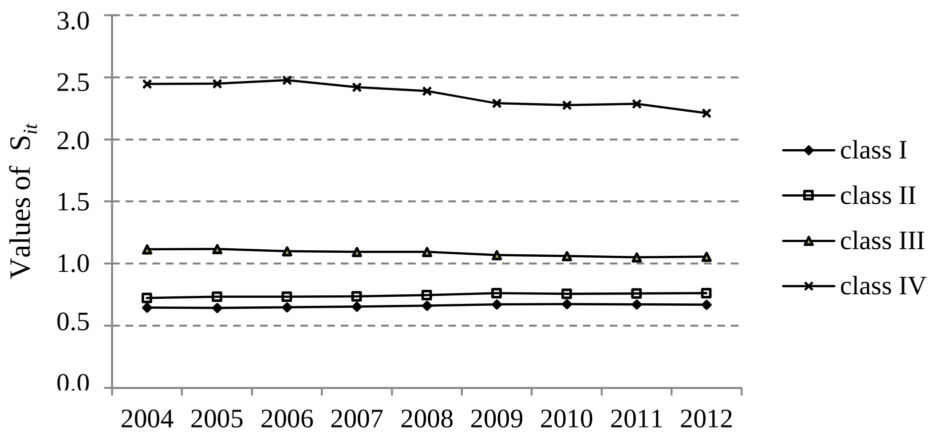
Technical infrastructure	Diagnostic features		Class of rural areas: I-high	II-medium-high	III-medium-low	IV- low	Poland
road network density [km/100 km²]	mean	2004	30.60	31.82	18.17	14.12	21.66
		2012	38.54	38.41	25.45	21.75	29.04
	Average Annual Growth Rate[3] (%)		2.49	1.51	3.28	4.28	2.83
	relative indices w_{ijt}	2004	1.60	1.68	0.82	0.56	1.46
		2012	1.43	1.42	0.83	0.66	1.47
	relative Average Annual Growth Rate[3] (%)		-0.73	-1.87	0.12	1.60	0.23
users of water supply network [%]	mean	2004	75.44	64.77	69.80	76.17	71.32
		2012	78.75	70.53	75.80	80.32	76.20
	Average Annual Growth Rate (%)		0.43	0.91	0.96	0.61	0.73
	relative indices w_{ijt}	2004	1.08	0.88	0.97	1.09	0.98
		2012	1.05	0.90	0.99	1.07	0.99
	relative Average Annual Growth Rate (%)		-0.40	0.23	0.30	-0.17	0.03
users of sewerage network [%]	mean	2004	21.56	19.59	13.39	15.94	17.26
		2012	36.59	32.26	24.99	24.51	29.45
	Average Annual Growth Rate (%)		5.52	5.27	6.65	4.65	5.64
	relative indices w_{ijt}	2004	1.35	1.20	0.71	0.91	1.12
		2012	1.36	1.16	0.83	0.80	1.08
	relative Average Annual Growth Rate (%)		-0.06	-0.35	1.40	-1.10	-0.52
users of gas network [%]	mean	2004	31.77	27.52	9.93	2.45	17.81
		2012	36.55	31.00	14.32	3.79	21.69
	Average Annual Growth Rate (%)		1.14	0.98	3.93	3.90	1.81
	relative indices w_{ijt}	2004	5.77	4.96	1.58	0.14	3.53
		2012	4.46	3.73	1.54	0.16	2.76
	relative Average Annual Growth Rate (%)		-2.26	-2.42	0.52	1.09	-2.06
users serviced by sewage treatment facilities [%]	mean	2004	22.49	20.77	15.17	16.29	18.43
		2012	41.24	36.23	28.07	27.42	33.09
	Average Annual Growth Rate (%)		6.45	6.11	6.86	6.60	6.53
	relative indices w_{ijt}	2004	1.32	1.19	0.78	0.86	1.10
		2012	1.37	1.16	0.83	0.80	1.08
	relative Average Annual Growth Rate (%)		-0.09	-0.50	0.37	0.18	-0.22

Source: as in Table 1.

[3] Average Annual Growth Rate was computed from all elements of the time series [Lira, Wysocki 2004]

The only clear progress was noted in the fourth class, where Sit dropped by 0.23. However, one should notice that the class itself was far from homogenous in terms of rural infrastructure development between 2004 and 2012. Łódzkie and Kujawsko-Pomorskie improved noticeably but the other two provinces showed further deterioration from an already low relative estimates of development.

Figure 1. Values of relative synthetic index S_{it} for relative classes in 2004-2012



Source: as in Table 1.

CONCLUSIONS

The use of dynamic approach to multi-dimensional methods of relative taxonomy facilitated the analysis of the process of smoothing out the differences in development of the rural technical infrastructure between the provinces. These methods proved to be a useful tool in the analysis of changes in the development of infrastructure between individual provinces relative to all others.

The existing disparities in rural technical infrastructure development between the classes of provinces based on 2004 relative synthetic index values were observed to even out slightly between 2004 and 2012. The advantage of the class of high relative development level remained unchanged, but improvement was noticed of the low relative class, spearheaded by two provinces, Kujawsko-Pomorskie and Łódzkie.

APPENDIX

```
# database is expected in the normal (or molten) form with columns
#  + named PERIOD, FEATURE, OBJECT, VAL
# set up the data matrices x
lPERIOD = length(levels(z$PERIOD))
lFEATURE = length(levels(z$FEATURE))
```

```
lOBJECT = length(levels(z$OBJECT))
x = array(z$VAL, dim=c(lOBJECT, lFEATURE, lPERIOD),
 + dimnames = list(levels(z$OBJECT), levels(z$FEATURE),
 + levels(z$PERIOD))
# calculate matrices of relative indices D
D = array(, dim=c(lOBJECT, lOBJECT, lFEATURE, lPERIOD),
 + dimnames = list(levels(z$OBJECT), levels(z$OBJECT),
 + levels(z$FEATURE), levels(z$PERIOD)))
for (t in 1:lPERIOD)
  for (j in 1:lFEATURE)
    D[,,j, t] = (1/x[,j, t]) %*% t(x[,j, t])
# set up the averaging matrix A
A = array(1 / (lOBJECT-1), dim=c(lOBJECT, lOBJECT)); diag(A) = 0
# calculate matrices W of relative indices
W = array(, dim=c(lOBJECT, lFEATURE, lPERIOD),
 + dimnames = list(levels(z$OBJECT), levels(z$FEATURE),
 + levels(z$PERIOD))
for (t in 1:lPERIOD)
  for (j in 1:lFEATURE)
# vectors of W are formed by the diagonal elements of scaled
# D matrices
    W[,j, t] = diag(A %*% D[,,j, t])
# calculate the S matrix: reciprocals of elements of W
# averaged over the diagnostic features
S = apply(1/W, c(1,3), mean)
```

REFERENCES

Local Data Bank, www.stat.gov.pl

Chudy W. (2011) Rozwój infrastruktury obszarów wiejskich, Infrastruktura i Ekologia Terenów Wiejskich, Nr 10/2011, PAN, Kraków, s. 97-106.

Grabiński T. (1984) Wielowymiarowa analiza porównawcza w badaniach dynamiki zjawisk ekonomicznych, Zeszyty Naukowe, Seria specjalna: Monografie, Nr 61, Wydawnictwo Akademii Ekonomicznej w Krakowie, Kraków.

Hydzik P. (2012) Zastosowanie metod taksonomicznych do oceny poziomu rozwoju społeczno-ekonomicznego powiatów województwa podkarpackiego, Zeszyty Naukowe Politechniki Rzeszowskiej Nr 286, Ekonomia i Nauki Humanistyczne z. 19 (2/2012), s. 17-32.

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

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, s. 631-645.

Zeliaś A. (2000) Taksonomiczna analiza przestrzennego zróżnicowania poziomu życia w Polsce w ujęciu dynamicznym, Wydawnictwo Akademii Ekonomicznej w Krakowie, Kraków.