ACTA UNIVERSITATIS LODZIENSIS
FOLIA OECONOMICA 196, 2006

*Agnieszka Ordon**

MULTIVARIATE ANALYSIS OF REGIONAL DIFFERENCES IN THE HIGHER EDUCATION SYSTEM IN POLAND

Abstract. This article discusses an attempt at analysis of regional diversity in Poland in 2001 with respect to the level of the higher educational system.

The first part of this report deals with ranking provinces with respect to the level of the higher educational system, measured by a synthetic variable. This variable is the result of 10 characteristics, weighted according to their influence on higher education. Selection of those characteristics was dictated by their use by experts as well as their availability in regional statistics data.

In the second part of the article, the author presents clusters formed by provinces in two-dimensional areas: the first dimension indicates the level of higher educational system whereas the other dimension describes the socio-economic situation in the regions. This situation is represented by factors that have been singled out (by principal component analysis) as key among 21 characteristics, which potentially influence the higher educational system.

The discovery of commonalities according to which those clusters are formed is the main purpose of the article.

Key words: level of the higher education system, factor analysis, cluster analysis.

1. INTRODUCTION

The condition of the higher education system is becoming a more and more important factor, showing a general social and economic status of both states and regions. The level of educational achievements is considered to be one of the three areas (beside income and life expectancy) from which the components of Human Development Index[1] are derived (see Nowak 2001). It is worth mentioning that the HDI for Poland is still rising: from

* Assistant, Chair of Quantitative Methods in Economics, University of Information Technology and Management in Rzeszów.

[1] Human Development Index – a synthetic measure, used to international comparison of social development.

the 52nd position in 1995 (cf. "An Access to Education" 1998) to the 35th position at present (Glennie, Wóycicka 2002).

The HDI has been also successfully used for regional comparison in the Report on Social Development ("An Access to Education" 1998). The results showed that „the best situation [as regards educational achievements] occurs in those provinces which gather students from colleges and universities". The great role of higher education is also represented by (Czyżewski et al. 2001) results, according to which the level of human knowledge as well as skill is a significant factor in so-called "human capital" which can be interpreted as a production factor.

As it shows, the level of education (especially at a higher level) influences significantly the general socio-economic development of states or regions. Hence this paper discovers the regularities in the relation between socio-economic factors and the level of higher education. It also analyzes clusters of Polish provinces created by those regularities. The research is set in 2001 owing to the fact that this is the last year with both complete and available data in the Polish Official Statistic resources.

The analysis is conducted in the following stages:

1. Ranking of voivodships according to the level of higher education (synthetic variable).

2. Isolating the principal components from among the socio-economic factors which influence the condition of higher education (factor analysis).

3. Classification of provinces at two dimensions: the level of higher education and the individual socio-economic factors (hierarchic cluster analysis).

4. Analysis of the regularities that occurred in the relation between socio-economic factors and the level of higher education.

2. FACTORS DESCRIBING THE LEVEL OF HIGHER EDUCATION

In the theory of econometrics it is not settled which factor characterizes the level of higher education most accurately. That is why the author has chosen 9 characteristics, guided by:

- their application by the experts in higher school rankings,
- possibility of accessing or estimating data,
- necessity of making the data independent from the size of the provinces.

2.1. Data

According to the criteria mentioned above, 10 characteristics were collected which best describe the level of higher education system in each of the regions in 2001.

Y_1 – number of colleges and universities per 1 thousand km^2,
Y_2 – number of colleges and universities per 1 million inhabitants,
Y_3 – number of students per 1 thousand inhabitants,
Y_4 – number of graduates per 1 thousand inhabitants,
Y_5 – percentage of the day study system students,
Y_6 – gross education ratio of students I – percentage of students aged 19–24,
Y_7 – gross education ratio of students II – percentage of first-year students aged 19,
Y_8 – number of academic teachers per 100 students,
Y_9 – relation of the number of Reader and PhD titles conferred in 2001 to the number of academic teachers.

Values of those characteristics for particular provinces are given in Table 1:

Table 1. Variables describing the level of higher education system

Voivodship	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6	Y_7	Y_8	Y_9
Dolnośląskie	1.15	7.74	48.26	7.547	0.47	0.47	0.73	5.05	5.56
Kujawsko-pomorskie	0.67	5.71	35.78	6.981	0.46	0.35	0.57	4.62	4.72
Lubelskie	0.68	7.63	41.07	7.437	0.48	0.41	0.57	5.91	8.18
Lubuskie	0.36	4.88	35.10	6.841	0.42	0.33	0.45	4.15	0.27
Łódzkie	1.15	7.98	42.55	6.793	0.42	0.45	0.63	5.20	5.10
Małopolskie	1.72	8.02	46.86	7.633	0.56	0.46	0.74	6.89	6.31
Mazowieckie	2.25	15.75	65.65	12.27	0.40	0.68	1.07	4.91	4.64
Opolskie	0.53	4.63	32.88	6.147	0.42	0.34	0.54	3.66	3.08
Podkarpackie	0.89	7.51	31.78	5.32	0.43	0.31	0.54	3.60	0.62
Podlaskie	0.64	10.66	40.29	6.781	0.44	0.41	0.61	5.23	4.47
Pomorskie	1.09	9.07	37.87	6.293	0.52	0.37	0.56	6.17	6.01
Śląskie	2.68	6.83	39.87	6.875	0.41	0.40	0.63	4.33	7.15
Świętokrzyskie	0.94	8.34	40.10	8.121	0.29	0.41	0.77	3.56	0.96
Warmińsko-mazurskie	0.33	5.44	35.18	5.714	0.47	0.33	0.51	3.80	4.13
Wielkopolskie	0.94	8.32	41.20	6.935	0.47	0.40	0.67	5.81	6.46
Zachodniopomorskie	0.70	9.23	52.69	11.42	0.43	0.50	0.73	4.37	4.13

Higher education system encompasses both state and private schools, including foreign languages teachers training colleges, where graduates are conferred Bachelor's degree.

The data about first-year students (Y_7) concerns the students of both Bachelor and Master degree studies. The name "academic teachers" refers to teachers employed on such positions: Professor, Reader, Senior Lecturer, Lecturer, Assistant Lecturer, and Instructor. The data about teachers concern those working full-time as well as part-time, counted in terms of the total number of full-time teachers.

Education ratios of students Y_6 and Y_7 are assessed owing to the specificity of higher education system. Students' environment, especially of extramural studies, is very diverse in terms of age. Therefore, it is impossible to establish a fixed range of age to which they belong, as is the case with high school and gimnazjum students. That is why it has been assumed that for schools and universities gross educational ratio shows (in %) the relation of the number of students to the number of people aged 19–24, who should attend a college or a university, on condition that they attended a four-year high school and that they study in a five-year system. Analogously, education ratio of students II is the relation of the number of first-year students to the total number of young people aged 19.

Such selection of characteristics is caused by the necessity of compromise between the data that best describe the level of higher education system and the data available in Polish regional statistics. For that reason such significant factors have been omitted as: the number of faculties, the level of academics.

All the variables $Y_1, \dots Y_9$ are stimulants to the level of higher education system: the higher the values, the higher the level of education. But in the form presented in Table 1 those variables are of different unit, so it would be difficult to compare them. Hence the necessity of normalization.

2.2. Normalization of Variables Describing the Level of Higher Education System

To make the variables $Y_1, \dots Y_9$ comparable and standardize their size, they were normalized (with respect to mean value 0 and standard deviation 1) in the following way:

$$Z_{kj} = \frac{Y_{kj} - \overline{Y}_j}{S_{Y_j}}, \tag{1}$$

where

Y_{kj} – means j^{th} intput value for k^{th} province,
Z_{kj} – means j^{th} standardized value for k^{th} province,
$\overline{Y}_j$ – means the average of Y_j,
S_{Y_j} – means the standard deviation of Y_j.

Thanks to such operation the constant range of variability was achieved: all the variables are situated in the range of (–3; 3). The standardized variables are non-unit, with equal mean 0 and standard deviation 1.

2.3. Construction of the Synthetic Variable Measuring the Level of Higher Education System

After the normalization, a synthetic variable was created, which was to measure the level of higher education system. This variable is the result of 9 standardized characteristics $Z_1, Z_2, ..., Z_9$, weighted according to their influence on other variables (see Grabiński 1992).

Thus to the j^{th} standardized variable Z_j there was the weight ascribed:

$$w_j = \frac{\sum_{i=1}^{9} |r_{ij}|}{\sum_{i,j} |r_{ij}|}, \tag{2}$$

where:

w_j – the weight of j-th standardized variable Z_j,

r_{ij} – coefficient of correlation between Y_i and Y_j.

The weights calculated according to model (1) are presented in Table 2:

Table 2. Weights assigned to the variables $Y_1, ..., Y_9$

Specification	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6	Y_7	Y_8	Y_9
Weights w_i	0.11	0.13	0.14	0.12	0.04	0.14	0.13	0.1	0.09

It is easy to observe that the variables such as: gross education ratio of students and the academic level of teachers have the most significant influence on the level of the higher educational system.

Finally, for a hypothetical k-th province the synthetic variable Y^s assumes the value:

$$Y_k^s = \sum_{j=1}^{10} w_j Z_{kj}, \tag{3}$$

where:

Y_k^s – value of the synthetic variable for a k-th voivodship,

w_j – weight of the j-th standardized variable Z_j,

Z_{kj} – value of the j-th standardized variable Z_j *for* k-th voivodship W.

The values of Y^s, which represent synthetic measures of the level of higher education system, are presented in Table 3. Those results are also shown on the map of Poland (Figure 1): the higher the level of higher education, the darker the colour with which the province is marked.

Table 3. The level of higher education system in the provinces of Poland in decreasing order

Voivodship	Y^s level of higher education system
Mazowieckie	2.06
Małopolskie	0.72
Zachodniopomorskie	0.59
Dolnośląskie	0.37
Wielkopolskie	0.16
Śląskie	0.14
Lubelskie	0.12
Łódzkie	0.07
Pomorskie	0.02
Podlaskie	0.00
Świętokrzyskie	–0.25
Kujawsko-pomorskie	–0.46
Warmińsko-mazurskie	–0.80
Opolskie	–0.88
Podkarpackie	–0.89
Lubuskie	–0.97

Fig. 1. The level of higher education system in the provinces of Poland

When analyzing regional diversity in Poland in 2001 in terms of the level of higher education system it can be observed that:

1. The highest level of higher education system is represented by the following voivodships: mazowieckie and małopolskie. This is confirmed by the rankings of universities, where colleges and universities from Warsaw (mazowieckie) and Kraków (małopolskie) are in the first places.

2. The lowest level of higher education system is represented by the following voivodships: podkarpackie, lubuskie, warmińsko-mazurskie i opolskie.

3. SOCIO-ECONOMIC FACTORS INFLUENCING HIGHER EDUCATION SYSTEM

One of the aims of the article is to discover which socio-economic factors influence the level of higher education. It is known that apart from the factors directly determining the development of higher education (as variables $Y_1, ..., Y_9$ discussed in Section 2) there are also some factors that influence it in a more indirect way, such as industry, unemployment rate or population growth. The selection of those factors as well as the analysis of their influence on higher education is the aim of the following section.

3.1. Data

To find out what kind of socio-economic factors influence the level of the higher education and in what way do they do it, 15 characteristics were gathered which could be connected with higher education:

X_1 – number of living births per 1 thousand inhabitants,

X_2 – population growth per 1 thousand inhabitants (difference between the number of births and the number of deaths divided by the number of inhabitants (given in thousands)),

X_3 – percentage of pre-working age population,[2]

X_4 – percentage of post-working age population,[3]

X_5 – percentage of population having higher education (over 15 years of age),

X_6 – number of employed people per 1 thousand inhabitants,

X_7 – average annual unemployment rate[4] (in %),

X_8 – registered unemployment rate (in %),

X_9 – expenditures on research and developmental activity per capita (in zł),

X_{10} – sold production of industry[5] per capita (in thousand zł),

X_{11} – average monthly gross pay (in zł),

X_{12} – average monthly income in households per capita (in zł),

X_{13} – expenditure of voivodoships' budgets on education (in % of total),

X_{14} – book collection in public libraries (volumes per 1 thousand inhabitants),

X_{15} – number of readers in public libraries per 1 thousand inhabitants.

Values of these characteristics for particular voivodships are presented in Table 4.

---

[2] Pre-working age population – population up to the age of 17.

[3] Post-working age population – males aged 65 and more, females aged 60 and more.

[4] The unemployment rate – relation of the number of unemployed people to the total economically active population.

[5] Sold production of industry – the value of products sold, work and services provided for a payment.

Table 4. Selected variables influencing the level of higher education indirectly

Voidoship	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6	Y_7	Y_8	Y_9	Y_{10}	Y_{11}	Y_{12}	Y_{13}	Y_{14}	Y_{15}
Dolnośląskie	8.7	–0.6	0.21	0.2	10.3	328	23.7	21.5	115	12.1	1973	652	9.01	3698	221
Kujawsko-pomorskie	10	0.9	0.24	0.1	8.6	349	20	21.9	62	12.1	1795	609	9.99	3648	177
Lubelskie	10	–0.2	0.24	0.2	9.7	431	14.8	15.7	66	6.11	1797	553	11.1	3313	203
Lubuskie	9.8	1.4	0.24	0.1	8.5	309	24.2	24.4	17	10	1789	635	5.91	3866	202
Łódzkie	8.5	–3.4	0.21	0.2	9.5	397	19.8	18.1	113	11.9	1783	645	13.3	3410	180
Małopolskie	10.5	1.6	0.24	0.1	10.5	401	13	14.1	142	10.5	1867	630	11.1	3271	206
Mazowieckie	9.3	–0.8	0.22	0.2	14.2	455	14.6	13	422	20.2	2682	771	8.92	3236	174
Opolskie	8.2	–0.2	0.22	0.1	8.3	326	18	18.2	36	11.6	1865	630	9.52	3919	170
Podkarpackie	10.3	1.9	0.26	0.1	8.8	435	18	17.4	48	9.14	1755	527	12.1	4085	191
Podlaskie	9.5	0	0.25	0.2	9.5	398	16	15.1	73	7.34	1787	583	9.28	3531	162
Pomorskie	10.7	2.6	0.24	0.1	11.3	326	18.5	19.6	92	13.5	1991	683	7.17	2937	172
Śląskie	8.4	–1.1	0.22	0.1	9.2	349	19.7	15.7	84	17.4	2094	689	10.3	3464	220
Świętokrzyskie	9.3	–0.8	0.23	0.2	9.4	440	18	18.4	15	8.32	1815	557	5.14	3340	166
Warmińsko-mazurskie	10.4	2.3	0.25	0.1	8.7	299	23.6	28.9	35	8.85	1797	579	8.89	3720	201
Wielkopolskie	10.3	1	0.24	0.1	9.6	389	17.6	15.4	100	15.3	1898	637	6.31	3507	187
Zachodnio-pomorskie	9.5	1.1	0.23	0.1	10.4	309	22.4	24.7	37	9.69	1897	643	8.42	4095	199

3.2. Remarks

Data: X_{1-4}, X_{6-15} are taken from *Statistical Annual of the Regions – Poland* (2002),
where:

$X_{1-4}, X_6, X_8, X_{14,15}$ represent situation at the end of the year (December 31st, 2001),

X_7, X_{9-13} describe the situation in the whole year 2001.

Owing to the rarity of adequate research, the data: X_5 come from the National Population Census 2002. Despite the fact that the Census was carried out in May 2002, the data obtained represent the situation in 2001. The explanation is that the majority of students graduate in June, so the number of the higher educated among the examined population approximates to their number at the end of the school year 2000/2001.

The selection of those socio-economic factors is guided both by their possible connection with higher education as well as their availability (the same as $Y_1, \dots Y_9$) in the regional statistics data.

3.3. Reduction of the Number of Variables

To gain a better understanding of the variables in a data set by detecting which variables are related to one another and to identify how they are related, they were reduced to a smaller number of so called "principal components", independent of each other (Grabiński 1992).

To ensure that it is appropriate to run a factor analysis of the data, sampling adequacy has been examined. It is measured by the Kaiser-Meyer-Olkin (KMO) statistics using SPSS program. Since the overall KMO appeared too low, the variables with the lowest individual KMO statistics (the diagonal elements on the anti-image correlation matrix) had to be dropped. Those were: X_{13}, X_{14} and X_{15}.

To the remaining 12 variables the following methods were applied (all of them using SPSS program): the factor analysis with the principal component method, VARIMAX rotation and Kaiser's normalization.

As a result, eigenvectors λ_i, $i = 1, \dots, 12$ were obtained as well as eigenvalues w_i, $i = 1, \dots, 12$, determining what part of the total variance of variables X_1–X_{12} is represented by the individual components: F_1–F_{12}.

The choice of the principal components number was guided by Cattel's criterion (cf. Ostasiewicz 1998) which says that the best set of components is corresponded with the point in the scree plot, where the slope of eigenvalues is getting gentle. As the following scree plot shows, for these data it is best to accept 3 principal components (Table 5 and Figure 2).

Table 5. Median vectors for two-dimensional samples

Criterion	PD1		PD2		PD3		PD4	
	DWM	no.	DWM	no.	DWM	no.	DWM	no.
L	(–0.025, 0.370)	2	(0.299, 0.669)	16	(0.390, 0.975)	3	(–0.633, –0.743)	21
M	(–0.025, 0.370)	2	(0.299, 0.669)	16	(–0.508, –0.385)	21	(–0.222, 0.066)	16
B	(–0.105, 0.043)	2, 20	(0.753, 1.053)	18, 23	(0.390, 0.975)	3	(–0.633, –0.743)	21
T	(–0.025, 0.370)	2	(0.299, 0.669)	16	(–0.304, 0.234)	2, 3, 15, 21	(–0.633, –0.743)	21
R boundary	(–0.184, 0.370)	–	(0.402, 0.669)	–	(–0.304, 0.221)	15	(–0.633, –0.528)	–

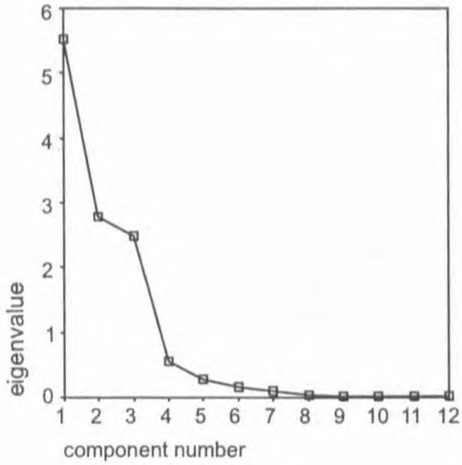

Fig. 2. Scree plot

The following matrix (Table 6) presents rotated components, in other words normalized eigenvectors (Niemczyk 2001), reduced to three selected principal components sorted according to the values of the factor loadings.

Table 6. Rotated component matrix

	Input characteristics	Component		
		1	2	3
X_{11}	average monthly gross pay	0.937	0.191	–0.117
X_{12}	average monthly income in households per capita	0.931	–0.142	–0.241
X_{10}	sold production of industry per capita	0.881	–0.039	–0.190
X_9	expenditures on research and developmental activity per capita	0.855	0.423	–0.069
X_5	percentage of population having higher education	0.839	0.353	0.071
X_6	number of employed people per 1 thousand inhabitants	0.037	0.947	–0.038
X_7	average annual unemployment rate	–0.147	–0.907	–0.168
X_8	registered unemployment rate	–0.262	–0.854	0.141
X_4	percentage of post-working age population	0.018	0.763	–0.590
X_1	number of living births per 1 thousand inhabitants	–0.068	0.074	0.953
X_2	population growth per 1 thousand inhabitants	–0.051	–0.313	0.924
X_3	percentage of pre-working age population	–0.412	0.091	0.884

The values presented in Table 6 can also be interpreted as coefficients of correlation between the principal components $F_{i,\ i=1,2,3}$, and initial values $X_{j,\ j=1,\ldots,12}$.

1. The first of the principal components (F_1) is correlated (positively) with 5 variables: $X_5, X_9, X_{10}, X_{11}, X_{12}$. Taking into consideration their character, the first component should be defined as *the economic standing* of voivodships.

2. The second principal component (F_2) is positively correlated with the percentage of post-working age population and the number of employed people, and negatively correlated with both kinds of unemployment rate, so *the job market* seems to be the proper name for it.

3. Since the third component is described by typical demographic variables (X_1, X_2, X_3), thus it is referred to as *the demographic potential* of the voivodships.

3.4. Classification of Provinces

One of the aims of the article is the detection of some regularities which occur with in the provinces with respect to the extracted principal components and the level of higher education (Y^s), determined in the first section. To achieve this aim, the division of provinces into clusters was applied through:

- assessment of the scatter plot,
- interpretation of the dendrograms obtained by applying hierarchical cluster analysis.

The first step to obtain the scatter plot was calculating the realizations of the principal components $F_{i,\ i=1,2,3}$ for each of 16 provinces. Here I used the fact that each principal component may be interpreted as a linear combination of standardized variables $X_1 - X_{12}$ with elements of the rotated component matrix (Table 6) used as coefficients (cf. Grabiński 1992, Ostasiewicz 1998):

$$F_i = a_{1i}Z_1 + a_{2i}Z_2 + \ldots + a_{12i}Z_{12} = \sum_{j=1}^{12} a_{ji}Z_j, \tag{4}$$

where

F_i – i-th principal component ($i = 1, 2, 3$),

a_{ji} – an element situated in an i-th column ($i = 1, 2, 3$) and in a j-th line ($j = 1, \ldots, 12$) of the rotated component matrix,

Z_j – j-th standarized input variable X_j (the process of standarization was conducted as in Section 2).

Realizations of the principal components $F_{i,\ i=1,2,3}$ obtained this way, along with the values of the level of the higher educational system (Y^s) obtained in the first chapter are presented in Table 7.

Table 7. Realizations of the principal components F_1, F_2, F_3 and the values of Y^s – the level of higher education system

Voivodship	F_1	F_2	F_3	Y^s
Dolnośląskie	0.36	–0.60	–1.00	0.49
Kujawsko-pomorskie	–0.45	–0.63	0.48	–0.61
Lubelskie	–0.67	1.01	0.23	0.15
Lubuskie	–0.66	–1.46	0.57	–1.27
Łódzkie	0.19	0.53	–1.87	0.10
Małopolskie	0.16	0.92	0.75	0.95
Mazowieckie	3.17	2.09	–1.31	2.70
Opolskie	–0.22	–0.29	–0.82	–1.15
Podkarpackie	–0.98	0.28	1.23	–1.16
Podlaskie	–0.55	0.74	0.09	0.00
Pomorskie	0.46	–0.46	1.11	0.03
Śląskie	0.80	–0.09	–1.26	0.19
Świętokrzyskie	–0.64	0.65	–0.45	–0.33
Warmińsko-mazurskie	–1.01	–1.70	1.41	–1.05
Wielkopolskie	0.21	0.16	0.58	0.20
Zachodniopomorskie	–0.18	–1.16	0.26	0.77

On the basis of the values presented in Table 7 three scatter plots were created (Figures 3a–5a) which present the distribution of provinces in two-dimensional space: of the principal components $F_{i,\ i=1,2,3}$ and of the synthetic variable Y^s representing the level of higher education system. They show clusters of provinces which result from the regularities occurring among them according to the two dimensions. It can be seen that the visual assessment is confirmed by the clusters formed by provinces in the hierarchical cluster analysis. The process of formation of those clusters is shown in dendrograms (Figures 3b–5b) below each of the scatter plots. The clusters obtained were marked in the scatter plots with red bordering:

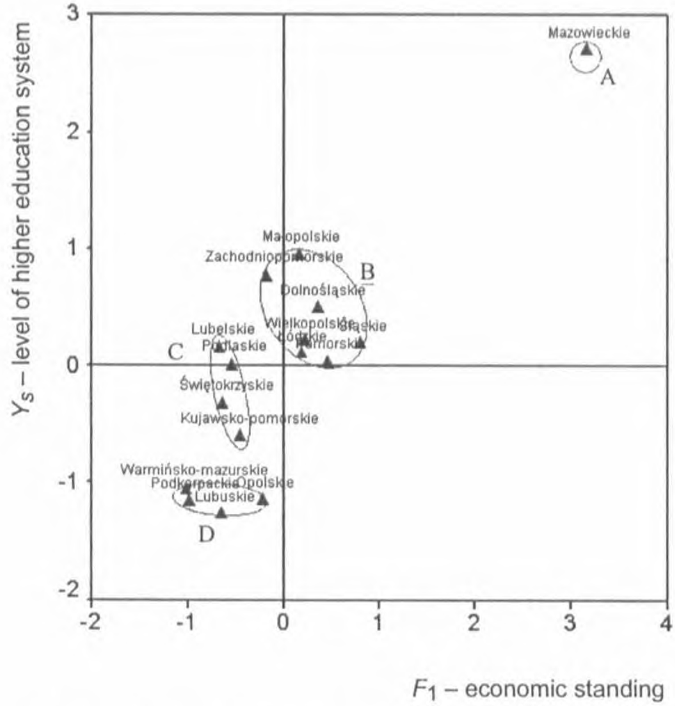

Fig. 3a. The distribution of the provinces according to the economic standing of the provinces (F_1) and the level of higher education system (Y^s)

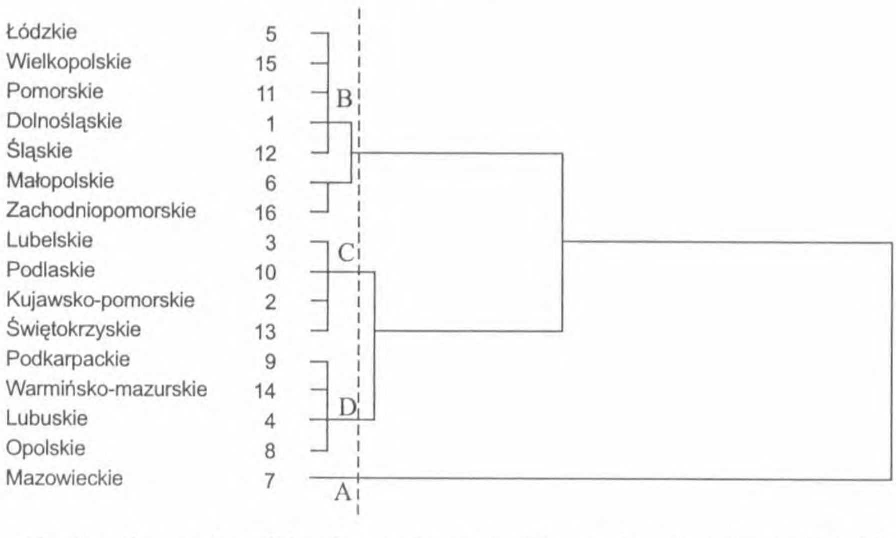

Fig. 3b. The clusters formed by the provinces according to the economic standing of the provinces (F_1) and the level of higher education system (Y^s)

In the graphs above, one can easily notice that the provinces form four clusters according to the economic standing of provinces (F_1) and the level of higher education system (Y^s). These are:

- cluster A: very high level of higher education, very high economic standing,
- cluster B: high level of higher education, high economic standing,
- cluster C: medium level of higher education, low economic standing,
- cluster D: low level of higher education, low economic standing.

The arrangement of clusters A–D seems to confirm the view that the level of education is converged with the level of economic development. This phenomenon can be observed when comparing different countries: the higher the gross domestic product value is, the better educated people are and vice versa ("An Access to Education" 1998). A similar regularity can be observed when comparing different provinces: voivodships in A and B clusters are characterized by good economic standing as well as by the high level of education. Meanwhile, in case of B and D clusters weaker economy is accompanied by lower level of higher education system.

The sensation is mazowieckie province, which is characterized both by the highest level of education and the best economic conditions. However, this fact does not surprise as the capital city is considered to be both cultural and industrial center of a country.

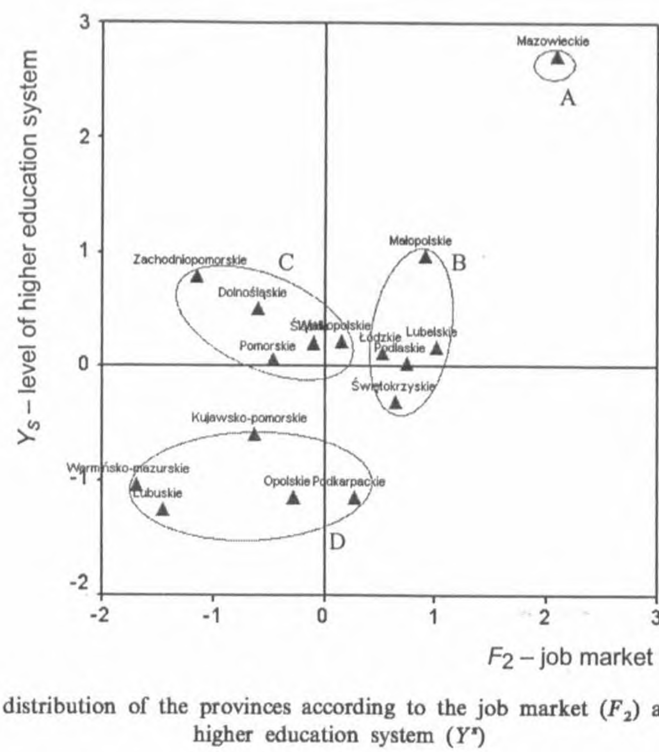

Fig. 4a. The distribution of the provinces according to the job market (F_2) and the level of higher education system (Y^s)

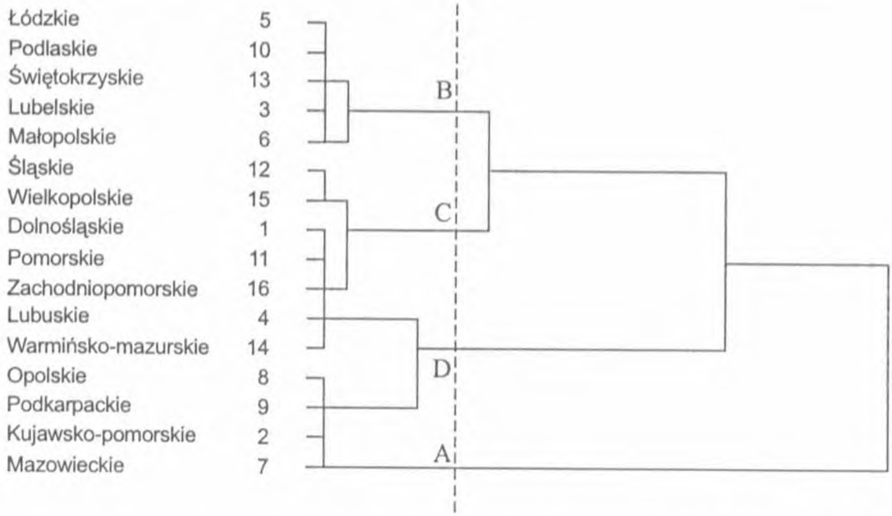

Fig. 4b. The clusters formed by the provinces according to the job market (F_2) and the level of higher education system (Y^s)

In the space of the job market (F_2) and the level of higher education system (Y^s) the objects (voivodships) are more dispersed and it is difficult to mark off any particular clusters visually. The hierarchical cluster analysis (Ward's method, square Euclidean distance) proved to be particularly useful at this point. The clusters obtained are presented in Figure 4b. They marked with red bordering on the scatter plot Figure 4a.

- cluster A: very high level of education, good situation on the job market,
- cluster B: high level of education, quite good situation on the job market,
- cluster C: high level of education, bad situation on the job market,
- cluster D: low level of education, rather bad situation on the job market.

Such a wide dispersion of the provinces in the space of the job market (F_2) and the level of higher education system (Y^s) indicates a weak correlation between those two factors, although it would appear that this relationship is strong. It is common knowledge that highly educated, qualified and specially trained employees are always in great demand, and that such people graduate mostly from schools with high level of education. The case of mazowieckie (having the highest values of both factors), and warmińsko-mazurskie provinces (the lowest values of both factors), would seem to confirm the view that high level of education plays a significant role in unemployment reduction.

Amongst the remaining provinces, though, such regularities cannot be observed. The reason might be the fact that the job market (at least when comparing provinces) is more determined by some other factors, not necessairly connected with higher education (e.g. the liquidation of state-owned farms in northern voivodships or the gradual liquidation of the coal industry in Silesia).

In the scatter plot (Figure 5a) presenting the distribution of the provinces in the space of the demographic potential (F_3) and the level of the higher education system (Y^s), the clusters of objects are more noticeable. Figure 5b confirms the visual assessment, so the following division was made.

- cluster A: high level of education, high demographic potential,
- cluster B: high level of education, very low demographic potential,
- cluster C: low level of education, high demographic potential,
- cluster D: diverse level of education, very low demographic potential.

Voivodships from A and C clusters are the most demographicly buoyant and very diverse as far as the level of education goes. Groups B and D are also educationally diverse, being simultaneously characterized by drastically low demographic rates (łódzkie province in particular with the lowest negative natural increase per 1 thousand inhabitants).

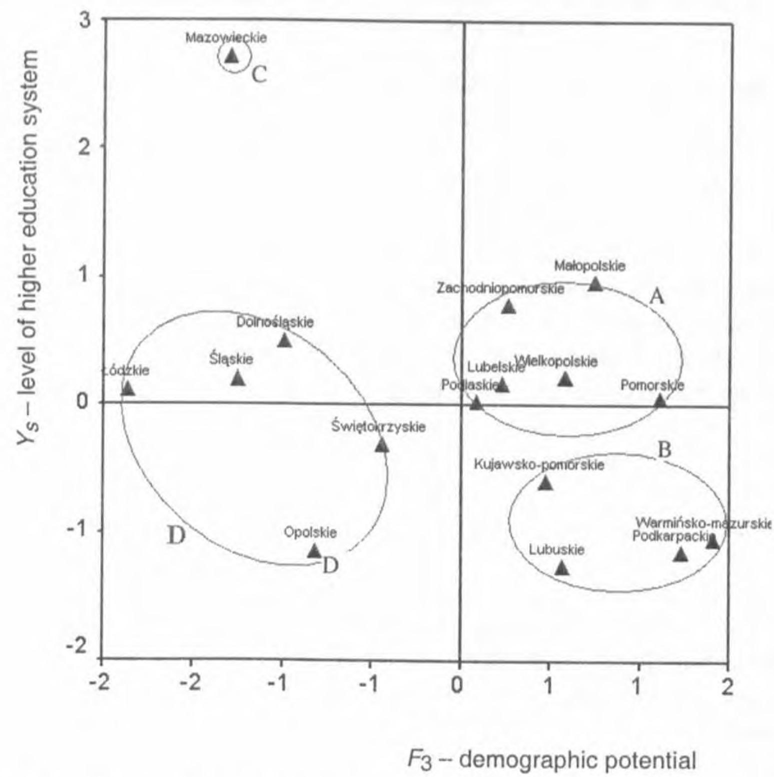

Fig. 5a. The distribution of the provinces according to the demographic potential (F_3) and the level of higher education system (Y^s)

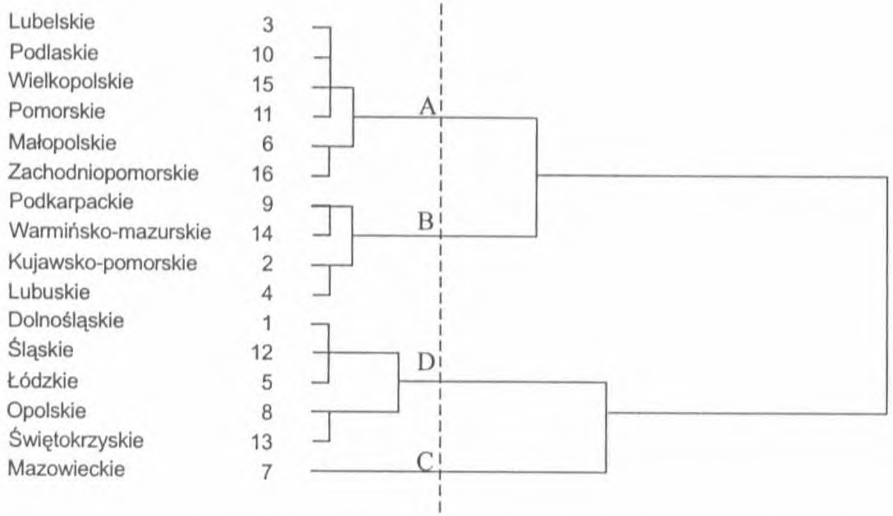

Fig. 5b. The clusters formed by the provinces according to the job market (F_2) and the level of higher education system (Y^*)

One can draw a conclusion that there is no significant causal relationship between the level of higher education system and the demographic potential of the voivodships (although the number of students is obviously determined by the population growth).

There is quite an opposite regularity presented in "An Access to Education" (1998). It shows that the voivodships (in the previous administrative division: 49 provinces, 1998) which are highly ranked in terms of the level of education have at the same time very low demographic rates and vice versa. However, in this research, maybe as a result of the administrative changes (16 voivodships, 2001) such a regularity cannot be traced.

In order to compare the provinces with respect to the three extracted factors, they were presented in Figure 6 in a form of columns against the background of the level of higher education (see Figure 1). For a better comparison, the factors F_1, F_2, F_3 were normalized with minimum value as a point of reference, and with the range as the scaling factor. Thanks to this operation the realizations of the factors F_1, F_2, F_3 take values from the range $[0, 1]$, with no change in the hierarchy of the provinces. The values of the main factors have no statistical interpretation, they only show:

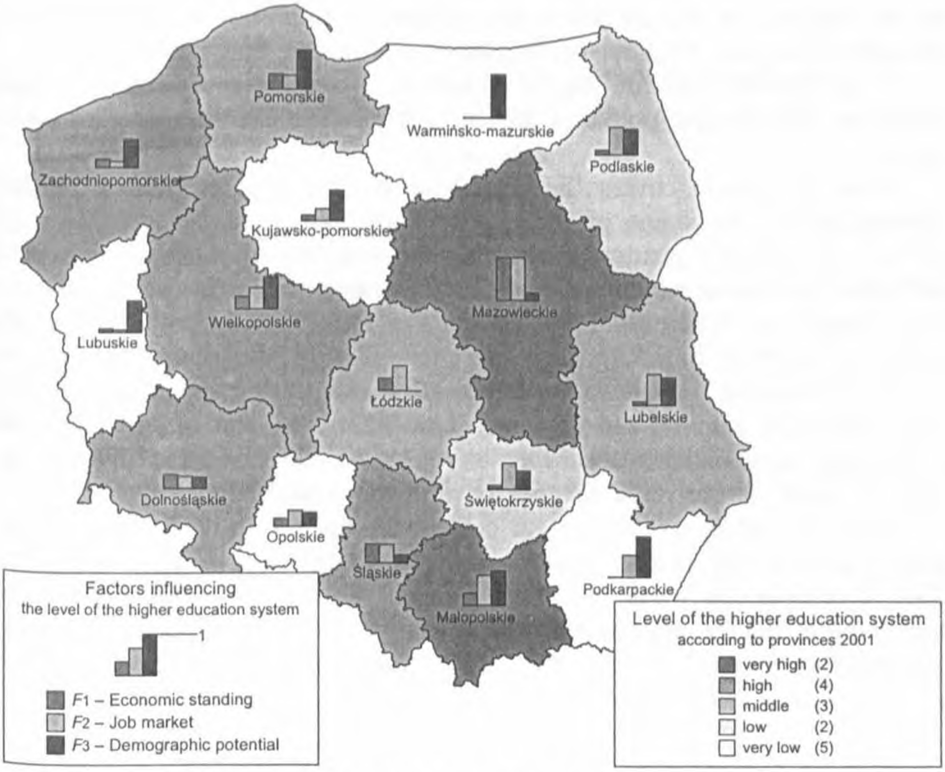

Fig. 6. Factors influencing the level of higher education system

- values closer to 1 (higher columns) – to advantage of a phenomenon represented by a given factor,
- values closer to 0 (low columns or no columns) – to disadvantage of a phenomenon represented by a given factor.

The detailed interpretation is presented (in the notes) in Figure 6.

4. CONCLUSIONS

The research shows that Poland is a very diverse area in terms of the level of higher education system. This phenomenon is a consequence of various factors, both geographical (location, natural resources) and socio-economic (job market, demographic potential, economic standing). Using the synthetic variable Y^s, the picture of the spatial location of the provinces with different level of higher education system was obtained.

1. The highest level of higher education system is represented by the following voivodships: mazowieckie and małopolskie. This result is confirmed

by the rankings of universities, where colleges and universities from Warsaw (mazowieckie) and Kraków (małopolskie) are in the first positions.

2. The lowest level of higher education system is represented by the following voivodships: podkarpackie, lubuskie, warmińsko-mazurskie i opolskie.

Since the article covers the period of one year, it is only a partial observation of the shape of higher education system in the provinces of Poland. A complete picture of not only the level but also the development of higher education would be possible if a similar research was conducted with respect to dynamics. Data covering a longer period of time would allow to improve both accuracy and reliability of the results. However, a task of this kind encounters varies obstacles such as: the limited availability and resourcefulness of the regional data base and the impossibility of converting some indices according to the new administrative division in case the time of the research goes beyond the year 1999.

There should be an attempt at comparing the results with those of an analogous research, based on the available data from years 1999–2003. Such a comparison would probably provide valuable new observations on the shape of higher education system in the voivodships of Poland. Further continuation as well as deepening of this topic can become the subject of another works in this field.

REFERENCES

Czyżewski A. B., Góralczyk-Modzelewska M., Saganowska E., Wojciechowska M. (2001), "Regionalne zróżnicowanie kapitału ludzkiego w Polsce", *Studia i Prace Zakładu Badań Statystyczno-Ekonomicznych*, **277**.

"Dostep do edukacji" ("An Access to Education") (1998), *Raport o rozwoju społecznym Polska* (*Repoet ou Social Development*), Ośrodek Informacji ONZ, Warszawa, http://www.unic.un.org.pl/nhdr/1998/index.php

Glennie C., Wóycicka I. (2002), "Milenijne cele rozwoju", *Raport o rozwoju społecznym*, Ośrodek Informacji ONZ, Warszawa, http://www.unic.un.org.pl/mcrp/

Grabiński T. (1992), *Metody taksonometrii*, Akademia Ekonomiczna, Kraków.

Niemczyk A. (2001), "Poziom rozwoju społecznego w nowym układzie administracyjnym Polski", *Przegląd Statystyczny*, **3**.

Nowak L. (2000), "Metodologia wyznaczania wskaźnika rozwoju społecznego (HDI)". In: *Wskaźniki statystyczne, Raporty UNDP* (Program Narodów Zjednoczonych ds. Rozwoju), http://www.undp.org.pl/pliki/nhdr2001/IXrozdzial.pdf

Ostasiewicz W. (1998), *Statystyczne metody analizy danych*, Akademia Ekonomiczna, Wrocław.

Raports of the National Population Census in Poland (2002), Central Statistical Office, Warsaw.

Statistical Annual of the Regions – Poland (2001), Central Statistical Office, Warsaw.

Statistical Annual of the Regions – Poland (2002), Central Statistical Office, Warsaw.

"Szkoły wyższe i ich finanse" (2001), *Informacje i opracowania statystyczne*, Warszawa.

Agnieszka Ordon

ANALIZA WIELOWYMIAROWA RÓŻNIC REGIONALNYCH W POLSCE POD WZGLĘDEM SZKOLNICTWA WYŻSZEGO

(Streszczenie)

Artykuł jest próbą analizy zróżnicowania regionalnego Polski w 2001 r. pod względem szkolnictwa wyższego.

W pierwszej części artykułu opisany jest proces hierarchizacji województw pod względem poziomu szkolnictwa wyższego. Za miarę tego poziomu przyjęto zmienną syntetyczną, będącą średnią ważoną z 10 cech mogących świadczyć o poziomie szkolnictwa wyższego. Przy wyborze tych cech kierowano się zarówno ich zastosowaniem przez ekspertów, jak i dostępnością danych w statystyce regionalnej.

W drugiej części artykułu autor przedstawia grupowanie województw w dwuwymiarowych przestrzeniach, gdzie pierwszy wymiar to poziom szkolnictwa wyższego, drugi natomiast opisuje sytuację społeczno-gospodarczą w regionach. Sytuację tę reprezentują czynniki wyodrębnione w analizie głównych składowych spośród 21 cech mających potencjalny wpływ na kształtowanie się szkolnictwa wyższego.

Celem artykułu jest wykrycie prawidłowości, według których tworzą się te skupiska.