Folia Oeconomica
Acta Universitatis Lodziensis
ISSN 0208-6018 e-ISSN 2353-7663
www.czasopisma.uni.lodz.pl/foe/
4(343) 2019

DOI: http://dx.doi.org/10.18778/0208-6018.343.14

Iwona Bąk
Westpomeranian University of Technology in Szczecin, Faculty of Economics
Department of Mathematical Applications in Economy, ibak@zut.edu.pl

Katarzyna Wawrzyniak
Westpomeranian University of Technology in Szczecin, Faculty of Economics
Department of Mathematical Applications in Economy, kwawrzyniak@zut.edu.pl

Maciej Oesterreich
Westpomeranian University of Technology in Szczecin, Faculty of Economics
Department of Mathematical Applications in Economy, moesterreich@zut.edu.pl

The Application of Statistical Methods to Identify Factors Determining Employment Effectiveness in District Labour Offices in Poland

Abstract: In Poland, institutions that perform public tasks in the field of the labour market are state employment services, including District Labour Offices (PUPs). They try to restore the unemployed to the open labour market using active forms of their activation. The activities of PUPs in the field of customer activation are subject to annual evaluation. According to the "Act on Employment Promotion and on Labour Market Institutions", four indicators were selected for the evaluation of the functioning of labour offices in 2015–2017. Among them, the employment effectiveness index is very important, as its level allows us to assess to what extent active forms of activation of the unemployed implemented by PUPs contribute to the return of the unemployed to employment. This assessment is so important that in Poland as well as in other countries research has been conducted on both active forms of activation of the unemployed and employment policies.

The aim of the research whose results are presented in the article was to identify factors influencing the employment effectiveness index achieved by District Labour Offices (PUPs) operating in Poland in 2016. Selected statistical methods were used, including correlation and regression analysis as well as multidimensional correspondence analysis. It turned out that the methods applied did not make it possible to unambiguously identify factors which had significantly affected the employment effectiveness index calculated on the basis of all forms of activation. It may turn out that clearer relations could be identified if this indicator was considered separately for particular forms of activation.

Keywords: correlation analysis, stepwise regression, zeroed unitarisation, multidimensional correspondence analysis, employment effectiveness, District Labour Offices

JEL: C38, C51, J49

1. Introduction

In Poland, in accordance with the Act of 20 April 2004 on Employment Promotion and on Labour Market Institutions (Journal of Laws 2017.1065, Chapter 3, Article 6), institutions that perform public tasks in the scope of the labour market include: public employment services, Volunteer Labour Corps, employment agencies, training institutions, and social dialogue institutions. Public employment services are created by employment bodies together with district and voivodship labour offices, an office servicing the minister competent for labour issues and voivodship offices carrying out tasks defined by law. The detailed scope of their tasks is provided by the Act in Article 4.

At the district level, institutions dealing with researching and analysing the labour market, providing information to unemployed people and professional agency for job seekers are District Labour Offices (PUP).

In accordance with Act on Employment Promotion and on Labour Market Institutions Article 109 (7j), four indicators were designated to assess the functioning of District Labour Offices (PUPs) in 2015–2017. Two of the indicators concern the PUP personnel potential and the extent of its use (the percentage indicator of employees of the district labour office performing the function of customer adviser in the total employment and the indicator of the number of unemployed persons per employee of the district labour office working as customer advisers), and the other two indicate the efficiency of employment (the employment efficiency index of the basic forms of activation) and the cost (the cost-effectiveness index of the basic forms of activation) of the PUP's operation.

Among the above-mentioned indicators, the employment efficiency index is very important because its level allows us to assess the extent to which the active forms of activation of the unemployed implemented by PUPs contribute to the reinstatement of the unemployed to the open labour market. The index should be understood as a percentage of the number of people who during or after participation in the basic forms of reactivation have been shown as employed in relation to the number of people who have completed their participation in the basic forms of reactivation. This assessment is so important that in numerous studies (reports) organisational and legal aspects as well as methodological aspects related to the measurement of the effectiveness of services and instruments of the labour market in Poland have been presented (Błędowski, 2008; Wiśniewski, Zawadzki, 2011; Maksim, Wiśniewski, 2012; *Analiza rozwiązań*…, 2017). Similar research concerning various European countries has been presented, among others, in the following works: L. Calmfors, A. Forslund and M. Hemström (2001), J. Kluve (2006), H. Lehmann and J. Kluve (2008), D. Card, J. Kluve and A. Weber (2009), G. Guzmán (2014), E. Rollnik-Sadowska (2014), V. Escudero (2018), A. Bánociová and S. Martinková (2017) as well as in reports of the European Commission (*European Semester: Thematic*…, 2017). Particularly worth mentioning, both in terms of methodology and scope of analysis, are the results of the research project "Study on the effectiveness of ALMPs" (Kluve et al., 2005) conducted for the European Commission, concerning the analysis of the effectiveness of active policies in the labour markets in such countries as: the Netherlands, Sweden, Austria, Germany, Italy, Denmark, Estonia, Poland, Spain, France, and the United Kingdom.

The diverse level of employment efficiency index in individual PUPs urges us to identify the causes of this state of affairs, and thus to answer the question about the factors which cause a situation when in some labour offices its level is high while in others it is low.

The article presents the results of research the objective of which was to identify the factors affecting the employment efficiency index achieved by District Labour Offices (PUPs) in Poland in the year 2016. The research hypothesis was formulated: the effectiveness of employment is a comprehensive resultant of interacting factors that characterise the economic, social and demographic situation in the area of PUPs' operation and their personnel potential. In order to verify this hypothesis, a research process consisting of three stages was employed. In the first stage, the strength and the direction of the relationship between the employment efficiency index and the diagnostic features adopted for the study were investigated. The next step was to identify the mechanism of the links between the employment efficiency index and the diagnostic features. However, in the final stage, multiple correspondence analysis was used in order to detect links between the categories of employment efficiency and the categories of synthetic measures calculated on the basis of the value of diagnostic features in the following areas: unemployment, demography, environment, entities, and the personnel potential of the PUP.

2. Characteristics of statistical data

The surveyed group consisted of 340 District Labour Offices (PUPs) operating in Poland in 2016. The information on the employment efficiency index and indicators characterising the potential of particular PUPs comes from the Ministry of Family, Labour and Social Policy. On the other hand, statistical data on features characterising the socio-economic situation in the areas of the functioning of labour offices were taken from the Local Data Bank of the Central Statistical Office. The potential diagnostic features that may affect the efficiency of employment are grouped into five areas: unemployment, demography, environment, economic entities, the personnel potential of the PUP, and their names along with the determination of their nature[1] is given in Table 1[2].

Table 1. Names and nature of indicators used in the study

Area	Indicators	Properties of indicators
Unemployment	X_1 – the number of unemployed with tertiary education in the total number of unemployed in %	Destimulants
	X_2 – the number of unemployed with tertiary education in the working age population in %	
	X_3 – the number of youth unemployed (24 years and less) in the total number of unemployed in %	
	X_4 – the number of unemployed aged 24 and less in the working age population in %	
	X_5 – the number of unemployed over 12 months in the total number of unemployed in %	
	X_6 – the number of unemployed from 12 to 24 months in the total number of unemployed in %	
	X_7 – the number of unemployed over 24 months in the total number of unemployed in %	
	X_8 – the number of unemployed women in the total number of unemployed in %	
	X_9 – the number of unemployed per 100 working age people	
	X_{10} – the registered unemployment rate in %	
	X_{11} – the number of job vacancies per 100 unemployed	Stimulants
Demography	X_{12} – the number of people of post-working age per 100 persons of working age	Destimulants
	X_{13} – the number of people of pre-working age per 100 people of working age	
	X_{14} – the number of people of post-working age to the number of people of pre-working age	
	X_{15} – people of post-working age in the total population in %	
	X_{16} – people of pre-productive age people in the total population in %	Stimulants
	X_{17} – people of working-age in the total population in %	

1 Stimulants – higher values of a given diagnostic feature indicate a higher level of the studied phenomenon. Destimulants – higher values of a given diagnostic feature indicate a lower level of the studied phenomenon.

2 The decision to adopt such a set of diagnostic features was made having taking into account the results of the CATI survey conducted among the directors of labour offices or persons designated by them (*Badanie ilościowe*..., 2018).

Area	Indicators	Properties of indicators
Environment	X_{18} – crimes detected by the Police per 1000 residents	Destimulants
	X_{19} – the number of low-qualified schools[a)] per 10,000 residents	Stimulants
	X_{20} – the number of graduates of low-qualified schools per 10,000 residents	
	X_{21} – the number of high-qualified schools per 100,000 residents	
	X_{22} – the number of highly qualified school graduates per 10,000 residents	
	X_{23} – the number of training institutions per 10,000 residents[b)]	
	X_{24} – (the number of employees in entities over 9 persons to the total population) · 100%	
	X_{25} – (the number of employees in entities over 9 persons to the number of working age people) · 100%	
	X_{26} – (the estimated number of employees[c)] to the total population) · 100%	
	X_{27} – (the estimated number of employees to the number of working-age people) · 100%	
	X_{28} – local and district roads with hard surface per 100 km^2	
	X_{29} – taxes on natural persons per capita in PLN	
	X_{30} – corporation taxes per capita in PLN	
Economic entities	X_{31} – the number of economic entities in the REGON register per 10,000 population	Stimulants
	X_{32} – the number of national economy entities per 10,000 working age people	
	X_{33} – the number of economic entities employing 0–9 persons per 10,000 working-age people	
	X_{34} – the number of economic entities employing 10–49 persons per 10,000 working-age people	
	X_{35} – the number of economic entities employing 50–249 persons per 10,000 working-age people	
	X_{36} – the number of economic entities employing 250 persons and more per 10,000 working-age people	
	X_{37} – the number of public sector entities per 10,000 working-age people	
	X_{38} – the number of private sector entities per 10,000 working-age people	
	X_{39} – the number of agriculture sector entities per 10,000 working-age people	
	X_{40} – the number of economic entities in the industry and construction sector per 10,000 working-age people	
	X_{41} – the number of economic entities in the sector other activities per 10,000 working-age people	

Area	Indicators	Properties of indicators
PUP staff potential	X_{42} – the number of client advisers in the total employment in PUPs in %	Stimulants
	X_{43} – the number of vocational counsellors in the total employment in PUPs in %	
	X_{44} – the number of unemployed per 1 employee of PUP	Destimulants
	X_{45} – the number of unemployed per one key worker of PUP	
	X_{46} – the number of unemployed per 1 employment agent in PUP	
	X_{47} – the number of unemployed per 1 vocational counsellor in PUP	
	X_{48} – the number of unemployed per one client adviser in PUP	

[a] Schools that educate students to obtain the Matura exam (secondary school final exam), i.e. schools: primary schools, lower secondary school, basic vocational school, general secondary schools, secondary vocational schools.
[b] The data come from the Register of Training Institutions – www.stor.praca.gov.pl [accessed: 20.09.2018].
[c] The estimated number of employees was calculated as follows: the number of economically active – the number of registered unemployed, where the number of economically active people was calculated as the ratio of the number of unemployed and the registered unemployment rate multiplied by 100%.

Source: own research based on data from the Local Data Bank of the Central Statistical Office of Poland (GUS) and the Ministry of Family, Labour and Social Policy (MRPiPS)

Phase I – the study of the correlation between the employment efficiency index and the diagnostic features

The study of the relationship between the employment efficiency index and the diagnostic features presented in Table 1 began with the calculation of Pearson's correlation coefficients. Due to the fact that the study is conducted as a full study, while identifying diagnostic features affecting employment efficiency in PUPs, it was decided that significant (explicit) relationships should be considered in the case of those for which the absolute value of the correlation coefficient was at least 0.3. In order to justify the accepted correlation coefficient level in Table 2, there are presented sample correlation coefficient values intervals along with the determination of the strength commonly discussed in the literature. Table 2 shows that according to three authors the relationship between the characteristics is at least average (clear) when the correlation coefficient exceeds 0.3.

Table 3 contains the correlation coefficients between the employment efficiency index and the diagnostic features of each area. It leads to the conclusion that no diagnostic feature had a clear impact on the level of employment efficiency achieved in the analysed labour offices in 2016.

Table 2. The strength of dependence between the diagnostic features for the sample ranges of values for the Pearson correlation coefficient

According to J.P. Guilford		According to A. Góralski		According to L.A. Gruszczyński		According to K. Zając	
$r < 0.2$	Weak correlation	$r = 0$	No correlation	$0 < r \leq 0.2$	Very weak correlation, unclear	$r \leq 0.3$	Unclear correlation
$0.2 < r < 0.4$	Low correlation	$0 < r < 0.1$	Slight correlation	$0.2 < r \leq 0.3$	Weak correlation	$0.3 < r < 0.5$	Average correlation
$0.4 < r < 0.7$	Average correlation	$0.1 < r < 0.3$	Weak correlation	$0.3 < r \leq 0.5$	Clear correlation with medium intensity	$r \geq 0.5$	Clear correlation
$0.7 < r < 0.9$	High correlation	$0.3 < r < 0.5$	Average correlation	$0.5 < r \leq 0.7$	High correlation		
$0.9 < r < 1.0$	Very high correlation	$0.5 < r < 0.7$	High correlation	$0.7 < r < 1.0$	Very high correlation		
		$0.7 < r < 0.9$	Very high correlation				
		$0.9 < r < 1.0$	Almost full correlation				
		$r = 1.0$	Full correlation				

Source: own elaboration based on Guilford, 1964: 157; Góralski, 1976: 71; Zając, 1982: 298; Gruszczyński, 1986: 168

Table 3. Pearson correlation coefficients between the employment effectiveness index and the diagnostic features from individual areas

Unemployment		**Demography**		**Environment**		**Economic entities**		**PUP staff potential**	
X_1	–0.042	X_{12}	–0.079	X_{18}	–0.152	X_{31}	–0.101	X_{42}	0.040
X_2	–0.141	X_{13}	0.082	X_{19}	0.161	X_{32}	–0.104	X_{43}	–0.006
X_3	0.148	X_{14}	–0.091	X_{20}	0.165	X_{33}	–0.104	X_{44}	–0.166
X_4	–0.085	X_{15}	–0.081	X_{21}	–0.102	X_{34}	–0.083	X_{45}	–0.206
X_5	–0.109	X_{16}	0.086	X_{22}	–0.130	X_{35}	–0.066	X_{46}	–0.134
X_6	–0.082	X_{17}	0.040	X_{23}	–0.129	X_{36}	–0.041	X_{47}	–0.061
X_7	–0.101			X_{24}	0.114	X_{37}	–0.042	X_{48}	–0.176
X_8	0.042			X_{25}	0.106	X_{38}	–0.101		
X_9	–0.164			X_{26}	0.082	X_{39}	0.139		
X_{10}	–0.178			X_{27}	0.072	X_{40}	0.038		
X_{11}	0.004			X_{28}	–0.189	X_{41}	–0.138		
				X_{29}	–0.161				
				X_{30}	–0.116				
				X_{31}	0.114				

Source: own calculations based on data from the Local Data Bank of GUS and MRPiPS

Phase II – The Identification of the linking mechanism between the employment efficiency index and the diagnostic features

Despite the lack of explicit relationship between the employment efficiency index achieved by the labour offices and the individual diagnostic features, an attempt was made to estimate the linear regression function using the forward stepwise regression[3]. The results of the estimation are provided in Table 4.

Table 4 shows that out of 48 diagnostic features only nine were included in the set of explanatory variables, only four of which (selected in the table in italics) had a significant impact on the level of employment efficiency. The estimated function explained 17.25% of the development of the studied phenomenon and it was the best function due to the highest values of *F*-statistics, the level of which is important when using stepwise regression. In the next step, the relevant explanatory variables were removed and the linear regression function was re-estimated (Table 5).

3 The essence of and the procedure for conducting stepwise regression can be found, for example, in the following papers: Stanisz, 2007: 143–159; Podolec, Ulman, Wałęga, 2008: 66–71.

Table 4. The results of linear regression function estimation of employment efficiency (EE) for all examined diagnostic features (explanatory variables are ordered by the decreasing value of *t*-statistics)

Explanatory variables	Coefficients	Std. error	*t*-ratio	*p*-value
Constant	91.022	5.915	15.389	0.000
X_{10}	–0.445	0.121	–3.662	0.000
X_{28}	–0.017	0.007	–2.579	0.010
X_{20}	0.031	0.013	2.453	0.015
X_{45}	–0.021	0.010	–2.199	0.029
X_{16}	–0.606	0.314	–1.930	0.054
X_{21}	–0.181	0.117	–1.540	0.125
X_{1}	–0.171	0.114	–1.502	0.134
X_{18}	–0.107	0.073	–1.467	0.143
X_{19}	0.401	0.296	1.355	0.176
$R = 0.4153$; $R^2 = 0.1725$; $F(9.330) = 7.6413$; $p < 0.0000$; $Se = 6.8577$				

Source: own calculations based on data from the Local Data Bank of GUS and MRPiPS

Table 5. The results of linear regression function estimation of employment efficiency (EE) for significant explanatory variables from the first model (explanatory variables are ordered by the decreasing value of *t*-statistics)

Explanatory variables	Coefficients	Std. error	*t*-ratio	*p*-value
Constant	81.427	3.134	25.986	0.000
X_{28}	–0.028	0.006	–4.864	0.000
X_{10}	–0.284	0.100	–2.833	0.005
X_{20}	0.026	0.009	2.759	0.006
X_{45}	–0.022	0.009	–2.360	0.019
$R = 0.3693$; $R^2 = 0.1364$; $F(4.335) = 13.223$; $p < 0.0000$; $Se = 6.9532$				

Source: own calculations based on data from the Local Data Bank of GUS and MRPiPS

After the re-assessment, the function was obtained which improved the significance of explanatory variables and the value of *F*-statistic. Other measures characterising the quality of the model deteriorated. Although the estimated function explains only 13.64% of the researched relationship, it is important to note that the most important impact on the efficiency of employment has been exerted by: communication accessibility (X_{28} – local and district roads with hard surface per the 100 km^2) unemployment (X_{10} – registered unemployment rate), the level of education of people in the pre-productive age (X_{20} – the number of graduates of low-qualified school per 10,000 residents), and the workload of the PUP personnel related to the number of unemployed (X_{45} – the number of unemployed per one key employee of a PUP).

Phase III – the detection of links between the categories of employment efficiency and the synthetic measure categories

The first two stages of the research, which used the actual values of the employment efficiency index and individual diagnostic features, did not allow us to identify the factors affecting the employment efficiency index achieved by the examined District Labour Offices in 2016. Therefore, it was decided to replace the real values of diagnostic features – characterising the five highlighted areas in the study (Table 1) – with synthetic measures determined for each area separately. For that purpose, the method of zeroed unitarisation (Kukuła, 2000: 60–92) was used, in which the normalisation of diagnostic features was preformed according to the formulas:

1) for stimulants:

$$z_{ij} = \frac{x_{ij} - \min_i x_{ij}}{\max_i x_{ij} - \min_i x_{ij}}, \tag{1}$$

2) for destimulants:

$$z_{ij} = \frac{\max_i x_{ij} - x_{ij}}{\max_i x_{ij} - \min_i x_{ij}} \tag{2}$$

and the synthetic measure (z_i) was calculated as the arithmetic mean of the normalised values of the diagnostic features:

$$z_i = \frac{1}{k}\sum_{j=1}^{k} z_{ij}, \tag{3}$$

where:

x_{ij} – value of the *j*-th diagnostic feature in the *i*-th object,

$\min_i x_{ij}$ – the minimum value of the *j*-th diagnostic feature,

$\max_i x_{ij}$ – the maximum value of the *j*-th diagnostic feature,

$\max_i x_{ij} \neq \min_i x_{ij}$

$i = 1, 2, \ldots, n$, n – the number of objects,

$i = 1, 2, \ldots, k$, k – the number of diagnostic features.

A synthetic measure (z_i) takes the values in the range [0, 1]. The closer the value of z_i is to unity, the situation i of the object is better from the point of view of k diagnostic features.

Then, correlation coefficients were calculated between the employment efficiency index (EE) and the values of synthetic measures determined for individual areas whose values were as follows:

1) for the unemployment area: 0.133,
2) for the demography area: 0.075,
3) for the environment area: 0.020,
4) for the economic entities area: 0.074,
5) for the PUP personnel potential area: 0.167.

The obtained correlation coefficients show the lack of a clear link between the employment efficiency index and the synthetic measures of the highlighted areas. This means that the replacement of individual diagnostic features with one synthetic measure also did not answer clearly the question which socio-economic factors had influenced the level of employment efficiency achieved by the examined labour offices in 2016.

Finally, in order to identify the factors that affect the level of employment efficiency, it was decided to use multiple[4] correspondence analysis, despite the fact that the use of this method requires the transformation of characteristics expressed on an interval scale or ratio scale onto a nominal or ordinal scale through their categorisation (Hand, 1996; Greenacre, 2007). The loss of information as a result of the transformation of the characteristics on the lower scale does not limit the applicability of this method in practice, which has been proven by numerous studies in which multiple correspondence analysis has been applied to identify the factors that affect the investigated phenomenon. For example, A. Ezzari and P. Verme (2012) have examined with the use of correspondence analysis the factors affecting poverty in Morocco in the late 2001–2007. I. Bąk and K. Wawrzyniak (2009), using correspondence analysis, have identified the factors that affected the choice of types of trips made by retirees and pensioners in 2005, while B. Batóg, M. Mojsiewicz and K. Wawrzyniak (2009) have used this method to answer the question of which factors impact households' decision to buy the III pillar insurance. In turn, K. Dębkowska and J. Kilon (2014) have analysed the factors affecting the development of enterprises in the e-commerce sector in the Podlasie Voivodeship, and K. Cheba and J. Hołub-Iwan (2014) have used correspondence analysis to examine the segmentation of the medical services market in Poland, using the data obtained from surveys on a sample of 1067 adult Poles. Multiple correspondence analysis is also used successfully in studies related to the labour market. For example, G. Autiero, B. Bruno and F. Mazzotta (2000) have analysed the effectiveness of actions taken by the institutions related to the labour market of 19 countries of the world in the years 1989–1994, using the characteristics related

4 Multiple correspondence analysis is a development of classical correspondence analysis which allows the analysis of more than two diagnostic features (Greenacre, Blasius, 2006). Its important advantage is an opportunity to present the results in a graphic form, facilitating the interpretation of the obtained results (Stanimir, 2005: 22, 76–81).

to the protection of the labour market and the coordination of institutions controlling this market with other actors. In contrast, B. Batóg and J. Batóg (2016) have applied correspondence analysis in order to determine the factors affecting the time of being unemployment. For that purpose, they used 6 features characterising the length of unemployment, sex, seniority, age, the level of education and the type of work undertaken by the unemployed registered in the District Labour Office in Szczecin in 2014–2015.

Having decided to apply multiple correspondence analysis in order to identify the factors that affect the level of employment efficiency achieved by the examined labour offices in 2016, the following variables were adopted to provide the basis for the analysis:

1) Employment Efficiency Index (EE),
2) a synthetic measure for the unemployment area (U),
3) a synthetic measure for the demography area (D),
4) a synthetic measure for the environment area (E),
5) a synthetic measure for the economic entities area (C),
6) a synthetic measure for the area concerning the PUP personnel potential (PP).

These variables were subjected to transformation by replacing their actual values with ordinal categories designated with the use of three means method[5]. Thanks to this method, for each variable, four categories were obtained, the first of which (1) is associated with the highest values of the variable and the fourth one (4) – with the lowest. The categories of individual variables are marked with the following symbols:

1) Employment Efficiency Index (EE), EE1 (the highest efficiency of employment), EE2, EE3, EE4 (the lowest efficiency of employment),
2) a synthetic measure for the unemployment area (U), U1 (the best situation in terms of unemployment), U2, U3, U4 (the worst situation in terms of unemployment),
3) a synthetic measure for the demography area (D), D1 (the best demographic situation), D2, D3, D4 (the worst demographic situation),
4) a synthetic measure for the environment (E), E1 (the best situation in terms of the environment), E2, E3, E4 (the worst situation in terms of the environment),
5) a synthetic measure for economic entities (C), C1 (the best situation in terms of economic entities), C2, C3, C4 (the worst situation in terms of economic entities),
6) a synthetic measure for the area of the PUP personnel potential (PP), PP1 (the best personnel potential), PP2, PP3, PP4 (the worst personnel potential).

5 The procedure in the three variables method used in order to determine the four typological groups of objects based on the value of the synthetic measure can be found, for example, in the following work (Nowak, 1990: 93).

Due to the number of variables ($Q = 6$) and the number of their categories ($J_q = 4$), the Burt matrix, which is a starting point in correspondence analysis, had the dimensions 24 × 24. While the actual dimension of the co-occurrence space was $K = 18$ and was calculated pursuant to the formula:

$$K = \sum_{q=1}^{Q} \left(J_q - 1 \right), \tag{4}$$

where:
J_q – the number of categories of diagnostic features q ($q = 1, 2, \ldots, Q$),
Q – the number of variables.

Table 6 contains, for the actual space of coexistence, the value of singularities (γ_k), the eigenvalues (λ_k), the degree of explanation of total inertia (λ) by eigenvalues for the k-dimension ($\lambda_k/\lambda \cdot 100\%$) and the degree of explanation of total inertia by eignevalues in the k-dimensional space ($\tau_k \cdot 100\%$).

In order to examine to what extent the eigenvalues of the lower-dimensional space explain the total inertia, the Greenacre criterion (Greenacre, 1994; Stanimir, 2005) was used. In accordance with this criterion, the eigenvalues which are larger than $1/Q$ shall be deemed relevant for the study. In view of the fact that the number of variables in the study is 6, the eigenvalues higher than 0.1667 shall be considered relevant. Table 6 shows that the eigenvalues that meet this condition are assigned to $k \leq 7$, whereas in the 7-dimensional space the degree of explanation of total inertia is 53.14%. However, in the two- and three-dimensional spaces, which are important due to the possibility of graphical presentation of correspondence analysis results, the degree of explanation of total inertia is 20.14% and 27.39%, respectively.

Table 6. Singular values, eigenvalues and the degree of explanation of total inertia by eigenvalues for the *k*-th dimension and for the *k*-dimensional space in the actual projection space

K	Singular values (γ_k)	Eigenvalues (λ_k)	(λ_k/λ) · 100%	τ_k · 100%
1	0.5929	0.3515	11.7165	11.7165
2	0.5028	0.2528	8.4261	20.1426
3	0.4663	0.2174	7.2463	27.3889
4	0.4596	0.2112	7.0399	34.4289
5	0.4439	0.1971	6.5693	40.9982
6	0.4358	0.1900	6.3320	47.3302
7	0.4175	0.1743	5.8111	53.1413
8	0.4069	0.1656	5.5194	58.6607
9	0.4003	0.1602	5.3409	64.0017
10	0.3977	0.1582	5.2733	69.2750
11	0.3928	0.1543	5.1433	74.4183

K	Singular values (γ_k)	Eigenvalues (λ_k)	$(\lambda_k/\lambda)\cdot 100\%$	$\tau_k \cdot 100\%$
12	0.3902	0.1522	5.0739	79.4922
13	0.3660	0.1340	4.4661	83.9584
14	0.3629	0.1317	4.3908	88.3491
15	0.3378	0.1141	3.8037	92.1528
16	0.3208	0.1029	3.4310	95.5838
17	0.2932	0.0860	2.8664	98.4502
18	0.2156	0.0465	1.5498	100.0000
		$\lambda = \sum_{k=1}^{K=18} \lambda_k = 3.0000$	100.0000	

Source: own research based on data from the Local Data Bank of GUS and MRPiPS

In order to improve the quality of the mapping of results in spaces of lower dimensions, the modification of eigenvalues was applied according to the Greenacre proposal (Stanimir, 2005) using the following transformation:

$$\tilde{\lambda}_k = \left(\frac{Q}{q-1}\right)^2 \cdot \left(\sqrt{\lambda_{B,k}} - \frac{1}{Q}\right)^2, \tag{5}$$

where:
Q – the number of variables,
$\lambda_{B,k}$ – the k-th eigenvalue ($k = 1, 2, \ldots, K$),
$\left(\sqrt{\lambda_{B,k}} = \gamma_{B,k}\right)$, $\gamma_{B,k}$ – the k-th singular value of B (Burt) matrix.

For $k \leq 7$, after applying Greenacre modifications, eigenvalues and the degree of explanation of total inertia by eigenvalues for the k-th dimension and for the k-dimensional space were presented in Table 7.

Table 7 shows that the modification has significantly improved the quality of mapping and, in the case of three-dimensional space, the degree of explanation of total inertia by eigenvalues doubled from 27.39% to 56.32%. In order to confirm that the three-dimensional space is sufficient for the graphical presentation of the coexistence of the categories of studied variables, a diagram of eigenvalues was drawn up (Figure 1) and the "elbow" criterion[6] was used, on the basis of which it was found that this dimension was appropriate.

6 The elbow criterion as the basis for choosing the number of dimensions of the co-existence space of the categories of variables was discussed in the paper (Stanimir, 2005: 85–86).

Table 7. Modified eigenvalues with measures characterising the degree of explanation of total inertia by eigenvalues for $\boldsymbol{k} \leq 7$

$\boldsymbol{k}$	$\tilde{\lambda}_k$	$\tilde{\lambda}_k / \tilde{\lambda} \cdot 100\%$	$\tilde{\tau}_k \cdot 100\%$
1	0.2616	26.6180	26.6180
2	0.1627	16.5541	43.1722
3	0.1292	13.1518	56.3240
4	0.1235	12.5711	68.8950
5	0.1107	11.2655	80.1606
6	0.1043	10.6175	90.7781
7	0.0906	9.2219	100.0000
	$\tilde{\lambda} = \sum_{k=1}^{7} \tilde{\lambda}_k = 0.9827$	100.0000	

Source: own elaboration based on Table 6

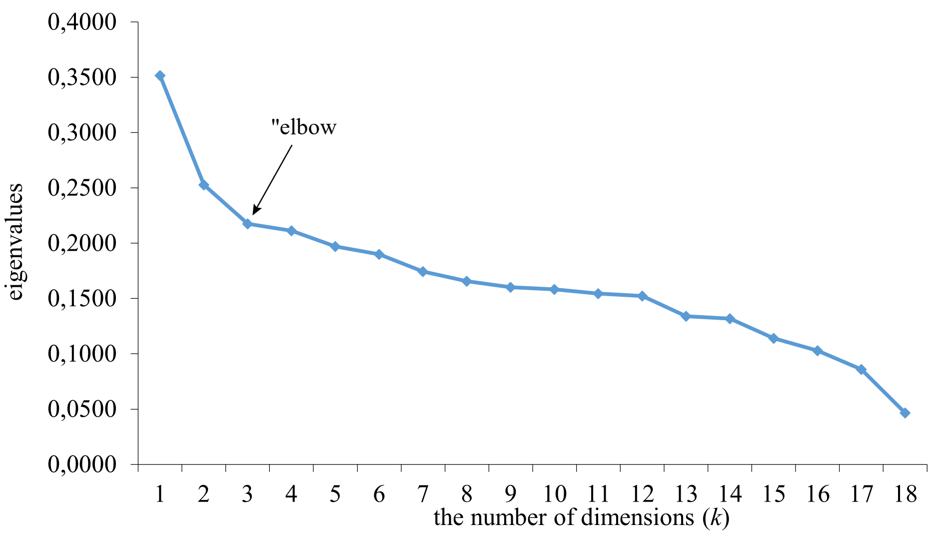

Figure 1. Plot of eigenvalues – elbow criterion

Source: own elaboration based on Table 6

The visualisation of multiple correspondence analysis results in three-dimensional space is possible, however, the interpretation of these results is difficult due to a large number of analysed variables and their categories. Therefore, in order to achieve a more meaningful interpretation, Ward's[7] method was implemented, which enabled the grouping of categories of tested variables pursuant to the mod-

7 The Ward method is one of agglomeration classification methods. It is used in empirical research both in terms of object and characteristics classification. In this method, the distance between groups is defined as the module of the difference between the sums of squares of the

ified values of three coordinates that specify their location in three-dimensional space. The new (modified) values of coordinates in three-dimensional space for different categories of variables are calculated according to the following formula:

$$\tilde{F} = F^{*} \cdot \tilde{A}^{-1} \cdot \tilde{\tilde{E}}, \tag{6}$$

where:
$\tilde{F}$ – a matrix (size 24 × 3) of new coordinates for variable categories,
F^{*} – a matrix (size 24 × 3) of the original coordinates for variable categories,
Γ^{-1} – a diagonal inverse matrix (size 3 × 3) of singular values,
$\tilde{\tilde{E}}$ – a diagonal matrix (size 3 × 3) of modified eigenvalues (dimension).

The results of combining the categories of studied variables into classes using the Ward method were presented in Figure 2.

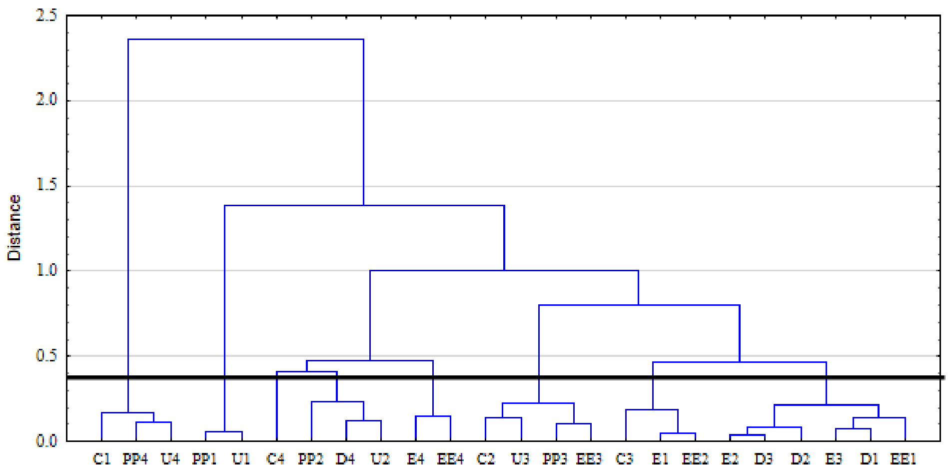

Figure 2. Classification of categories of analysed variables using the Ward method
Source: own elaboration

The horizontal line marked on the graph at the height of a binding equal 0.4109[8] indicates the stage at which combining classes was aborted. In this way, eight classes that contain the following categories of analysed variables were obtained:

distance of points from the means of the groups to which these points belong (Pociecha et al., 1988: 83; Malina, 2004: 62–63; Młodak, 2006: 72; Balicki, 2009: 276–277).

8 For that purpose, the measure proposed by T. Grabiński was used (Stanisz, 2007: 141–142; Panek, 2009: 120–123).

1) class I: C1, PP4, U4,
2) class II: PP1, U1,
3) class III: C4,
4) class IV: PP2, D4, U2,
5) class V: E4, EE4,
6) class VI: C2, U3, PP3, EE3,
7) class VII: C3, E1, EE2,
8) class VIII: E2, D3, D2, E3, D1, EE1.

The first four classes do not contain any category highlighted for employment efficiency, so due to the objective of the study, the co-occurrence of defined variables categories in those classes will not be interpreted. However, on the basis of the coexistence of the variable categories in classes V to VIII, the following relationships can be observed:

1) class V – the lowest values of the employment efficiency index (EE4 – employment efficiency below 76.90%) were achieved in the labour offices where the worst situation in terms of environmental characteristics was observed;
2) class VI – characterised by employment efficiency between 76.90% and 80.42% (EE3) was determined by a fairly difficult situation in terms of characteristics regarding unemployment and the PUP personnel potential as well as an acceptable situation in the area of economic entities;
3) class VII – characterised by employment efficiency between 80.43% and 84.66% (EE3) was determined by an excellent situation in terms of the characteristics regarding the environment and a relatively difficult situation in the area of economic entities;
4) class VIII – the highest values of the employment efficiency index (EE1 – employment efficiency equal to at least 84.67%) were influenced predominately by the demographic situation and conditions related to the environment.

3. Conclusions

In order to verify the research hypothesis adopted in the article, different statistical methods were applied to determine the strength and direction of the relationship between the employment efficiency index and the examined set of diagnostic features characterising the situation relating to: unemployment, demography, environment, economic entities and the personnel potential of District Labour Offices in the area of PUPs' operation. It appeared that the analysis of correlation and regression did not identify the factors which had significantly affected the employment efficiency index. However, it needs to be emphasised that the use of regression function indicated four factors (out of 48) significantly affecting the examined index, although the quality of the estimated function was low.

In the next step, an attempt was made to replace the individual diagnostic features by one synthetic measure (zeroed unitarisation method) calculated for the highlighted areas. However, also this solution could not provide an answer to the question which areas had clearly affected the level of employment efficiency achieved by the examined labour offices in 2016.

Looking for regularities in the relationship between the employment efficiency index and socio-economic factors as well as the PUP personnel potential, it was decided to use multi correspondence analysis. It was found out that the coexistence of 24 variable categories adequately characterised three-dimensional space. Due to a large number of analysed categories, the interpretation of results in that space was difficult. It was therefore decided to use the Ward method, whose final results enabled to find a relationship between the categories of employment efficiency and the categories of synthetic measures of the highlighted areas. It appeared that only in the case of low employment efficiency index (EE3 and EE4), it was possible to indicate relatively clear dependencies, namely that low employment efficiency is a consequence of a bad situation in terms of the characteristics regarding the environment, unemployment and the PUP personnel potential. The identification of factors affecting the Employment Efficiency Index of labour offices is associated with many problems concerning, first of all, a very diverse socio-economic situation in the area of functioning of labour offices and the maladjustment of the PUP personnel potential to the situation in the labour market.

In conclusion, it should be noted that the lack of a clear link between the employment efficiency index and the characteristics selected for examination may be a consequence of the fact that the analysis was limited to the employment efficiency calculated on the basis of the six basic forms of activation of unemployed persons in total. Perhaps clearer dependencies could be identified if employment efficiency was considered separately for each individual form and therefore further studies in this field should continue in this direction.

References

Analiza rozwiązań, o których mowa w art. 108 ust. 1i ustawy o promocji zatrudnienia i instytucjach rynku pracy (2017), Ministerstwo Rodziny, Pracy i Polityki Społecznej, Departament Funduszy, Wydział Wydatków Fakultatywnych i Analiz Efektywności, Warszawa, https://www.gov.pl/web/rodzina/analiza-rozwiazan-o-ktorych-mowa-w-art108-ust–1l-ustawy-o-promocji--zatrudnienia-i-instrumentach-rynku-pracy [accessed: 25.09.2018].

Autiero G., Bruno B., Mazzotta F. (2000), *A Correspondence Analysis of Labour Market Institutions*, "CELPE Discussion Papers", no. 57, pp. 7–42, http://www.unisa.it/uploads/2551/57_dp.pdf [accessed: 24.06.2019].

Badanie ilościowe w powiatowych urzędach pracy. Raport badawczy do projektu pt. "Wypracowanie metodologii i wdrożenie monitorowania efektywności zatrudnieniowej podstawowych form aktywizacji zawodowej bezrobotnych w okresie dłuższym niż 12 miesięcy od zakończe-

nia działań urzędu pracy" (2018), Stowarzyszenie Czas Przestrzeń Tożsamość, Grupa BST Sp. z o.o., Katowice.

Balicki A. (2009), *Statystyczna analiza wielowymiarowa i jej zastosowania społeczno-ekonomiczne*, Wydawnictwo Uniwersytetu Gdańskiego, Gdańsk.

Bánociová A., Martinková S. (2017), *Active Labour Market Policies of Selected European Countries and Their Competitiveness*, "Journal of Competitiveness", vol. 9, issue 3, pp. 5–24, http://doi.org/10.7441/joc.2017.03.01h.

Batóg B., Batóg J. (2016), *Application of correspondence analysis to the identification of the influence of features of unemployed persons on the unemployment duration*, "Economic and Business Review", vol. 2(16), no. 4, pp. 25–44, http://doi.org/10.18559/ebr.2016.4.2.

Batóg B., Mojsiewicz M., Wawrzyniak K. (2009), *Badanie rynku ubezpieczeń III filara z zastosowaniem analizy korespondencji*, [in:] K. Jajuga, M. Walesiak (eds.), *Taksonomia 16. Klasyfikacja i analiza danych – teoria i zastosowania*, "Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu", no. 47, pp. 149–156.

Bąk I., Wawrzyniak K. (2009), *Zastosowanie analizy korespondencji w badaniach związanych z motywami wyboru rodzajów wyjazdów turystycznych przez emerytów i rencistów w 2005 r.*, [in:] K. Jajuga, M. Walesiak (eds.), *Taksonomia 16. Klasyfikacja i analiza danych – teoria i zastosowania*, "Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu", no. 47, pp. 324–332.

Błędowski P. (ed.) (2008), *Efektywność usług i instrumentów rynku pracy służących podnoszeniu kwalifikacji bezrobotnych w Polsce i wybranych krajach Unii Europejskiej*, Ministerstwo Pracy i Polityki Społecznej, Departament Rynku, Warszawa.

Calmfors L., Forslund A., Hemström M. (2001), *Does active labour market policy work? Lessons from the Swedish experiences*, "Swedish Economic Policy Review", no. 85(2001), pp. 61–124, https://www.government.se/contentassets/d61c0e625bf54a54b59decdcb0bd6acd/lars-calmfors-anders-forslund--maria-hemstrom-does-active-labour-market-policy-work-lessons-from-the-swedish-experiences, [accessed: 26.04.2019].

Card D., Kluve J., Weber A. (2009), *Active Labor Market Policy Evaluations: A Meta-Analysis*, "IZA Discussion Paper Series", no. 4002, http://ftp.iza.org/dp4002.pdf [accessed: 26.04.2019].

Cheba K., Hołub-Iwan J. (2014), *Wykorzystanie analizy korespondencji w segmentacji rynku usług medycznych*, "Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu", no. 328, Taksonomia 23, pp. 230–237.

Dębkowska K., Kilon J. (2014), *Application of correspondence analysis to the identification of the determinants of the economic condition of the e-services sector enterprises in Podlaskie region*, "Ekonomia i Zarządzanie", no. 1, pp. 38–56, http://yadda.icm.edu.pl/yadda/element/bwmeta1.element.ekon-element–000171315133 [accessed: 26.04.2019].

Escudero V. (2018), *Are active labour market policies effective in activating and integrating low-skilled individuals? An international comparison*, "IZA Journal of Labor Policy", no. 7: 4, https://doi.org/10.1186/s40173–018–0097–5.

European Semester: Thematic factsheet – Active labour market policies (2017), https://ec.europa.eu/info/sites/info/files/european-semester_thematic-factsheet_active-labour-market-policies_en_0.pdf [accessed: 26.04.2019].

Ezzari A., Verme P. (2012), *A Multiple Correspondence Analysis Approach to the Measurement of Multidimensional Poverty in Morocco, 2001–2007*, "Policy Research Working Paper", no. 6087, World Bank, Washington.

Góralski A. (1976), *Metody opisu i wnioskowania statystycznego w psychologii*, Państwowe Wydawnictwo Naukowe, Warszawa.

Greenacre M. (1994), *Multiple and Joint Correspondence Analysis*, [in:] M. Greenacre, J. Blasius (eds.), *Correspondence Analysis in Social Sciences. Recent Developments and Applications*, Academic Press, San Diego, pp. 141–161.

Greenacre M. (2007), *Correspondence Analysis in Practice*, 2nd Edition, Chapman & Hall, London.

Greenacre M., Blasius J. (2006), *Multiple Correspondence Analysis and Related Methods*, Chapman & Hall, London.

Gruszczyński L.A. (1986), *Elementy statystyki dla socjologów*, Uniwersytet Śląski, Katowice.

Guilford J.P. (1964), *Podstawowe metody statystyczne w psychologii i pedagogice*, Państwowe Wydawnictwo Naukowe, Warszawa.

Guzmán G. (2014), *How effective are active employment policies to reduce unemployment in EU countries*?, "Atlantic Review of Economics", vol. 2, https://dialnet.unirioja.es/descarga/articulo/5318291.pdf [accessed: 26.04.2019].

Hand D.J. (1996), *Statistics and Theory of Measurement*, "Journal of the Royal Statistical Society", ser. A, no. 3, pp. 445–492, http://doi.org/10.2307/2983326.

Kluve J. (2006), *The Effectiveness of European Active Labor Market Policy*, "IZA Discussion Paper Series", no. 2018, http://ftp.iza.org/dp2018.pdf [accessed: 26.04.2019].

Kluve J., Fertig M., Jacobi L., Nima L., Schaffner S., Schmidt Ch., Card D., Góra M., Jensen P., Leetmaa R., Patacchini E., Klaauw B. van der, Weber A. (2005), *Study on the effectiveness of ALMPs: Research project for the European Commission, DG Employment, Social Affairs and Equal Opportunities. Final report*, RWI – Leibniz-Institut für Wirtschaftsforschung, http://www.rwi-essen.de/media/content/pages/publikationen/rwi-projektberichte/PB_ALMP.pdf [accessed: 26.04.2019].

Kukuła K. (2000), *Metoda unitaryzacji zerowanej*, Wydawnictwo Naukowe PWN, Warszawa.

Lehmann H., Kluve J. (2008), *Assessing Active Labor Market Policies in Transition Economies*, "Working Papers" Dipartimento Scienze Economiche, Universita' di Bologna, https://econpapers.repec.org/scripts/redir.pf?u=http%3A%2F%2Famsacta.unibo.it%2F4601%2F1%2F646.pdf;h=repec: bol: bodewp: 646 [accessed: 26.04.2019].

Maksim M., Wiśniewski Z. (eds.) (2012), *Metody i narzędzia badania efektywności aktywnej polityki rynku pracy*, Centrum Rozwoju Zasobów Ludzkich, Warszawa.

Malina A. (2004), *Wielowymiarowa analiza przestrzennego zróżnicowania struktury gospodarki Polski według województw*, "Seria Specjalna: Monografie", no. 162, Wydawnictwo Akademii Ekonomicznej w Krakowie, Kraków.

Młodak A. (2006), *Analiza taksonomiczna w statystyce regionalnej*, Difin, Warszawa.

Nowak E. (1990), *Metody taksonomiczne w klasyfikacji obiektów społeczno-gospodarczych*, Polskie Wydawnictwo Ekonomiczne, Warszawa.

Panek T. (2009), *Statystyczne metody wielowymiarowej analizy porównawczej*, Szkoła Główna Handlowa w Warszawie – Oficyna Wydawnicza, Warszawa.

Pociecha J., Podolec B., Sokołowski A., Zając K. (1988), *Metody taksonomiczne w badaniach społeczno-ekonomicznych*, Państwowe Wydawnictwo Naukowe, Warszawa.

Podolec B., Ulman P., Wałęga A. (2008), *Aktywność ekonomiczna a sytuacja materialna gospodarstw domowych*, Wydawnictwo Uniwersytetu Ekonomicznego w Krakowie, Kraków.

Rollnik-Sadowska E. (2014), *Transformation of European labour market policy models – exemplified by Denmark, Germany and The United Kingdom*, "Optimum. Studia Ekonomiczne", no. 5(71), pp. 38–51, https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=12&cad=rja&uact=8&ved=2ahUKEwjcqe-C3–7hAhWNblAKHWwfBFk4ChAWMAF6BAgCEAI&url=http%3A%2F%2Fcejsh.icm.edu.pl%2Fcejsh%2Felement%2Fbwmeta1.element.hdl_11320_2904%2Fc%2F04_Ewa_ROLLNIK-SADOWSKA.pdf&usg=AOvVaw1Qd2MbgQWT8-qA9-bM550O [accessed: 26.04.2019].

Stanimir A. (2005), *Analiza korespondencji jako narzędzie do badania zjawisk ekonomicznych*, Wydawnictwo Akademii Ekonomicznej we Wrocławiu, Wrocław.

Stanisz A. (2007), *Przystępny kurs statystyki z zastosowaniem STATISTICA PL na przykładach z medycyny, Tom 2. Modele liniowe i nieliniowe*, StatSoft Polska Sp. z o. o., Kraków.

Ustawa z dnia 20 kwietnia 2004 r. o promocji zatrudnienia i instytucjach rynku pracy (Dz.U. 2017, poz. 1065, Załącznik do obwieszczenia Marszałka Sejmu Rzeczypospolitej Polskiej z dnia

26 maja 2017 r.) [Act of 20 April 2004 on Employment Promotion and on Labour Market Institutions, Journal of Laws 2017, item 1065, Annex to the announcement of the Speaker of the Sejm of the Republic of Poland of 26 May 2017].
Wiśniewski Z., Zawadzki K. (eds.) (2011), *Efektywność polityki rynku pracy w Polsce*, Wojewódzki Urząd Pracy w Toruniu, Uniwersytet Mikołaja Kopernika, Toruń.
Zając K. (1982), *Zarys metod statystycznych*, PWE, Warszawa.

Wykorzystanie metod statystycznych do identyfikacji czynników determinujących efektywność zatrudnieniową w powiatowych urzędach pracy w Polsce

Streszczenie: W Polsce do instytucji realizujących zadania publiczne w zakresie rynku pracy należą publiczne służby zatrudnienia, w tym powiatowe urzędy pracy (PUP-y). Urzędy te, stosując aktywne formy aktywizacji osób bezrobotnych, starają się przywrócić je na otwarty rynek pracy. Działania PUP-ów w zakresie aktywizacji klientów podlegają corocznej ocenie. Zgodnie z ustawą o promocji zatrudnienia i instytucjach rynku pracy do oceny funkcjonowania urzędów pracy w latach 2015–2017 wybrane zostały cztery wskaźniki. Wśród nich istotne znaczenie ma wskaźnik efektywności zatrudnieniowej, gdyż jego poziom pozwala ocenić, w jakim stopniu aktywne formy aktywizacji osób bezrobotnych, realizowane przez PUP-y, przyczyniają się do powrotu bezrobotnych do zatrudnienia. Ta ocena jest na tyle istotna, że w Polsce i w innych krajach prowadzone są badania dotyczące zarówno działań aktywizujących bezrobotnych, jak i polityki zatrudnieniowej.

Celem badań, których wyniki zaprezentowano w artykule, była identyfikacja czynników wpływających na wskaźnik efektywności zatrudnieniowej uzyskiwany przez powiatowe urzędy pracy (PUP-y) funkcjonujące w Polsce w 2016 roku. Wykorzystano wybrane metody statystyczne, w tym analizę korelacji i regresji oraz wielowymiarową analizę korespondencji. Okazało się, że zastosowane metody nie pozwoliły na jednoznaczne zidentyfikowanie czynników, które w znaczący sposób wpływają na wskaźnik efektywności zatrudnieniowej, obliczanej na podstawie wszystkich form aktywizacji. Może się okazać, że wyraźniejsze zależności udałoby się zidentyfikować, gdyby rozpatrywać ten wskaźnik odrębnie dla poszczególnych form.

Słowa kluczowe: analiza korelacji, regresja krokowa, unitaryzacja zerowana, wielowymiarowa analiza korespondencji, efektywność zatrudnieniowa, powiatowe urzędy pracy

JEL: C38, C51, J49

Received: 2019-01-11; verified: 2019-05-14. Accepted: 2019-08-16