10.2478/v10103-012-0032-8

ELŻBIETA LITWIŃSKA*

Spatial Analysis of the Labour Market by Using Econometric Tools. The Case of Lower Silesia Region (Dolnośląskie voivodship)

Abstract

This paper presents the application of econometric techniques to examine the labour market in Lower Silesia. First the analysis was performed on a data set for variables connected with labour market recorded in poviats (NUTS 4). In order to determine the existence of spatial autocorrelation Moran's statistics I was calculated. Then the spatial regression model was used to describe the relationship between the rate of unemployment and other variables. Next, LISA cluster maps were generated for units at NUTS 5 level. The results indicate the spatial dimension of the unemployment and its tendency to creating concentrations.

1. Introduction

Studies on European regional labour markets show a significant degree of spatial dependence among them. These investigations use Moran's statistics I and spatial regression models. Both regions marked by high unemployment as well as areas characterized by low unemployment have tendency to cluster in space (Niebuhr 2003). Overman and Puga (2000) underline that existence of clusters of units with similar level of unemployment rate corresponds with polarization processes observed in Europe. However this fact is not often taken into consideration by European policy-makers.

* Ph.D., Wroclaw University of Technology

Poland has experienced changes of labour market both as a result of the transformation in Polish economy and the integration process with the European Union.

The paper aims to present an analysis of relationship among different variables connected with labour market and intra-regional unemployment disparities in Lower Silesia, one of the Polish region. The spatial analysis was made by means of GeoDa software (Anselin 2005). The source of statistical data is GUS (Główny Urząd Statystyczny, Central Statistical Office) (*Local Data Bank, Województwo Dolnośląskie* 2011).

2. Lower Silesia

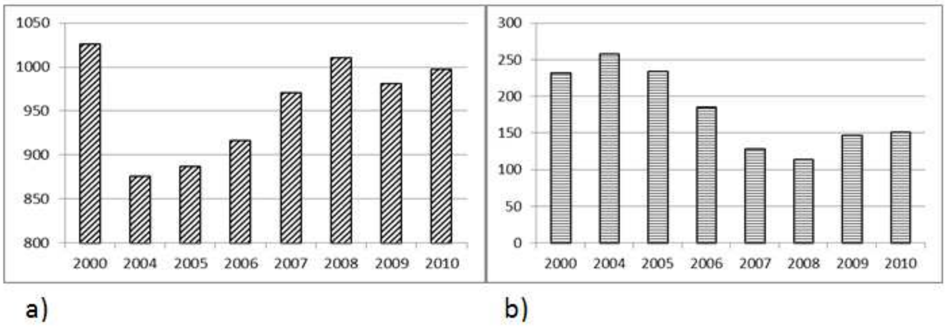
Lower Silesia (Dolnośląskie voivodship) is situated in the south-western corner of Poland. It neighbours Germany and Czech Republic. The region is strategically located in Poland and in Europe, at the intersection of transport routes leading from the east to the west, and from the south to the north. It covers an area of 19,947 sq.km and has a population of nearly 2.9 million.

The Lower Silesian natural resources (including copper ores and lignite) form a base for the development of mining and other industries. Modern industrial processing and services are also developing. In addition, the region's geographical location, environmental conditions, medicinal waters and its rich cultural heritage make tourism and spa healthcare one of the leading industries.

A significant part of Lower Silesia is or was in the past heavily industrialized. During the transformation many production facilities have been liquidated as a consequence of technological, and market changes, as well as conflicts with environmental requirements. The largest fall has affected the textile industry and the coal mining. However key to the Lower Silesian economy is the extraction of copper, silver and lignite. In the northern part of the region there is the Legnica-Głogów Copper Mining District. The lignite mine is located in the gmina Bogatynia, in the south-western corner of the region.

Changes in economy have effected labour market, especially in the beginning of 1990s. Fluctuations in number of jobs and unemployed people in period 2000-2010 are presented on Figure 1. In 2000, the industry provided jobs to 343 thousand people, whereas in 2010 only 248 thousand. It results from a decay in many branches of industry, as well as obsolete technologies and production facilities.

Figure 1. Lower Silesia



Note: umber of work places (a) and unemployed persons (b) (in thousand) in the period 2000-2010

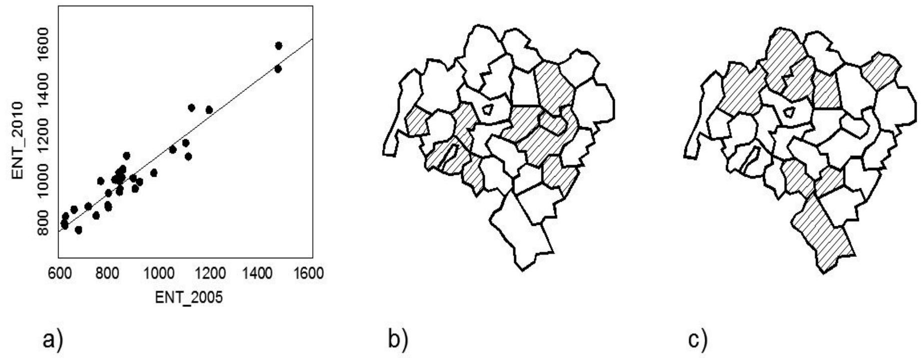
Source: Own elaboration based on GUS data.

3. NUTS 4 level (poviats)

In the first part of analysis it was taken into consideration the division of Lower Silesia into 29 poviats (NUTS 4). The following variables connected with labour market were taking into account: population (POP), population in working age (W_AGE), work places (JOBS), jobs in agriculture (AGR), industry (IND) and services (SER), registered rate of unemployment (UNEMP), the number of entities of national economy recorded in REGON register (ENT) and wages (WAGE). Wages were expressed by average monthly gross wages to the national average (Poland = 100). Data set was for 2005 and 2010.

Changes between 2005 and 2010 may be presented by scatter plots (Figure 2.). In the period 2005-2010 the number of entities increased about 12%. Selection of poviats in which the increase of the variable was noticed (points lying above line) allows to show chosen units on the map. The growth was noticed in Wroclaw and neighbouring poviats and in the group of units near Jelenia Góra (with exception of this city). The decline was recorded in Kotlina Kłodzka and in the northern part of voivodship. The images reflect the real economic situation.

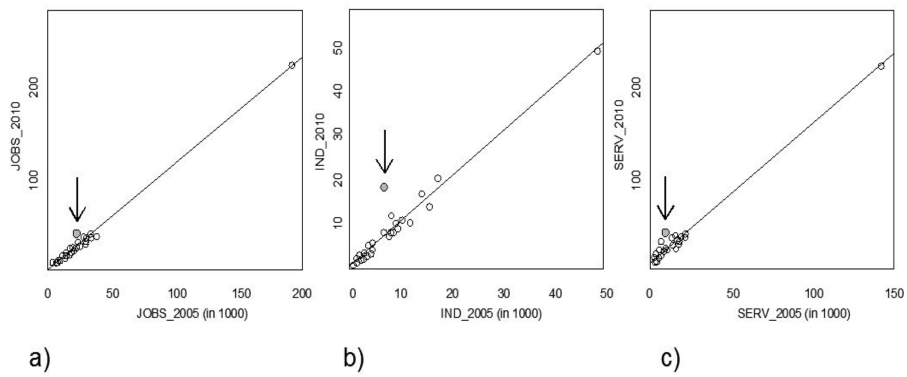
Figure 2. Lower Silesia (NUTS 4). The number of entities in 2005 and 2010. Scatter plot (a) areas with growth (b) decreased areas (c)



Source: Own elaboration by using GeoDa.

Scatter plots for the number of all work places, jobs in industry and in services show that in the period 2005-2010 considerable growth was recorded only in wroclawski poviat what demonstrates the impact of Wrocław on economy of neighbouring units (Figure 3).

Figure 3. Lower Silesia (NUTS 4). Scatter plots for all work places (a) jobs in industry (b) jobs in services (c) in 2005 and 2010



Note: The arrows point the unit with the biggest growth

Source: Own elaboration by using GeoDa.

The first law of geography according to Tobler (1970) is: "All places are related but nearby places are more related than distant places”. In other words, spatial autocorrelation means that dependency exists between values of a variable in neighbouring locations.

The spatial autocorrelation analysis is handled by means of Moran's I statistics (Moran 1950) which requires creating a weight matrix describing a local neighbourhood of each geographic unit. Most common is using binary connectivity based on contiguity: w_{ij} = 1 if units i and j are adjacent, w_{ij} = 0 otherwise.

The statistics I is based on cross-products of the deviations from the mean and is calculated for observations on a variable x at n locations:

$$I = \frac{n}{S_0} \frac{\sum_i \sum_j w_{ij}(x_i - \overline{x})(x_j - \overline{x})}{\sum_i (x_i - \overline{x})^2} \tag{1}$$

where $\overline{x}$ is the mean of the x variable, w_{ij} are the elements of the weight matrix, i, j are number of units, and S_0 is the sum of the elements of the weight matrix: $S_0 = \sum_i \sum_j w_{ij}$ (i, j = 1, 2,… n).

The Moran's statistics I is a single value which applies to the entire data set. It varies from -1 to +1. High negative spatial autocorrelation (-1) means dispersed pattern, and high positive spatial autocorrelation (+1) – clustered pattern. A random arrangement gives a value that is close to 0. It can be interpreted as the correlation between variable X and the "spatially lagged" of X formed by averaging all the values of X for the neighbouring units (W_X).

The Moran scatter plot is a plot with the variable X on the x-axis and the spatially lagged on the y-axis. The variables are standardized so that the units in the graph correspond to standard deviation. The slope of the regression line equals the Moran's I statistics.

The Moran's I statistics were generated for following variables: population, population in working age, work places, jobs in agriculture, industry and services, rate of unemployment, wages and the number of entities for 2005 and 2010 as well (Table 1.). Inference for Moran's I was based on a permutation approach, in which a reference distribution was calculated for spatially random layouts with the same values as observed variables (Anselin 1986). 49999 permutations were used in generating the reference distribution.

Weights were constructed based on contiguity to poviat's boundaries (the first order rook contiguity.

Table 1. Poviats. Moran's I statistics

Variable	2005		2010	
	Moran's I	p-value	Moran's I	p-value
Population (POP)	-0.0008	0.73	0.0095	0.22
Working age population (W_AGE)	-0.0047	0.71	0.0099	0.22
Work places (JOBS)	-0.0175	0.66	0.0138	0.15
Jobs iIn agriculture (AGR)	0.1288	0.09	0.1185	0.11
Jobs in industry (IND)	-0.0362	0.53	0.0748	0.12
Jobs in services (SER)	-0.0291	0.59	-0.0171	0.71
Rate of unemployment (UNEMP)	0.3050	**0.005**	0.2267	**0.024**
Wages (WAGE)	-0.0529	0.46	-0.0166	0.630
Entities (ENT)	0.1235	0.09	0.2177	**0.028**

Source: Own calculation based on GUS data and by using GeoDa.

The Moran's I statistics is statistically significant only for the unemployment rate in 2005 and for the unemployment rate and the number of entities in 2010. Further analysis was conducted for 2010.

In the next step it was decided to apply multivariate regression analysis (Anselin 1988; Anselin, Bera 1998; Suchecki 2010) with the rate of unemployment as the dependent variable and following independent variables: the number of jobs in agriculture, industry and services, the level of wages and the number of entities. The remained variables were excluded from the set because of the high correlation among them.

The regression linear function is:

$$y = x\beta + \varepsilon, \tag{2}$$

where dependent variable y (n×1) is vector of observed dependent variable in n spatial units, x (n×k) represents k independent, explanatory variables. Vector β (k×1) contains unknown coefficients, and vector ε (n×1) represents error term.

In the 1st variant of calculation unknown coefficients were solved by using standard linear regression (Ordinary Least Squares - OLS). Not all coefficients were significant (Table 2.) and in addition multicollinearity condition number was greater than 20 what is not acceptable. AGR_10 (employment in agriculture) was rejected from the set of independent variables.

In the 2nd variant the value of multicollinearity condition number diminished but it was still bigger than 20. So the next independent value was excluded. It was SER_10 (employment in services).

In the 3rd variant there were three following independent variables: IND_10 (employment in industry), WAGE_10 (wages) and ENT_10 (the number of entities). The multicollinearity condition number was proper (18). A test for non-normality (Jarque-Bera) said that errors had normal distribution (p=0.40) and three diagnostics for heteroskedasticity (Breusch-Pagan, Koenker-Bassett, and White) showed a constant variance of errors (p=0.95, 0.87, and 0.46). But tests did not confirm spatial dependence. Therefore it was decided to limit a set of independent variables. Only the level of wages and the number of entities were taken into consideration. The first variable measures the attractiveness of poviat for employees, the second one – expresses the level of economic activity of local population.

In 4th variant of modeling both coefficients were significant. The multicollinearity condition number was 14.3. Test Jarque-Bera said that errors had normal distribution (p=0.51). Homoskedasticity was confirmed (p=0.91, 0.84 and 0.29) for three tests. Both Lagrange Multipliers, M-Lag and LM-Error statistics, rejected the null hypothesis (p=0.020 and 0.0197), what told about the spatial dependence. Also Moran's test statistics I confirmed spatial autocorrelation (p=0.0037).

Table 2. Poviats. Regression analysis. Significance of independent variable coefficients (*p*-value)

	Variant 1 OLS	Variant 2 OLS	Variant 3 OLS	Variant 4 OLS	Variant 5 Spatial Lag Model
W_UNEM_10	-	-	-	-	0.004
AGR_10	0.245	-	-	-	-
IND_10	0.011	0.08	0.03	-	-
SER_10	0.007	0.17	-	-	-
WAGE_10	0.052	0.06	0.065	0.014	0.007
ENT_10	0.082	0.14	0.5	0.035	0.002

Source: Own calculation based on GUS data and by using GeoDa.

Then it was applied the spatially lagged model (5^{th} variant):

$$y = \rho Wy + x\beta + \varepsilon. \quad (3)$$

The spatially lagged rate of unemployment W_UNEM_10 appeared as the additional indicator. Its coefficient parameter (ρ=0.504) reflected the spatial dependence, measuring the average influence on observation in neighbouring units. Breusch-Pagan test detected hetroskedasticity. The final diagnostics was for spatial autocorrelation. Likelihood Ratio Test was significant (p=0.019).

Finally the equation has the form:

$$UNEM\_10 = 0.504 \cdot W\_UNEM\_10 - 0.064 \cdot WAGE\_10 - 0.000499 \cdot ENT\_10 + 15.17$$

what confirms spatial dependence of the unemployment rate and, what it is obvious, the negative relationship between the unemployment and the level of wages and the number of entities.

4. NUTS 5 level (gminas)

One of the problems with areal data is scale effect – spatial data at different scales produce different results. In contrast, the next analysis examined the spatial aspects of Lower Silesian labour market using smaller units of observation – municipalities (gminas - NUTS 5). There are 169 gminas (36 urban gminas, 54 urban-rural and 79 rural ones).

For the year 2010 there were taken into account: population (POP), population in working age (W_AGE), work places (JOBS), rate of unemployment (UNEMP), budget revenue *per capita* (REV), the number of entities (ENT) and the variable describing attraction of labour market (ATR). The last variable was equal the number of people coming to work from other gminas divided by the number of inhabitants working outside their own gmina in 2006 (*Dojazdy do pracy*… 2010).

In the first place, three types of weight matrices were generated using both rook and queen criteria, and 5 nearest neighbours (5nn) as well. Next Moran's I statistics was calculated (Table 3.).

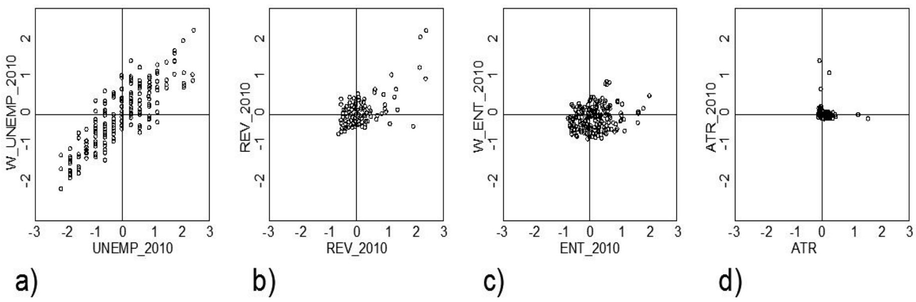
Table 3. Gminas. Moran's I statistics for different weight matrices

W_matrix	POP	W_AGE	JOBS	UNEMP	REV	ENT	ATR
Rook p- value	0.0094 (0.148)	0.0072 (0.162)	0.0151 (0.105)	**0.7285** **(0.00002)**	**0.5097** **(0.00002)**	**0.1342** **(0.0042)**	**0.0198** **(0.0039)**
Queen	0.0087	0.0067	0.0154	0.7154	0.5072	0.1414	0.0210
5nn	0.0025	0.0031	-0.030	0.6717	0.2895	0.1295	-0.0844

Source: Own calculation based on GUS data and by using GeoDa.

The highest value of Moran's statistics was for the unemployment rate (0.7285). Kossowski and Hauke (2012) calculated Moran's coefficient for all Polish gminas and got value 0.7426. Besides the unemployment rate Moran's statistics is statistically significant for the budget revenue *per capita*, the number of entities and the attraction. But the last ones had value near zero (Figure 4.).

Figure 4. Lower Silesia (NUTS 5). Moran scatter plots for unemployment rate (a), budget revenue *per capita* (b), entities (c) and attraction of gmina (d)



Source: Own elaboration by using GeoDa.

In order to determine whether or not geographic patterns exist it was performed LISA (the local indicator of spatial association), the local version of Moran's I (Anselin 1995). For each unit *i* the value of indicator is calculated according to formula:

$$I_i = z_i \sum_j w_{ij} z_j \tag{4}$$

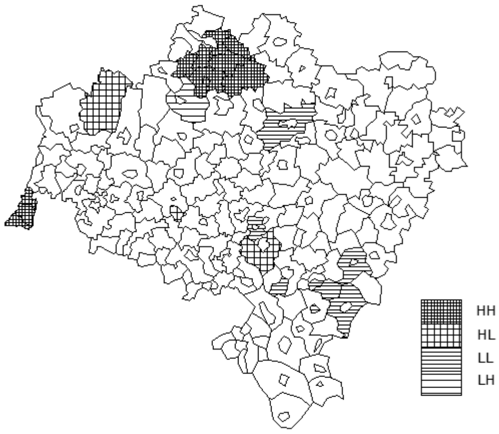
where z_i is the standardized form of original variable x_i, w_{ij} is the element of the spatial weight matrix and *j* describes neighbours of unit *i*.

This measure can be used to identify local clusters or spatial outliers – areas different from their neighbours.

To illustrate the local spatial autocorrelation there were created LISA cluster maps. This map shows a statistically significant relationship with neighbours and four types of spatial autocorrelation. The positive spatial association is when units with high values is surrounded by ones with high values (high-high, H-H) or units with low values are bordered by similar units (low-low, L-L). The negative autocorrelation says that units are different from their neighbours (high-low, HL or low-high, LH).

On Figure 5, the LISA cluster map is reported for the budget revenue *per capita*. The adopted level of significance was p=0.05. Two clusters of the highest budget revenues overlap with Legnica-Głogów Copper Industrial District in the northern part of region and gmina Bogatynia situated in the west. These areas have industrial character. The municipalities (Strzelin, Ciepłowody, Niemcza, Piława Górna and Ziębice) with the lowest budget create cluster in the southern east part of the region.

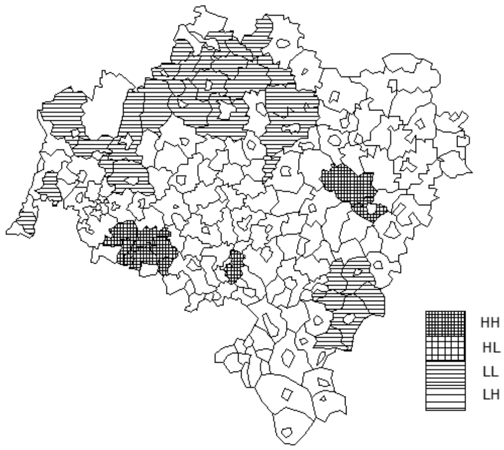
Figure 5. LISA cluster map for budget revenue *per capita*



Source: Own elaboration by using GeoDa.

The LISA map for the number of entities (Figure 6.) shows two groups of low values. The first cluster is located in the northern part of region, where either there are less developed municipalities or municipalities have little entities, but employing a lot of employees (Legnica-Głogów Industrial District). The second one consists of following gminas: Strzelin, Kondratowice, Przeworno, Ciepłowody, Ziębice and Kamieniec Ząbkowicki. Clusters of units with high values is created by gminas near Jelenia Gora where tourism is well developed and Wrocław with gmina Święta Katarzyna.

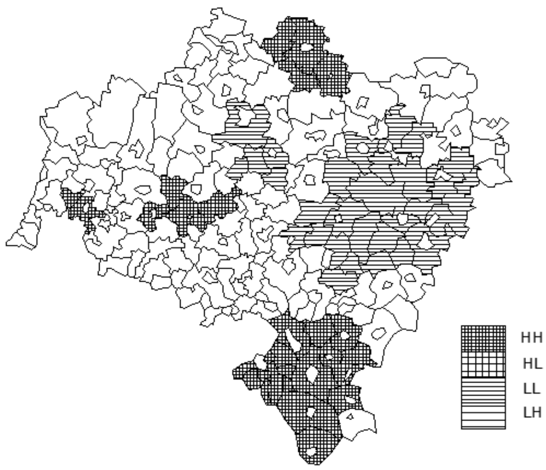
Figure 6. LISA cluster map for the number of entities



Source: Own elaboration by using GeoDa.

In the next step LISA cluster map was created for the rate of unemployment (Figure 7.). There are two clusters of low unemployment, the big one covers Wroclaw Metropolitan Area and smaller one is situated near Lubin. High level of unemployment is recorded in Kotlina Kłodzka, in the northern part of the region (Góra, Niechlów, Wąsosz and Jemielno), and in two groups in the western part.

Figure 7. LISA cluster map for rate of unemployment

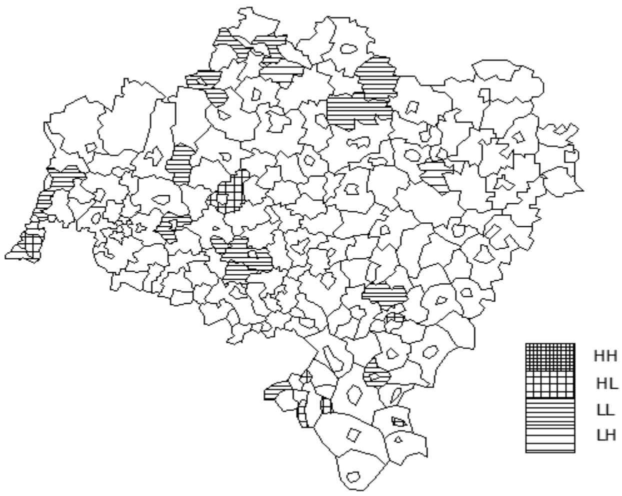


Source: Own elaboration by using GeoDa.

This image shows the spatial structure of regional unemployment disparities. On the one hand it is seen the important role of Wrocław as pole of growth and on the other hand it is easy to notice that despite the great touristic potential municipalities in Kotlina Kłodzka are not able to use it sufficiently.

The attraction of the gmina was defined as the number of people coming to work from other gminas divided by the number of inhabitants working outside their own gmina. The variable defined in such a way does not create concentrations (Figure 8).

Figure 8. LISA cluster map for attraction of gmina



Source: own elaboration by using GeoDa.

5. Conclusions

In this paper it was presented the spatial analysis of the labour market in Lower Silesia. The first part of research was run on a data set for poviats (NUTS 4). The Moran's statistics was calculated and then the spatial regression model for the unemployment was obtained.

In turn, division of Lower Silesia into gminas (NUTS 5) was taken into account. Among different variables connected with labour market it were chosen four of them with statistically significant Moran's statistics I and then LISA cluster maps were created. The biggest clusters were found for the unemployment and the budget revenue *per capita*.

All these calculations confirm that the unemployment is spatially dependent and tends to concentrate. The rate of unemployment seems to be one of the most important indicators which characterize the labour market.

References

Anselin L. (1986), *MicroQAP: a Microcomputer Implementation of Generalized Measures of Spatial Association*, Working Paper, Department of Geography, University of California, Santa Barbara

Anselin L. (1988), *Spatial Econometrics: Methods and Models* Kluwer Academic Publishers, Dordrecht

Anselin L. (1995), *Local Indicators of Spatial Association — LISA*, Geographical Analysis, 27: 93-115

Anselin L. (2005), *Exploring Spatial Data with GeoDa™: A Workbook*, Center for Spatially Integrated Social Science, Urbana, IL

Anselin L., Bera A. K. (1998), *Spatial Dependence in Linear Regression Models with an Introduction to Spatial Econometrics*. [in:] A. Ullah and D. Giles (eds), *Handbook of Applied Economic Statistics*. New York: Marcel Dekker, pp.237-289

Dojazdy do pracy w Polsce. Terytorialna identyfikacja przepływów ludności związanych z zatrudnieniem [2010] Urząd Statystyczny w Poznaniu, Poznań

Moran P.A.P. (1950), *Notes on Continuous Stochastic Phenomena* Biometrika 37:17-23

Kossowski T. Hauke J. (2012), *Analysis of the Labour Market in Metropolitan Areas: a Spatial Filtering Approach*, Quaestiones Geographicae, 31[2], Bogucki Wydawnictwo Naukowe, Poznań, pp. 39-48

Local Data Bank. Central Statistics Office of Poland < http://www.stat.gov.pl/>

Niebuhr A. (2003), *Spatial Interaction and Regional Unemployment in Europe* European Journal of Spatial Development, 5 <http:www.nordregio.se/EJSD/>

Overman H.G. Puga D. (2000), *Unemployment clusters across European Regions and Countries* Centre for Economic Performance, Discussion Paper 434

Suchecki B. (2010), *Ekonometria przestrzenna. Metody i modele analizy danych przestrzennych*, Wydawnictwo C.H. Beck, Warszawa

Tobler W. (1970), *A Computer Movie Simulating Urban Growth in the Detroit Region* Economic Geography, 46[2]:234-240

Województwo Dolnośląskie. Podregiony. Powiaty. Gminy (2011) Urząd Statystyczny we Wrocławiu, Wrocław

Streszczenie

PRZESTRZENNA ANALIZA RYNKU PRACY PRZY UŻYCIU NARZĘDZI EKONOMETRYCZNYCH. PRZYKŁAD WOJEWÓDZTWA DOLNOŚLĄSKIEGO

Artykuł prezentuje wykorzystanie technik ekonometrycznych do badania rynku pracy na Dolnym Śląsku. Najpierw analiza dotyczyła zbioru danych związanych z rynkiem pracy na poziomie powiatów (NUTS 4). W celu zbadania istnienia autokorelacji przestrzennej policzono statystykę Morana I, potem został użyty model regresji przestrzennej do opisu zależności pomiędzy stopą bezrobocia a innymi zmiennymi. W następnym kroku wygenerowano mapy obrazujące lokalny wskaźnik zależności przestrzennej LISA dla gmin (NUTS5). Wyniki wskazują na przestrzenny wymiar bezrobocia i jego tendencję do tworzenia koncentracji.