



Dorota Dejaniak

State Higher School of Technology and Economics in Jarosław, Institute of Technical Engineering,
dorota.dejaniak@pwste.edu.pl

The Application of Spatial Analysis Methods in the Real Estate Market in South-Eastern Poland

Abstract: The aim of the article is to apply the method of spatial analysis to research the real estate property market in south-eastern Poland. The methods of spatial statistics will be used to model the space differences of prices per one square metre of dwelling surface located in districts of south-eastern Poland and to investigate spatial autocorrelation. The databases will be presented in a graphical form. The results may be used to set the spatial regularities and relations. The methods presented may be applied while making strategic decisions.

Keywords: spatial econometrics, spatial autocorrelation, weight matrix, global correlation coefficient

JEL: C21, C31, C51, R12, R32

1. Introduction

A gradual stabilization of the housing markets has recently been observed after the turbulent years of 2006 to 2008. The stabilization is supported by the partial weakening of economic activity and a significant reduction in lending rates. Currently, the supply dynamically responds to the high demand, supported by low interest rates on housing loans, an increase in salaries and the implementation of the government program for the Young (M_{dM}). Housing prices remain stable despite the increase in supply and demand (The report on the situation on real estate and commercial property markets in Poland in 2014). Reducing interest rates on housing loans is a consequence of cuts in interest rates by the Monetary Policy Council. The observed positive growth in lending and strong competition among real estate developers have both had an impact on the stabilization of real estate prices. However, housing markets are local in nature. The impact of macroeconomic factors can be affected by local factors influencing supply and demand. The structural feature of the real estate markets is their periodicity. Housing cycles are often not synchronized with the business cycle, however, they remain under the influence of economic factors, including the financial and longer cycles and they tend to have a local character. Reports on real estate markets conducted by the Polish National Bank created data based on the analysis of transactions show a strong dependence of the price of residential units on the level of wages, unemployment, and the cost of construction. Also, an increase in housing prices in previous periods has a positive effect on future periods. In this context, interestingly, the issue of spatial development of the housing market seems particularly interesting. The purpose of this analysis is to obtain information about the spatial dependence regions and the interactions between the values of the variables tested in different locations. Spatial analysis enables the determination of similarities and differences between regions, the use of its methods and tools make it possible to distinguish groups of regions that are similar to each other and also find significantly different regions of neighbouring countries. With the estimation models taking into account the spatial factor, it is possible to determine the spatial relationship between observations in different locations, as well as demonstrating the validity of using this spatial factor for differentiating the examined phenomenon between locations (Kopczewska, 2006).

2. Methodology of research

Understanding the diversity of space allows us to predict changes and shape the policies of regional economic development. Analysis of the space takes place at various levels: analysis of location, transfer-spatial interactions, economies

of scale, spatial autocorrelation. Econometrics takes into account the spatial aspect of the position of the object in space, as opposed to classical econometrics, which deals with the setting using mathematical-statistical, and quantitative regularity methods.

The occurrence of spatial dependence results from two reasons (Janc, 2006). The first concerns the analysis of data in spatial studies with spatial units (country, state, county, municipality). The second reason is that the socio-economic activities of people are influenced by distance and location. The phenomenon of spatial autocorrelation is associated with the First Law of Geography by Tobler, which says that in a space where everything is related to everything else, the closer things are more related than the distant things (Miller, 2004). Spatial autocorrelation stands for the degree of correlation of the observed value of the variable in a given location with the value of the same variable in a different location. This means that the test variable determines and at the same time is determined by its embodiments in other locations (Sucheckı, 2010: 103).

For the time series, we refer to the delay in time and the phenomenon of time autocorrelation, while for spatial data we refer to the spatial delay caused by the criterion of neighborhood. The spatial structure of the neighbourhood is defined by the spatial scales, recorded with a matrix or graph (Janc, 2006). In the case of the recording matrix, an adjacency matrix is created in the first place. It is a binary matrix. A value of zero means no neighbourhood between regions, and a value of 1 is awarded for an element that satisfies the neighbourhood condition. Then the matrix is standardized by lines, so that the sum of each row equals 1. The matrix of the neighbourhood is the most common type of matrix used. The more sophisticated weights matrices are: the Cliff and Ord matrix, the Dacey matrix, the social distance matrix, the economic distance matrix (Janc, 2006; Miller, 2004). One of the most commonly used metrics for determining the strength and character of spatial autocorrelation are global and local spatial statistics. Of these, the most common are the global and local Moran's I statistics. It is also possible to calculate Geary's, and Getis-Ord coefficients. Moran's I global statistics is used to test the existence of global spatial autocorrelation and it is expressed by formula (Sucheckı, 2010: 112):

$$I = \frac{1}{\sum_{i=1}^n \sum_{j=1}^n w_{ij}} \cdot \frac{\sum_{i=1}^n \sum_{j=1}^n w_{ij} (x_i - \bar{x})(x_j - \bar{x})}{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2}, \quad (1)$$

where: x_i is an observation in a region and \bar{x} is an average from all the studied regions, n is the number of regions, and w_{ij} is the element of the spatial weights matrix.

If the weights matrix is standardized by lines, the sum of all the elements of such a matrix is equal to the number of rows $S_0 = \sum_i \sum_j w_{ij}^* = n$:

$$I_W = \frac{\sum_{i=1}^n \sum_{j=1}^n w_{ij}^* (x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2}. \quad (2)$$

Moran's statistics can take two forms, depending on whether it should assess normality or randomization. Therefore moments to test the null hypothesis are calculated assuming normality or randomization (Bivand, 1980). Moran's statistics has a value in the range from -1 to 1 . A value of 0 means no autocorrelation, negative values signify negative autocorrelation, which proves the existence of different values next to each other. Positive autocorrelation means that the values are concentrated in space, and the neighbouring regions are similar. This means that we are dealing with clusters, specifically, spatial clusters. This is a process comparable to diffusion. In the case of negative spatial autocorrelation neighbouring areas are different, more than it would appear from a random distribution. This is called a checkerboard effect.

The graphical representation of Moran's global statistics is a scatter chart, which is used to visualize the local spatial relationships. The graph on the horizontal axis has a standardized, analyzed variable, and on the vertical axis, a tested, standardized variable that is spatially delayed (Anselin, 1995; Kopczewska, 2006). The graph allows for a regression line and is divided into four quadrants (HL, HH, LL, LH) in relation to the zero point.

HH and LL squares indicate the clustering of regions with similar values. The slope coefficient of the regression line is associated with the Moran's I global statistic for standardized weights matrix lines.

The statistics used to determine the spatial autocorrelation can be used to identify spatial systems. For this purpose, the local ratings of spatial relationships LISA were used, proposed by Anselin in 1995, they allow to determine the similarity of the spatial entity with respect to the neighbours and examine the statistical significance of the compound (Anselin, 1995). The LISA for each observation indicates the degree of importance of the spatial concentration of similar values around the analyzed spatial unit, the sum of LISA for all observations is proportional to a global indicator of spatial relationships. In the article, Moran's local statistics I_i was used as LISA.

Moran's local statistics I_i measures whether the region is surrounded by neighbouring regions of similar or different values of the test variable relative to a random distribution of those values in the space. I_i is the smoothed index for individual observation, which can be used to find local clusters. Local statistics is expressed by the formula (Sucheckı, 2010: 123):

$$I_i(w) = \frac{(x_i - \bar{x}) \sum_{j=1}^n w_{ij}^* (x_j - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} , \quad (3)$$

where: elements w_{ij}^* are derived from the spatial weights matrix standardized with lines.

Tests of statistics significance are based on the distribution resulting from conditional randomization or permutation. Standardized local Moran's statistics has a value that is significantly negative when the region is surrounded by other regions that have significantly different values of the variable under consideration, which is interpreted as a negative autocorrelation. Accepting values that are significantly positive means that the region is surrounded by similar regions and neighbouring region clustering occurs. Values of p that are less than 0.05 indicate a significant positive autocorrelation values of p that are above 0.95 indicate a significant negative spatial autocorrelation. The absolute value of the local Moran's statistics can be interpreted as the degree of similarity and diversity.

3. An example of the use of statistics based on spatial dependence

For the analysis presented in this article, we have used the data concerning the average price per one square metre of dwelling surface in seventy-one districts in south-eastern Poland, located in the provinces of Lublin, Małopolska, and Podkarpackie. The data source was the contracts of sale on the primary and secondary markets. The summary was generated from the AMRON database, which is the only Polish interbank, a standardised database of real estate and their prices. The collected data comes from deeds, values estimated by appraisers in the appraisal reports or other reliable sources of information with J Recommendation. Dates of transactions examined covered the period from 2014.04.20 to 2016.04.20. In addition, the report has been enriched with macroeconomic data: the population of working age, the registered unemployment rate and the total housing put into use. Analyses were performed using Statistica software and Rcran. To describe the spatial relationship, a spatial matrix was generated. During the analysis, two types of neighbourhood matrix were used; a basic binary matrix and a matrix of the first row standardized with rows. Thus prepared the database has been used to calculate the autocorrelation of the global and local Moran's statistics. Basic descriptive statistics on the mean, median and standard deviation of the variable tested and selected macroeconomic data for the region were pre-generated (Table 1).

Table 1. The values of the basic descriptive statistics

Variable	Descriptive statistics			
	Average	Minimum	Maximum	Standard deviations
average price of 1 sq m [PLN]	3184.57	1095.41	9268.7	1135.95
working-age population	74021.02	15251.00	496352.0	64720.98
the unemployment rate [%]	27.39	5.20	905.0	110.61
total housing provided	412.83	40.00	7346.0	924.43

Source: own calculation

In the analysed period for Warsaw, variables assumed the following values (Table 2):

Table 2. The values of simple descriptive statistics

average price of 1 sq m [PLN]	7752.37
working-age population	1091242
the unemployment rate [%]	4.3
total housing provided	14964

Source: own calculation

Table 3. The values of correlation coefficients

Correlations (Marked correlation coefficients are significant with $p < 0.05$)			
the average price per 1 square meter [PLN]	working-age population	unemployment rate	total residential units provided for use
1.000000	0.445389	-0.341701	0.438008
0.445389	1.000000	-0.444744	0.932176
-0.341701	-0.444744	1.000000	-0.411705
0.438008	0.932176	-0.411705	1.000000

Source: own calculation

Data analysis was preceded by calculation of the correlation coefficients for the tested variable and two macroeconomic data. The values obtained show a significant effect of the number of working age population in a given region on the price of a dwelling.

Statistically significant are coefficients of correlations between the variable of average price per 1 sq m of housing space, and variable of working-age population and residential units generally provided for use. This confirms the opinion that the property market is a market operating locally. In the least developed provinces there are the least economically developed sub-regions in terms of GDP per capita (*Statistical Yearbook of the Republic of Poland 2015*, 2015). These include

Following from that, a scatter chart of the Moran’s global statistics was drawn, along with a matching straight line, the slope of which is equivalent to the calculated statistics.

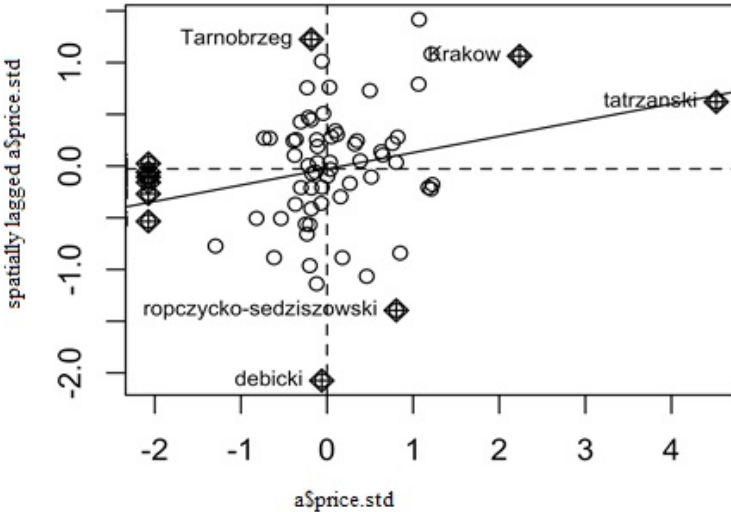


Figure 2. Scatter plots of Moran’s local statistics

Source: own elaboration

Table 4. The vital values of the local Moran’s statistics

Region	Ii	E.Ii	Var.Ii	Z.Ii	Pr(z > 0)
nowosadecki	1.3318979	-0.0142857	0.1397377	3.6012026	0.0001584
nowotarski	1.5357997	-0.0142857	0.1699803	3.7597271	0.0000850
tatrzański	2.8433606	-0.0142857	0.1699803	6.9312117	0.0000000
wielicki	0.8541393	-0.0142857	0.1019344	2.7200195	0.0032639
Krakow	2.4142710	-0.0142857	0.8958036	2.5659125	0.0051452

Source: own elaboration

As the next step, the local Moran’s statistics have been appointed (Table 5). Local statistics are used to determine whether a region is surrounded by high or low values, so it is possible to identify the so-called *hotspots* and local clusters. Table 4 shows the vital local Moran’s statistics.

Based on Table 5 graphs of statistics values and of vital local Moran’s statistics were generated.

The darkest colour is used for selected districts, most correlated with their neighbours. The examples are the districts of Krakow and Wieliczka and a group of Tatra, Nowy Sacz, Nowy Targ districts. They are more similar to each other than they would be as a result of the stochastic nature of the studied phenomenon.

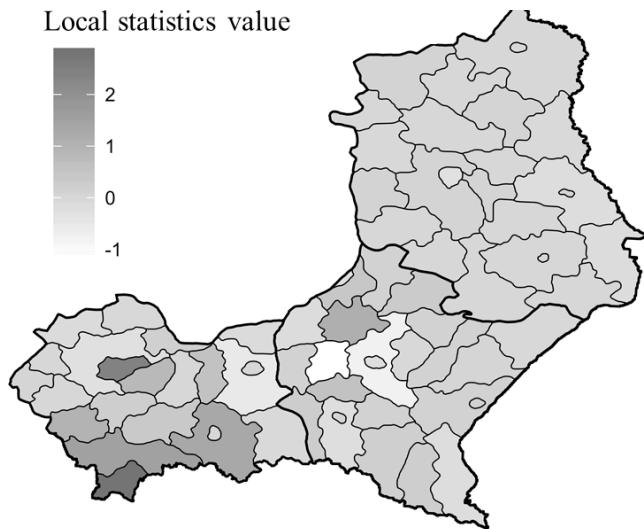


Figure 3. Chart of Moran's local statistics for the studied district
Source: own elaboration

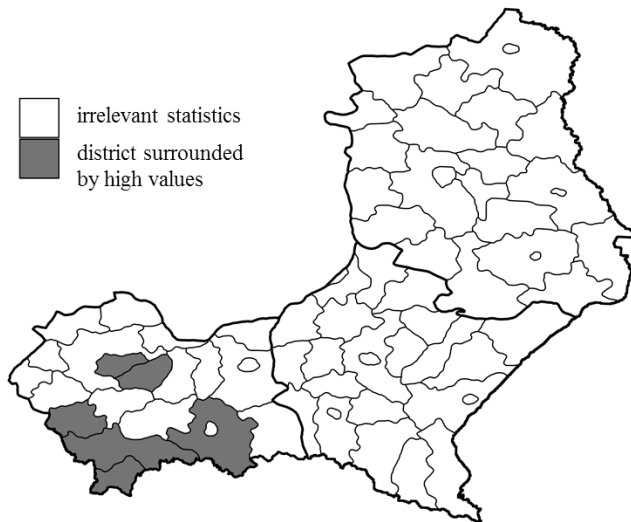


Figure 4. Chart of essential values of the Moran's local statistics
Source: own elaboration

4. Conclusion

Analysis of global and local indicators of spatial dependency can successfully be used in the economic analysis, including market research of real estate. Statistics' spatial autocorrelation, which indicates the type and strength of spatial dependence, allows for an expansion of the traditionally used measures. The analyses enable the comparison of economic processes and they become the basis for business decisions. The key issue is the selection of the matrix weights, strongly associated with the tested regions. The performed analysis confirms the opinion that the prices of residential properties depend on their spatial position. Of little impact on the average price of housing in the studied region of south – east Poland are macroeconomic variables. The analysis confirmed that in a very attractive part of the Podhale, there are important local business initiatives affecting the development of the region. This translates into the highest housing prices in the surveyed districts. Such a big influence of local conditions on the real estate market could mean that a narrowing of the research area e.g. to the region, could positively affect the conducted analysis. Thus, decisions to buy property would be optimal.

References

- Aldstadt J., Getis A. (2004), *Constructing the Spatial Weights Matrix Using a Local Statistics*, "Geographical Analysis", vol. 36, no. 2, pp. 90–104.
- Anselin L. (1995), *Local Indicators of Spatial Association – LISA*, "Geographical Analysis", vol. 27, no. 2, pp. 93–115.
- Basu S., Thibodeau T. (1998), *Analysis of spatial autocorrelation in house prices*, "Journal of Real Estate Finance and Economics", vol. 17(1), pp. 61–85.
- Bivand R. (1980), *Autokorelacja przestrzenna a metody analizy statystycznej w geografii*, [in:] Z. Chojnicki (ed.), *Analiza regresji w geografii*, PWN, Poznań.
- Dydenko J. (ed.) (2015), *Szacowanie nieruchomości. Rzeczoznawstwo majątkowe*, Wolters Kluwer, Warszawa.
- Getis A., Ord J.K. (1995), *Local spatial autocorrelation statistic: Distributional issues and an application*, "Geographical Analysis", vol. 27(4), pp. 286–306.
- Haining R. (2003), *Spatial data analysis. Theory and practice*, Oxford University Press, Oxford.
- Ismail S. (2006), *Spatial Autocorrelation and Real Estate Studies: A Literature Review*, "Malaysian Journal of Real Estate", vol. 1, pp. 1–13
- Janc K. (2006), *Zjawisko autokorelacji przestrzennej na przykładzie statystyki I Morana oraz lokalnych wskaźników zależności przestrzennej (LISA) – wybrane zagadnienia metodyczne*, Instytut Geografii i Rozwoju Regionalnego Uniwersytetu Wrocławskiego, Wrocław.
- Kopczewska K. (2006), *Ekonometria i statystyka przestrzenna z wykorzystaniem programu R CRAN*, Wydawnictwa Fachowe CeDeWu, Warszawa.
- Miller H.J. (2004), *Tobler's First Law and Spatial Analysis*, "Annals of the Association of American Geographers", vol. 94, no. 2, pp. 284–289.
- Statistical Yearbook of the Republic of Poland 2015* (2015), Central Statistical Office of Poland, Warsaw.

Sucecka J. (ed.) (2014), *Statystyka przestrzenna. Metody analiz struktur przestrzennych*, Wydawnictwo C.H. Beck, Warszawa.

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

The report on the situation on real estate and commercial property markets in Poland in 2014, The Polish National Bank, September 2015., https://www.nbp.pl/publikacje/rynek_nieruchomosci/raport_2014.pdf [accessed: 8.05.2016].

Appendix

Table 5. The values of the local Moran's statistics

Region	Ii	E.Ii	Var.Ii	Z.Ii	Pr(z > 0)
bialski	-0.0393104	-0.0142857	0.2909509	-0.0463937	0.5185018
bilgorajski	-0.0209933	-0.0142857	0.1699803	-0.0162693	0.5064902
chelmski	-0.1804075	-0.0142857	0.1397377	-0.4443957	0.6716217
hrubieszowski	-0.1321792	-0.0142857	0.2909509	-0.2185648	0.5865055
janowski	-0.0304741	-0.0142857	0.1699803	-0.0392649	0.5156604
krasnostawski	0.0125673	-0.0142857	0.1397377	0.0718350	0.4713666
krasnicki	0.0144445	-0.0142857	0.1699803	0.0696851	0.4722222
lubartowski	0.0198106	-0.0142857	0.8958036	0.0360247	0.4856313
lubelski	0.0693224	-0.0142857	0.2909509	0.1550024	0.4384097
leczyński	-0.0452335	-0.0142857	0.4421641	-0.0465412	0.5185605
lukowski	0.0862784	-0.0142857	0.1699803	0.2439178	0.4036472
opolski	0.0886534	-0.0142857	0.2153443	0.2218267	0.4122244
parczewski	0.0642712	-0.0142857	0.1699803	0.1905395	0.4244432
pulawski	-0.0552307	-0.0142857	0.1397377	-0.1095327	0.5436100
radzynski	-0.0004764	-0.0142857	0.1181358	0.0401774	0.4839758
rycki	-0.0010635	-0.0142857	0.8958036	0.0139700	0.4944269
swidnicki	0.0690821	-0.0142857	0.1181358	0.2425535	0.4041756
tomaszowski	0.0079035	-0.0142857	0.1699803	0.0538198	0.4785394
wlodawski	-0.0937414	-0.0142857	0.0792524	-0.2822402	0.6111203
zamojski	0.0248725	-0.0142857	0.8958036	0.0413729	0.4834993
Biała Podlaska	-0.0013178	-0.0142857	0.1699803	0.0314535	0.4874539
Chelm	0.0764279	-0.0142857	0.1699803	0.2200256	0.4129256
Lublin	-0.2730243	-0.0142857	0.1397377	-0.6921568	0.7555806
Zamosc	0.0012936	-0.0142857	0.2153443	0.0335723	0.4866091
bochenski	0.0287791	-0.0142857	0.8958036	0.0455005	0.4818542
brzeski	0.5662397	-0.0142857	0.2153443	1.2509921	0.1054687
chrzanowski	-0.0035964	-0.0142857	0.2909509	0.0198170	0.4920947
dabrowski	-0.0964880	-0.0142857	0.2909509	-0.1523962	0.5605628
gorlicki	-0.0835651	-0.0142857	0.0893333	-0.2317915	0.5916500
krakowski	-0.2476717	-0.0142857	0.4421641	-0.3509806	0.6371986
limanowski	0.3687137	-0.0142857	0.1699803	0.9289638	0.1764539
miechowski	-0.0626107	-0.0142857	0.2909509	-0.0895904	0.5356936

Region	Ii	E.Ii	Var.Ii	Z.Ii	Pr(z > 0)
myslenicki	0.2327284	-0.0142857	0.1181358	0.7186724	0.2361714
nowosadecki	1.3318979	-0.0142857	0.1397377	3.6012026	0.0001584
nowotarski	1.5357997	-0.0142857	0.1699803	3.7597271	0.0000850
olkuski	-0.1018276	-0.0142857	0.1397377	-0.2341852	0.5925794
oswiecimski	0.0394077	-0.0142857	0.8958036	0.0567302	0.4773800
proszowicki	0.1684519	-0.0142857	0.1699803	0.4432294	0.3287999
suski	1.0156461	-0.0142857	0.2909509	1.9094076	0.0281048
tarnowski	-0.4984938	-0.0142857	0.4421641	-0.7281826	0.7667491
tatrzański	2.8433606	-0.0142857	0.1699803	6.9312117	0.0000000
wadowicki	-0.0220786	-0.0142857	0.2909509	-0.0144473	0.5057634
wielicki	0.8541393	-0.0142857	0.1019344	2.7200195	0.0032639
Krakow	2.4142710	-0.0142857	0.8958036	2.5659125	0.0051452
Nowy Sacz	0.0219266	-0.0142857	0.8958036	0.0382605	0.4847400
Tarnow	0.0119878	-0.0142857	0.1699803	0.0637262	0.4745941
bieszczadzki	-0.1766039	-0.0142857	0.2153443	-0.3497844	0.6367497
brzozowski	0.1272158	-0.0142857	0.1699803	0.3432115	0.3657197
debicki	0.1275421	-0.0142857	0.8958036	0.1498494	0.4404417
jaroslawski	0.0338584	-0.0142857	0.0893333	0.1610783	0.4360159
jasielski	0.2750799	-0.0142857	0.1181358	0.8418916	0.1999243
kolbuszowski	1.1249678	-0.0142857	0.8958036	1.2036881	0.1143551
krosnienski	-0.1990068	-0.0142857	0.2909509	-0.3424575	0.6339967
lezajski	0.1376309	-0.0142857	0.1699803	0.3684732	0.3562602
lubaczowski	-0.0512464	-0.0142857	0.2153443	-0.0796477	0.5317413
lancucki	-0.0474441	-0.0142857	0.1699803	-0.0804256	0.5320506
mielecki	-0.1604231	-0.0142857	0.1397377	-0.3909351	0.6520774
nizanski	0.3305611	-0.0142857	0.1181358	1.0033107	0.1578555
przemyski	0.1166214	-0.0142857	0.0792524	0.4650043	0.3209642
przeworski	0.1388697	-0.0142857	0.4421641	0.2303248	0.4089197
ropczycko-siedziszowski	-1.1389553	-0.0142857	0.2153443	-2.4235851	0.9923159
rzeszowski	-0.7251225	-0.0142857	0.2909509	-1.3178321	0.9062201
sanocki	0.1955495	-0.0142857	0.1699803	0.5089547	0.3053920
stalowowolski	0.1407509	-0.0142857	0.2153443	0.3340932	0.3691546
strzyzowski	0.5520968	-0.0142857	0.2153443	1.2205151	0.1111348
tarnobrzesci	0.4198139	-0.0142857	0.2909509	0.8047845	0.2104720
leski	0.2196958	-0.0142857	0.1699803	0.5675213	0.2851800
Krosno	0.0362255	-0.0142857	0.2153443	0.1088482	0.4566614
Przemysl	0.1547051	-0.0142857	0.2153443	0.3641636	0.3578679
Rzeszow	-0.2223948	-0.0142857	0.1397377	-0.5567167	0.7111395
Tarnobrzeg	-0.2259730	-0.0142857	0.8958036	-0.2236600	0.5884891


Source: own elaboration

Zastosowanie metod analizy przestrzennej do badania rynku nieruchomości południowo-wschodniej Polski

Streszczenie: W artykule pokazano możliwości zastosowania metod analizy przestrzennej do badania rynku nieruchomości. Metody te zostały wykorzystane do zbadania zróżnicowania ceny 1 m² nieruchomości mieszkaniowej w Polsce południowo-wschodniej. Analizie poddano współczynniki korelacji globalnej i lokalnej Morana. Otrzymane wyniki zostały wzbogacone danymi makroekonomicznymi. Przeprowadzona analiza potwierdza opinię, że ceny nieruchomości mieszkalnych zależą od położenie przestrzenne. Natomiast oddziaływanie czynników makroekonomicznych na rynek nieruchomości może zostać zaburzone przez czynniki lokalne kształtujące popyt i podaż. Statystyki autokorelacji przestrzennej, które informują o rodzaju i sile zależności przestrzennej, umożliwiają poszerzenie tradycyjnie stosowanych miar, a tym samym mogą stać się podstawą decyzji biznesowych.

Słowa kluczowe: ekonometria przestrzenna, autokorelacja przestrzenna, macierz wag, współczynnik korelacji globalnej

JEL: C21, C31, C51, R12, R32

	<p>© by the author, licensee Łódź University – Łódź University Press, Łódź, Poland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license CC-BY (http://creativecommons.org/licenses/by/3.0/)</p> <p>Received: 2016-08-14; verified: 2017-10-18. Accepted: 2017-12-18</p>
---	--