INFORMATION
SYSTEMS IN
MANAGEMENT

Information Systems in Management (2016) Vol. 5 (1) 15–23

THE APPLICATION OF SPATIAL ANALYSIS METHODS IN THE REAL ESTATE MARKET IN SUBCARPATHIAN REGION

DOROTA DEJNIAK

Institute of Technical Engineering, State Higher School of Technology and Economics in Jaroslaw, Poland

The aim of the article is to apply the method of spatial analysis to research the real estate property market in Subcarpathian region in Poland. The methods of spatial statistics will be used to model the space differences of prices per $1m^2$ of a residential unit located in 26 districts of the Subcarpathian region and to investigate spatial autocorrelation. The databases will be presented in the graphical form. The results may be used to set the spatial regularities and relations. The methods presented may be applied while taking strategic decisions.

Keywords: spatial autocorrelation, property markets, spatial heterogeneity

1. Introduction

The aim of spatial analysis is to obtain information about the spatial dependence regions and interactions between the values of the variables tested in different locations. Spatial analysis allows determining the similarities and differences between regions, the use of such methods and tools enables to distinguish a group of regions similar to each other and find regions significantly different from its neighbours. Thanks to 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 to demonstrate the existence of spatial factor differentiating the studied phenomenon between locations [6].

Understanding the diversity of space allows us to predict a change and shape the policies of regional economic development. The space analysis takes place at different levels: location analysis, spatial interactions, economies of scale, spatial autocorrelation. Space effects can be divided into:

- Spatial heterogeneity - structural relationships, changing along with the location of the object,
- Spatial autocorrelation relating to the systematic spatial changes.

Spatial econometrics takes into account the aspect of the position of the object in space, unlike the classic econometrics, which deals with the setting using mathematical methods - statistical, quantitative regularity.

The occurrence of spatial dependence results from two reasons [4]. The first concerns the analysis of spatial data in studies with spatial units (country, county, municipality, district). The second reason is the fact of the socio – economic activities of people shaped by distance and locations. The phenomenon of spatial autocorrelation associated with the First Law of Geography of Tobler, saying that in space everything is related to everything else, the closer things are more related than distant things [5].

The article presents the spatial differentiation of prices per $1m^2$ of a residential unit. For spatial units, according to which we analysed, differences in selected districts of the Subcarpathian county. In addition, it presents the possibility of practical application of spatial dependence indicators for economic analysis.

2. Research Methodology

Analysis of the spatial autocorrelation is based on the finding that the intensification of the phenomena in the spatial entity depends on the level of such phenomenon in adjacent units. For time series we talk about a time delay and the phenomenon of temporal autocorrelation, while for spatial data we talk about the delay caused by the criterion of spatial neighbourhood. The spatial structure of the neighbourhood is defined by the spatial scales, recorded with matrix or graph [4].

With the saving matrix, initially formed is adjacency matrix - a binary matrix. A value of zero means no neighbourhood between regions, a value of 1 is awarded for an element that satisfies neighbourhood condition. Then the matrix is standardized by rows, so that the sum of each row equals 1. Adjacency matrix is the most common type used in the array. The group of more sophisticated matrix of weights include Cliff and Ord matrix, Dacey matrix, social distance matrix, economic distance matrix, [2, 6].

One of the common measures used to determine the strength and character of spatial autocorrelation are global and local spatial statistics. One of the commonly used are global and local I Moran's statistics. We can also calculate Geary, Getis and Horde coefficients. Global I Moran's statistics is used to test the existence of global spatial autocorrelation and it is given by:

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

where, $s^2 = \frac{1}{n}\sum_i (x_i - \overline{x})^2$ is an observation in the region, and $\overline{x}$ is the average of all regions studied, n is the number of the regions, w_{ij} centipede is an element of a spatial array by weight.

Moran's statistics can take two forms, depending on its assessment - normal or of randomization [4]. Therefore moments to test the null hypothesis are calculated on the assumption of normality or randomization [3]. The Moran's statistics has a value in the range of -1 to 1. A value of 0 means no autocorrelation, negative values - negative autocorrelation, which means there are 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, spatial clusters. This is comparable to the diffusion process. In the case of a negative spatial autocorrelation neighbouring areas are different, more than it would appear from a random distribution. This is called. checkerboard effect.

A graphical presentation of global Moran's statistics is a scatter chart, which is used to visualize the local spatial relationships. The chart on the horizontal axis is standardized, analysed variable on the vertical axis test, standardized with spatially delayed variable [6]. The chart allows for the regression line and it is divided into four quadrants (HL, HH, LL, LH) versus zero point.

Table 1. Graphic presentation of global Moran's statistic

	Low values in the neighbouring regions (L)	High values in the neighbouring regions (H)
High values in the region i (H)	Negative spatial autocorrelation (square HL)	Positive spatial autocorrelation (square HH)
Low values in the region *i* (L)	Positive spatial autocorrelation (square LL)	Negative spatial autocorrelation (square LH)

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

Statistics for determining spatial autocorrelation can be used to identify spatial arrangements. For this purpose, a local ratings spatial relationships LISA proposed by Anselin in 1995 that allow determination of the similarity of the spatial entity with respect to its neighbours and to examine the statistical significance of the compound [1]. LISA for each observation indicates the degree of importance of the spatial concentration of similar values around the analysed spatial unit, for all observations the sum LISA is proportional to a global indicator of spatial dependence. The article as LISA there was local I Moran's statistics used.

Local I Moran's statistics measures whether the region is surrounded by the neighbouring regions of similar or different values of the test variable with respect to the random distribution of the values in the space [6]. li is a smoothed index for individual observation, which can be used to find local clusters. Local statistics is given by:

$$I_i = \frac{(x_i - \bar{x})\sum_j w_{ij}(x_j - \bar{x})}{\sum_i \frac{(x_i - \bar{x})^2}{n}}$$

where elements w_{ij} of the centipede come from spatial weights matrix standardized by lines. Tests of significance statistics are based on distributions arising from the conditional randomization or permutation. Standardized, local Moran's statistics has a value significantly negative when the region is surrounded by regions with significantly different values of the test variable, which is interpreted as a negative autocorrelation. Acceptance of significantly positive values means that the region is surrounded by similar neighbouring regions and there occurs regional clustering. The absolute value of the local Moran's statistics can be interpreted as the degree of similarity or differentiation.

3. An example of the use of statistics spatial dependence

To carry out the analysis presented in the article here was applied the data on the average price of 1m2 of residential premises in 26 districts of the Subcarpathian county. The data source was the contract of sale on the primary market, the secondary market and the one presented in the offer. The summary was generated from the AMRON database. The terms of the transaction involved the entire 2014 year. In addition, the report was enriched with the macroeconomic data: the population of working age and the registered unemployment rate. Spatial distributions illustrated in the Arc View GIS analysis were performed using Statistica software, PQStat. To describe the spatial relationship there was the matrix based on space generated. During the analysis, the two types of neighbourhood matrices; basic binary matrix and the first row matrix standardized by rows. Such prepared the database was used to calculate the global and local Moran's autocorrelation. Pre-generated was the basic descriptive statistics on the average, median and standard deviation of the test variable and selected macroeconomic data for the region (Table 2).

Table 2. Values of basic descriptive statistics

	Average	Mean	Standard deviation
Price 1m2 [PLN]	1966.40	2430.98	1448.6
Population at the working age	89731.52	49029.00	172716.5
The registered unemployment rate	16.24	16.30	4.4

Data analysis was preceded by calculating the correlation coefficients for the test variable and two macroeconomic data. The values obtained indicate a slight impact of the number of people of working age for each region per the price of a dwelling. More important is the local unemployment rate. It is a negative correlation, confirming the opinions of Subcarpathian region as being poor (Table 3).

Table 3. The values of correlation coefficients

	Price 1m2 [PLN]	Population at the working age	The registered unemployment rate
Price 1m2 [PLN]	1.00000	0.127055	–0.225055
Population at the working age	0.12705	1.000000	–0.243818
The registered unemployment rate	–0.22500	–0.243818	1.000000

With the assumed significance level of 0.05, there was global correlation Moran's coefficient set, which for the test variable is $I_g = -0.121676$. There was a scatter chart of global Moran's statistics drawn. Points are placed in only three squares HL, HH, LH. The district of Rzeszow is an outlier. Distribution points indicate negative autocorrelation, or price differences due to the counties.

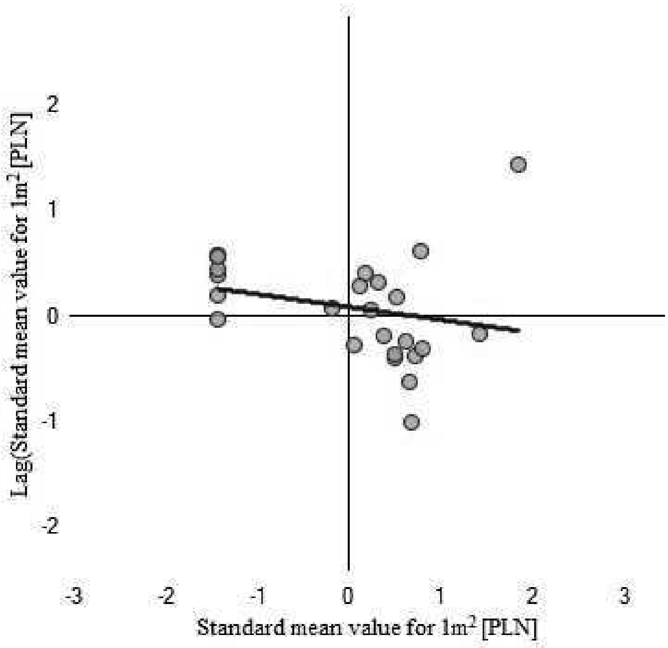

Figure 1. Scatter chart of the global Moran's statistics

Subsequently, the local Moran's statistics were determined (Table 4). Rzeszow district adopts the essential, positive, which means that it is surrounded by counties with similar values (cluster).

4. Conclusion

Analysis of global and local indicators of spatial dependence can be successfully used in the economic analysis, including real estate market research. Spatial autocorrelation statistics which indicate the type and strength of spatial dependence allow expansion of the traditionally used measures. These statistics allow observations of changes taking place in the regions. The analyses allow us to compare the economic processes, they become the basis for business decisions. The key issue is the choice of weights matrix, strongly associated with the tested regions.

The conducted spatial analysis showed the differences between the mean price per $1m^2$ of a residential unit in Subcarpathian region. The highest mean prices of the residential unit are in Rzeszow and they significantly impact the lower prices in neighbouring district of the region.

Table 4. Values of the local Moran's statistics

Location [counties]	Local Moran's statistics values	p-value
bieszczadzki	–0.22	0.8289
brzozowski	–1.95	0.0515
debicki	–0.32	0.7507
jaroslawski	–0.39	0.6932
jasielski	–1.04	0.2963
kolbuszowski	0.32	0.7464
krosnieński	0.07	0.9446
lezajski	0.15	0.8792
lubaczowski	–0.12	0.2642
lancucki	0.95	0.3430
mielecki	–0.84	0.4033
nizanski	–0.51	0.6085
przemyski	0.34	0.7345
przeworski	0.38	0.7012
ropczycko-sedziszowski	–1.24	0.2146
rzeszowski	–0.90	0.3691
sanocki	–0.39	0.6950
stalowowolski	–1.19	0.2322
strzyzowski	–0.67	0.5049
tarnobrzeski	0.06	0.9512
leski	–1.10	0.2695
Krosno	0.05	0.9567
Przemysl	0.13	0.8950
Rzeszow	**2.73**	**0.0064**
Tarnobrzeg	–0.03	0.9732

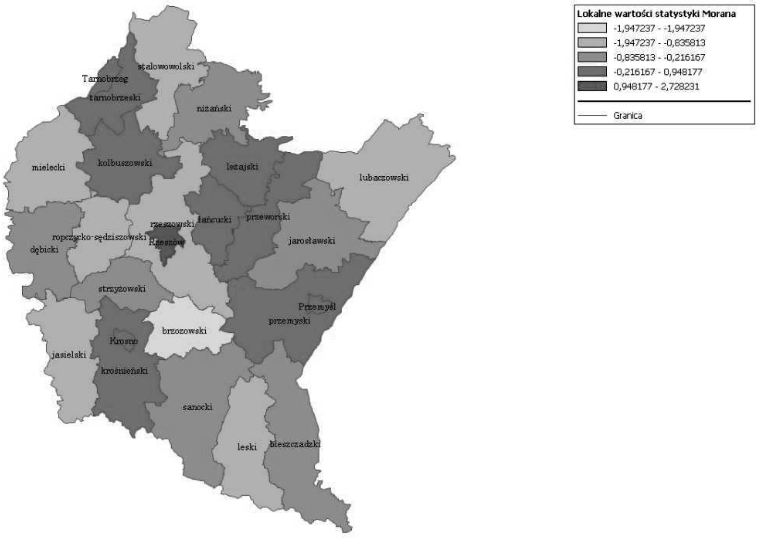

Figure 2. The chart of the local Moran's statistics for Subcarpathian counties

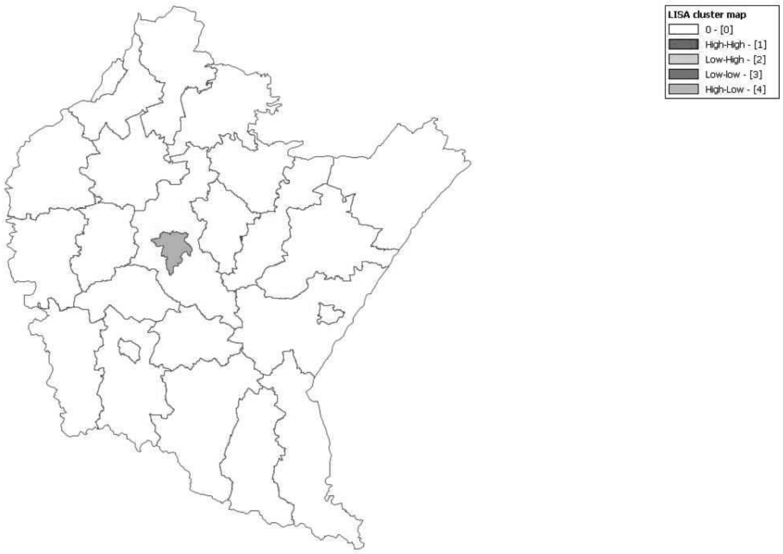

Figure 3. The chart of essential, local Moran's statistics values. The analysis supports the assessment that prices of residential properties depend on the spatial position. The high average price of $1m^2$ of a residential unit in Rzeszow and Rzeszow district generates lower prices in neighbouring counties. Macroeconomic variables have little impact on the average price

REFERENCES

[1] Anselin L. (1995) Local Indicators of Spatial Association – LISA, Geographical Analysis, vol. 27, no. 2, pp. 93-115.

[2] Aldstadt J., Getis A. (2004) Constructing the Spatial Weights Matrix Using a Local Statistics, Geographical Analysis, vol. 36, no. 2, pp. 90-104.

[3] Bivand R. (1980) Autokorelacja przestrzenna a metody analizy statystycznej w geografii [w:] Chojnicki Z. (red.), Analiza regresji w geografii, PWN, Poznań, pp. 23-38.

[4] 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.

[5] 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.

[6] Kopczewska K. (2006) Ekonometria i statystyka przestrzenna z wykorzystaniem programu R CRAN, Wydawnictwa Fachowe CeDeWu, Warszawa.