K. M. Berezka, O.M. Berezsky

Ternopil National Economic University

THE ANALYSIS OF STUDENTS ADMISSION RESULTS TO THE EDUCATIONAL QUALIFICATION LEVEL ‘’BACHELOR’S DEGREE’’ ON THE BASIS OF METHODS OF SPATIAL ECONOMETRICS

Abstract: The article deals with the disproportion estimation of undergraduates district distributions for educational qualification level "Bachelor in Science of Computer Engineering" of Ternopil National Economic University, based on the methods of spatial econometrics. There has been implemented clustering of Ternopil districts by the number of enrollees, which is based on statistical data using Moran’s statistics.

Keywords: enrollees, spatial weights matrix, Moran’s statistics, scatter plot, spatial clusters, EQL, region, district, applicant.

1. The statement of the problem

The sharp increase in the number of universities, disclosure of the same type of training areas and specialties in one city have led to a sharp drop in the number of applicants.

In order to save the incoming contingent, it is necessary to analyze the number of applicants from administrative&territorial units during previous years.

The experience shows that the vast majority of universities are mostly formed from both the locality, which the institution is located in, and the surrounding areas. Therefore, regarding the empirical analysis of the data, we can’t ignore the spatial interactions between districts, as this may lead to incorrect conclusions about the value and significance of the influence of the studied factors. Thus, the vital problem is the study and evaluation of disproportion of district distribution of enrollees for Bachelor’s Degree in Computer Engineering, basing on spatial econometrics.

In the world over the past 30 years there are popular concepts and models of new economic geography (NEG) [1-5]. One of the theoretical NEG conclusions is that in the economic development of the regions they observe agglomeration effects and spatial heterogeneity (such as the "center" - "periphery"). In terms of spatial specification of a conditional convergence model, this means that the equilibrium trajectory proportional growth (steady-state growth) regions will vary greatly depending on what spatial cluster fall into these regions.

2. The remaining part of the problem

To reflect the empirical models of spatial relationships we use spatial econometrics. It is assumed that those regions, closer to each other, integrated with each other more than those that are farther away. The main prerequisites of spatial econometrics are those, which can be correlated between the studied parameters of different objects due to geography and spatial factors. To investigate the spatial effects we use the (statistics) coefficients like Moran's (I), Jiri' (C), and others ones. [1-6].

The spatial weight matrix plays the main role in spatial analysis, showing the relations between regions and their intensity, thus formalizes the assumption that the region is spatially associated with the neighboring regions. There are various kinds of such matrix boundary neighbors, the nearest neighbors' distances etc. [1-6]. It is worth to mention that since the spatial weight matrix is formed exogenously, its specification is the most complex and controversial issue.

The aim of the study is to estimate the level of disproportion in distribution of the number of enrollees to TNEU for the specialty "Computer engineering" in the Ternopil region on the basis of spatial econometrics methods.

The methods of research are the methods of spatial econometrics, Moran's statistics.

3. The main results of the study

Determining the rate of general spatial autocorrelation, the main component is the matrix of spatial scales. Its elements w_{ij} include the influence of the region j on the region i. The matrices are quadratic; there are zeros along the main diagonal, as the region is not affected itself.

Using the matrix of the boundary neighbors we assume that the economy of the region will only affect the regions that have common borders; however the impact of the neighbors of both the 2nd and the 3rd range is insignificant. This approach is not always realistic economically and geographically. The elements of the matrix are as follows:

$$w_{ij} = \begin{cases} 1, \text{ IF region } i \text{ and } j \text{ have joint border,} \\ 0, \text{ IF } i = j, \\ 0, \text{ IF region } i \text{ and } j \text{ haven't got joint border.} \end{cases} \tag{1}$$

Another approach to the neighborhood is implemented in matrix k of the nearest neighbors:

$$w_{ij} = \begin{cases} 1, \text{ IF } d_{ij} \le d_i(k), \\ 0, \text{ IF } i = j, \\ 0, \text{ IF } d_{ij} > d_i(k), \end{cases} \tag{2}$$

where d_{ij} – the distance from region i to region j (calculated as the distance between regional centers); $d_i(k)$ – the largest k of the smallest distances.

The spatial weight matrix of distances is calculated as follows:

$$w_{ij} = \begin{cases} 1/d_{ij}^{\gamma}, & \text{IF} \quad d_{ij} \le D(q), \\ 0, & \text{IF} \quad i = j, \\ 0, & \text{IF} \quad d_{ij} > D(q), \end{cases} \tag{3}$$

where $D(q)$ – quartile distances. If $q = 4$, the matrix will consider all distances, and zeros will occur only on the main diagonal. As a rule $\gamma = 2$. For d_{ij} you can choose the distance between regional centers or the minimum time spent to cover the distance between roads.

You can also use the matrix market potential, which weights are calculated using the formula:

$$w_{ij} = \begin{cases} A_j / d_{ij}^{\gamma}, & \text{IF} \quad d_{ij} \le D(q), \\ 0, & \text{IF} \quad i = j, \\ 0, & \text{IF} \quad d_{ij} > D(q), \end{cases} \tag{4}$$

A_j – an indicator of the size or capacity of the region *j*.

Moran's statistics (general spatial autocorrelation coefficient) can be obtained by the formula:

$$I = \frac{n}{S_0} \cdot \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} (Y_i - \overline{Y})(Y_j - \overline{Y})}{\sum_{i=1}^{n} (Y_i - \overline{Y})^2}, \tag{5}$$

where $S_0 = \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}$, $\overline{Y} = \frac{1}{n} \sum_{i=1}^{n} Y_i$,

or in the matrix form:

$$I = \frac{n}{S_0} \cdot \frac{Z'WZ}{Z'Z}, \tag{6}$$

where $Z = Y - \overline{Y}$.

If $I > 0$, then there is a positive spatial autocorrelation, i.e. in total, values of observations in neighboring regions are similar; if $I < 0$ – negative autocorrelation, i.e. in total, values of observations in neighboring regions are different; if $I = 0$ – the values of observations in neighboring regions are placed randomly. The coefficient I shows the degree of linear correlation between the vectors *Z* of the centered values of *Y* indication and vectors *WZ* of the spatially centered weighted values of *Y* indication in neighboring regions, called as a spatial lag.

With the help of the spatial dispersion diagram we can make a visualization of the spread of values relating to spatial lag symptoms. On the *x*-axis we put the value of the vector *Z*, on the vertical axis we put the value of *WZ*. The regression line *WZ* on *Z* is depicted on the diagram, tangent of the slope angle of which is equal to the coefficient of the total spatial autocorrelation (if the matrix of spatial scales is standardized in rows (sum of the weights of the order is equal to one unit).

Our investigation has covered the period from 2010 to 2014, which is a sufficient time period to detect the dependencies. We have allocated 17 districts of Ternopil region in Ukraine (numbered areas are shown in Table 1) and the estimated number of enrollees for each of them.

Table 1. List of districts of Ternopil region

1	Ternopil district	10	Kremenets district
2	Berezhany district	11	Lanovtsi district
3	Borshchiv district	12	Monastyrysk district
4	Buchach district	13	Pidvolochysk district
5	Husyatyn district	14	Podghayetskyi district
6	Zalishchyky area	15	Terebovlya district
7	Zbarazh district	16	Chortkiv district
8	Zborov district	17	Shumsky district
9	Kozova district		

The scatter diagram of Moran, concerning distribution of foreign investments and using the neighboring boundary matrix, is shown in Fig. 1.

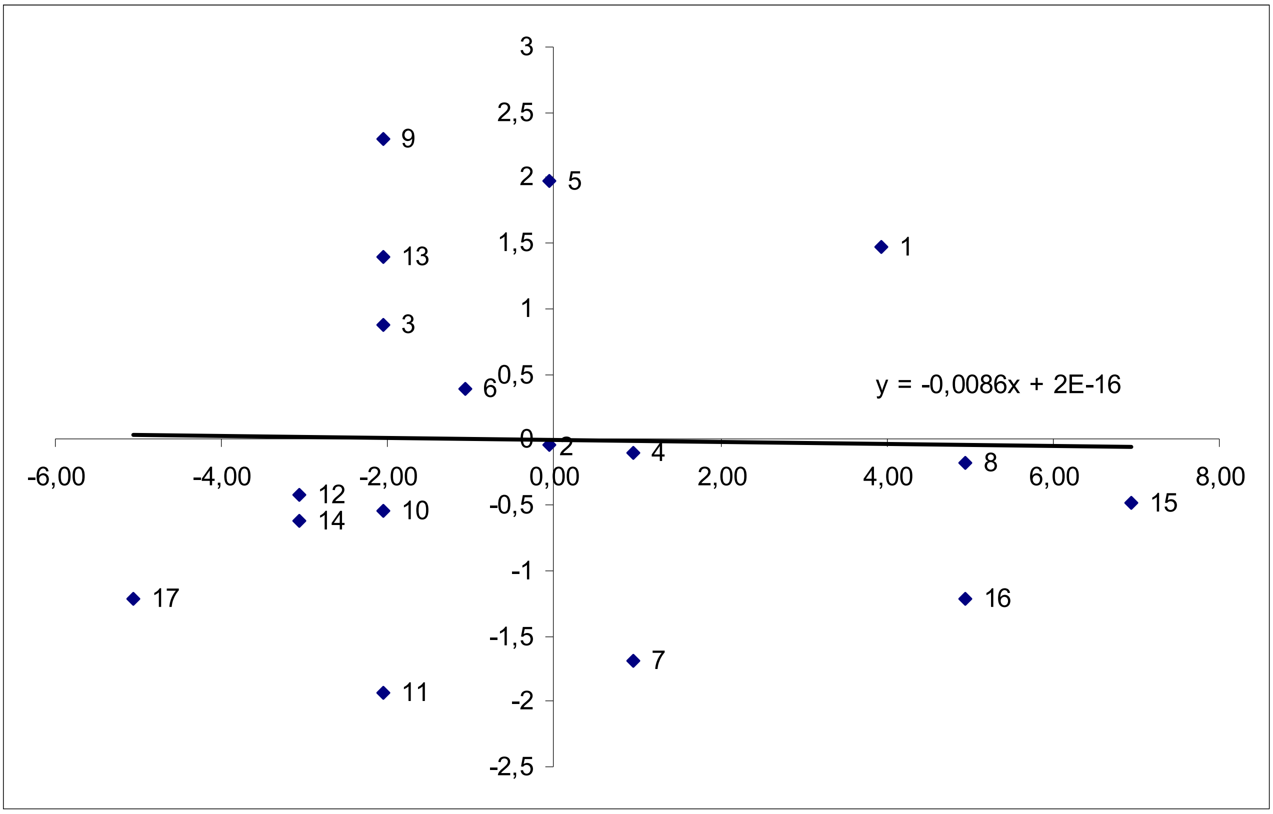

Fig. 1. Scatter diagram of Moran for the number of students qualifying EQL "Bachelor" , specialty "Computer Engineering" over 2010 - 2014 by the matrix of boundary neighbors
Source: Compiled by the authors themselves.

In the first cluster — the districts with a large number of applicants, which are surrounded by areas with a big number of enrollees – only the Ternopil district occurred in the boundary neighbors matrix.

In the second cluster – districts with few applicants, that are surrounded by areas with many enrollees - the boundary neighbors matrix has got Borshchiv, Zalishchyky, Husyatyn, Pidvolochysk, and Kozova areas.

In the third cluster III – areas with few applicants, which are also surrounded by areas with few enrollees — included Berezhansky, Monastyrysk, Podhayetskyy, Kremenets, Lanovetsi, and Shumsk regions.

In the IV-th cluster - districts with a large number of applicants, which are surrounded by areas with a small number of applicants – Buchach, Zbarazh, Zborov, Terebovlya, and Chortkiv districts occurred in the boundary neighbors matrix.

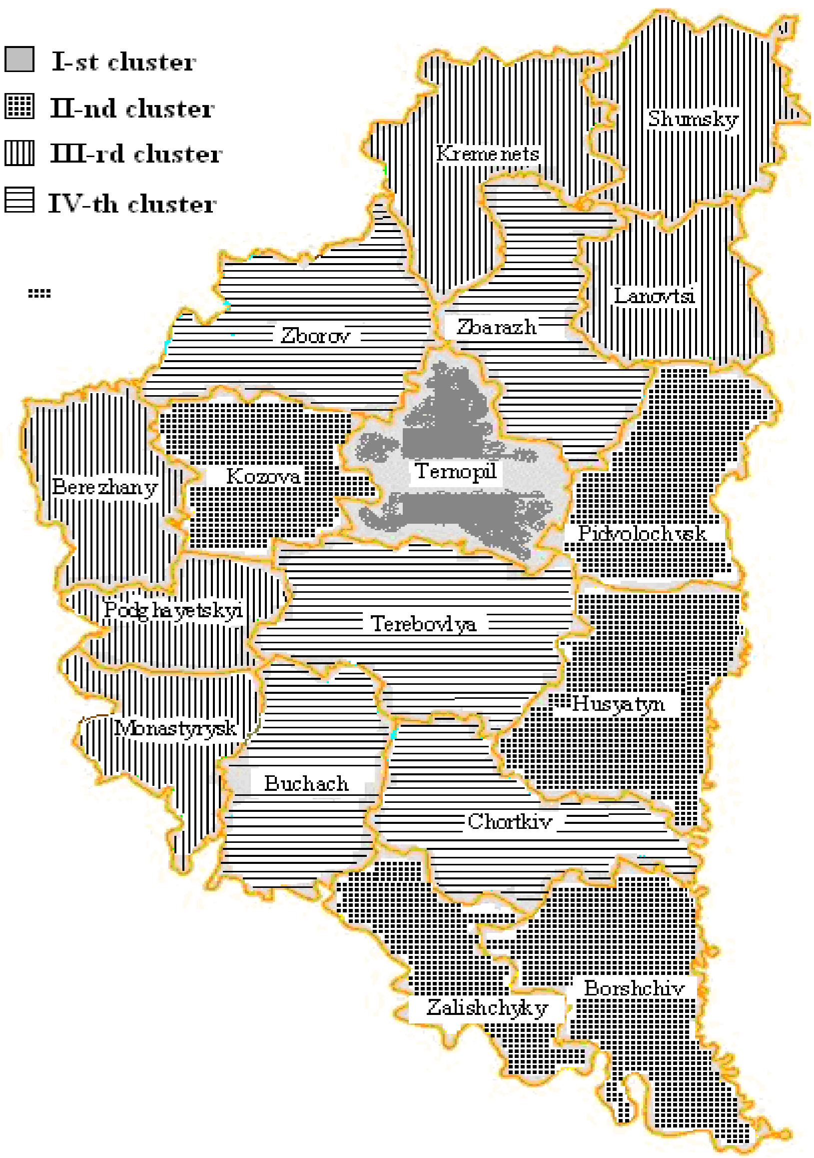

Fig. 2. Clustering regions in the Moran's diagram (for matrix boundary neighbors) (authoring)

Source: Compiled by the authors themselves.

Based on the Moran's statistics it is shown that there is negative autocorrelation, i.e. the number of applicants in neighboring areas varies. Since, $I \approx 0$ thus, it can be asserted that the number of applicants in the neighboring areas is placed randomly.

For better visualization the results of clustering regions can be displayed with the help of the following cartograms (Fig. 2).

4. Conclusions

In conclusion, the clustering of districts of Ternopil region by the number of applicants is carried out. It has been done on the basis of statistical data of the results relating to admission of incoming applicants for Bachelor's Degree in Computer Engineering with the use of Moran's statistics.

In order to increase the number of applicants first of all it is necessary to conduct the career guidance in the areas that have fallen into the II-nd and III-rd clusters.

The identification of spatial correlation (clustering of regions) with the help of both the general spatial autocorrelation coefficient and the scattering diagram is only the first step in the spatial analysis. Statistics I shows, that the values of the variable Y are spatially clustered in more than a random distribution, but it does not explain why it happens. Therefore, a perspective direction of the research is to test the hypotheses about the nature of spatial interactions.

References

1.L.Anselin: Spatial econometrics: methods and models, 1988 Dordrecht [u.a.], Kluwer.

2.J.P.LeSage, R. K. Pace: Introduction to spatial econometrics, 2009 Boca Raton, Fla. [u.a.], CRC Press.

3.Дубровина Н.А. Применение методов пространственной эконометрики в региональных исследованиях / Н.А. Дубровина // Бізнес Інформ – 2010 - №5(2) – С. 12-16.

4.Зверев Д.В. Субфедеральная фискальная политика в России: межрегиональные различия и связи / Д.В. Зверев, Е.А. Коломак // Серия «Научные доклады: независимый экономический анализ», № 209. Москва, Московский общественный научный фонд; Сибирский центр прикладных экономических исследований. – 2010. – 102 с.

5.Экономико-географические и институциональные аспекты экономического роста в регионах / Консорциум по вопр. приклад. экон. исслед., Канадское агентство по международ. развитию [и др.]; [О. Луговой и др.]. – М.: ИЭПП, 2007. – 164 с.

6.Березька К.М. Оцінка диспропорцій регіонального розподілу іноземних інвестицій в Україні / К.М. Березька, О. М. Березький, В.В. Маслій // Актуальні проблеми економіки. – 2013. – № 12(150). – С. 106 – 114.

АНАЛІЗ РЕЗУЛЬТАТІВ ВСТУПУ СТУДЕНТІВ НА ОКР «БАКАЛАВР» НА ОСНОВІ МЕТОДІВ ПРОСТОРОВОЇ ЕКОНОМЕТРИКИ

Анотація. У статті здійснено оцінку диспропорції районного розподілу вступників на освітньо-кваліфікаційний рівень "Бакалавр""Комп'ютерна інженерія" Тернопільського національного економічного університету на основі методів просторової економетрики. На основі статистичних даних з використанням статистики Морана здійснено кластеризацію районів Тернопільської області за кількістю вступників.

Ключові слова: вступники, матриця просторових ваг, статистика Морана, діаграма розсіювання, просторові кластери.