Spatial and Econometric Modeling of the Demographic Situation in the Kemerovo Region

M Novoseltseva, S Gutova, K Glinchikov

Kemerovo State University, Kemerovo, Russia

E-mail: man300674@gmail.com

Abstract. The article presents an analysis of the demographic situation in 34 municipal formations in the Kemerovo Region of the Russian Federation (16 urban districts and 18 municipal districts). The methods of spatial-econometrics modelling are used to study the factors affecting depopulation in the region and problems of the spatial arrangement and heterogeneity of the 24 single-industry cities. Calculating the local and global Moran’s index revealed positive autocorrelation between the population of the regions and districts, which makes it necessary to consider spatial interconnections while creating prediction models. In order to downweight the attribute space, four integrated indexes, that characterize integrated assessments of the levels of education, ecology, healthcare and housing conditions were created. The SAR-model was built, reflecting the effect complex assessment of the social-ecological infrastructure development has on the relative population growth in municipal formations of the Kemerovo region. The assessment of factors, affecting population in the territorial units, have shown that the most influential are healthcare and housing. Based on the data obtained, it was concluded that differentiation, taking place in the development of regions and district of the Kuznetsk Basin, is mostly due to internal factors. The results testify to the need for a regionally differentiated approach when managing the demographic situation in the Kemerovo region.

1. Introduction

Economic space cannot be heterogeneous since economic activity is concentrated in particular areas, each having their specific competitive advantages, and this activity causes the movement of people and capital from one territory, and concentration thereof in the others. Hence, research into depopulation processes in constituent entities of the Russian Federation must be based on the assessment of this connection.

By its population, the Kemerovo Region (KR) occupies the 17th place among the 85 constituent entities of the Russian Federation and the 3rd place among other regions of the Siberian Federal District. Depopulation in the regions largely stems from the outflow of the population due to interregional and intraregional migration. During strategic planning of social and economic development of municipal formations (MFs) of KR, one must select promising avenues for demographic development of the region and assess the factors affecting the depopulation process.

2. Assessing spatial autocorrelation among the municipal formations of the Kemerovo Region

In recent years, the methods of spatial econometrics have become widespread when examining regional processes [1-10]. These methods greatly extend the possibilities of statistical analysis, as they allow considering the relative position of the objects and spatial variability of the considered entity.
In order to test a hypothesis on presence of spatial autocorrelation, the Moran’s index is used: global $I_M$ and local I (LISA) [1-3, 17]. $I_M$ ranges from -1 to 1. If $I_M$ is significant and positive, then there is positive spatial autocorrelation. In economic sense, this is consistent with the clusterization of the districts of KR with similar levels of the considered parameter. If the index is negative, then the territories are isolated, and the neighboring regions differ from one another. The local Moran’s indices reveal the presence and nature of the connection between a particular region and the others.

The global Moran’s index is calculated based on the spatial weight matrix $W$ [1-3, 5, 11, 14-16]. Significance of $I_M$ is found by means of the Z-criterion [1-3, 16], whose null hypothesis means the analyzed indicators are allocated randomly. The alternative hypothesis is that the spatial effect does indeed exist.

While constructing the predictive autoregressive models, three basic types of spatial dependence are used [3, 15, 17, 19, 22]. Spatial lag might be included in the dependent variable (SAR – Spatial Autoregressive Model), the explanatory variables (SDM – Spatial Durbin Model) or residuals (SEM – Spatial Error Model).

The article considers creating a predictive model of relative population growth in KR on indicators in 2016 in 16 metropolitan districts (MDs) and 18 municipal areas (MAs).

The matrix of inverse distance weights was selected as the spatial weights matrix $W$ [13, 15, 18]. $I_M$ was 0.615 and significant. As a result, we accept the hypothesis of spatial autocorrelation. Given this fact, it is assessed that there is non-random spatial relationship between the municipal formations (MFs) of the Kuznetsk Basin in its growth and decline of the relative population growth.

The MF clusterization results are presented in Table 1. The HH cluster has 12 MFs with high indicators of the relative population growth, which causes population growth in their neighboring areas. The Kemerovo urban district (UD) and Kemerovo MAs are outliers, having distinctively high indicators compared to other administrative units. They are the centers of attraction, positively affecting the population in the surrounding MDs. The LL cluster consists of 11 MDs with very low population numbers, which has a negative effect on the indicators of the neighboring municipal formations.

Clusters LH and HL correspond with negative spatial correlation, which was indicative of the isolation of the territories. LH is the cluster of MDs with low indicators, surrounded by MDs with high indicators. Due to negative autocorrelation, we observe the outflow of the population into the neighboring districts, such as the Kemerovo MFs. The MFs of the HL cluster have high indicators, being surrounded by MFs with the low ones. Development there takes places not by means of interaction, but the outflow from the neighboring regions with low indicators.

| Cluster | Municipal formations of the Kemerovo Region |
|---------|---------------------------------------------|
| LH      | Kiselevsky UD, Myskovsky UD, Osinnikovskiy UD, Prokopyevsky UD, Guryevsky MD, Promyshlennovskiy MD, Yurginsky MD, Yashkinsky MD |
| HH      | Berezovsky UD, Kaltan UD, Kemerovo UD, Krasnobrodsky UD, Mezhdurechensky UD, Novokuznetsk UD, Taiginsky UD, Kemerovo MD, Krapivinsky MD, Novokuznetsk MD, Tashtadul MD, Topkinsky MD |
| LL      | Anzhero-Sudzhensky UD, Belovo UD, Lenin-Kuznetskiy UD, Belovsky MD, Izhmorsky MD, Lenin-Kuznetskiy MD, Mariinsky MD, Tisulsksy MD, Tyazhinsky MD, Chebulin MD, Yayskiy MD |
| HL      | Polysaevsky UD, Yurginsky UD, Prokopyevsky MD |

LISA provides a more thorough insight into the links between the population sizes in regions [16]. Negative LISA indicates the isolation of an MF. The higher the value, the stronger the links between
the MF. It is worth noting that LISA of 11 MF are less than zero. The LISA ranges from -0.02 to 0 and is localized in the LH and HL clusters. These are not typical territories, isolated and distanced from the favorable MF of the HH quadrant. Development of said regions takes place only thanks to internal population growth.

The interval from 0 to 0.02 consists of 19 fairly prosperous MFs of HH and LL clusters. The highest LISA values indicate the points of population growth, which occurs due to them being in close proximity to Kemerovo and Novokuznetsk. Four pronounced centers of attraction, the MD and UD of Novokuznetsk, the MD and UD of Kemerovo continue to lead with their LISA being 0.02 and higher.

The Kuznetsk Basin is the region where 24 cities are single-industry, and their development is a pressing challenge and depends on a specific industry. The single-industry cities of KR may be divided into 3 groups. The first group is the cities with the most difficult socio-economic situation: Yurga, Anzhero-Sudzhensk, Prokopyevsk, Salair, Tashtagol. The second group is the cities with the risk of deterioration of the socio-economic situation: Mariinsk, Guryevsk, Topki, Yashkino, Myski, Taiga, Mezhdurechensk, Osinniki, Leninsk-Kuznetsk, Berezovsky, Polysaev, Belovo, Kiselevsk, and the urban villages of Mundybash, Sheregesh and Krasnobrodsky. The third group is the cities with stable socio-economic situation: Novokuznetsk, Kaltan, the Belogorsk urban village.

The LH and HL clusters with negative spatial autocorrelation and lack of connections consist of 60% of single-industry cities with difficult socio-economic situation. They are isolated, each with their own specific features, and there is a possibility of the outflow of population into more prosperous cities. The single-industry cities of the second ground are mostly (62.5%) within the quadrant with a positive spatial correlation. All the cities of the third group fall into either LL or HH. For instance, the Belogorsk urban village of the Tisulsk MF faces decreasing population growth. Population growth in Kaltan and Novokuznetsk, on the contrary, causes population growth in the neighboring areas. It is partially due to creation and funding of the MD development projects.

Calculation of the local indices has shown that the most favorable single-industry cities of the third group have positive LISA values. The single-industry cities of the second group declusterized with LISA values from -0.02 to 0.02. Three cities of the first group, specifically, Yurga, Prokopyevsk and Salair have negative LISA values, which accounts for their territorial isolation and difficult demographic situation. In these cities, the negative dynamic of the index of industrial production persists, and draining of the human resource into the nearby economically favorable cities causes depopulation. Two cities of the first group, namely Anzhero-Sudzhensk and Tashtagol fall within the LISA range from 0 to 0.02 which is further corroborated by strengthening spatial linkage of said territories due to fairly high rates of socio-economic development of these cities.

In order to formulate the strategic solutions, it is necessary to predict the relative population growth. Regression analysis is traditionally used to create predictive models. However, positive spatial autocorrelation of the relative population growth in KR results in the need to consider the spatial interaction while creating the regression models. The SAR-model with spatial autoregression lag \([1, 3, 8, 19, 22]\) is the following:

\[
z_t = \rho Wz_{t-1} + X\beta + \varepsilon,
\]

where \(z\) – the vector of the examined parameter, \(Wz\) – its spatial lag, \(X\) – the vector of independent feature variables, \(\varepsilon\) – the vector of the model residual, \(\beta\) – the vector of the assessed parameters, \(\rho\) – the coefficient that characterizes spatial autocorrelation. In order to downweight the attribute space, four integrated indexes \([20-21]\), that characterize integrated assessments of the level of education, ecology, healthcare and housing conditions were created. The modelling results are presented in Table 2.
Assessment of the factors that define the relative population growth in MF has shown that with the confidence figure of 0.95 the most significant factors are healthcare and housing conditions. The $\rho$ index of 0.00004, characterizing spatial autocorrelation, is close to zero. Consequently, differentiation, which takes place in the development of the Kuznetsk Basin MD, is mostly caused by internal factors.

3. Conclusions
The study have shown positive correlation between the levels of population size in the Kuznetsk Basin MD, which indicates the territorial link and clusters with high and low values of relative population growth. Spatial analysis of heterogeneity of single-industry cities in KR for the demographic indicator was carried out. It was found that healthcare and housing conditions have a positive effect on the population growth in the territorial units of KR. However, sustainable links between the MD are primarily caused by internal factors and their further development must be promoted on the level of Kuznetsk Basin region MF. This testifies to the need for a regionally differentiated approach when managing the demographic situation in the Kemerovo region.

4. References
[1] Anselin L 2006 Spatial Econometrics. In: Palgrave handbook of econometrics: 1 Econometrics theory 901–941 (Palgrave Macmillan, Basingstoke)
[2] Anselin L 2007 Spatial econometrics Palgrave handbook of econometrics (Palgrave Macmillan, Basingstoke)
[3] LeSage J, Pace R K 2009 Introduction to Spatial Econometrics Chapman & Hall/CRC (Florida)
[4] Kelejian H H, Prucha I R 2010 Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances Journal of Econometrics 157 (1) 53–67
[5] Aldstadt J, Getis A 2006 Using AMOEBA to Create a Spatial Weights Matrix and Identify Spatial Clusters Geographical Analysis 38 (4) 327–343
[6] Baltagi B H, Fingleton B, Pirotte A 2014 Estimating and Forecasting with a Dynamic Spatial Panel Data Model Oxford Bulletin of Economics and Statistics 76 (1) 112–138
[7] Debarsy N, Ertur C 2010 Testing for spatial autocorrelation in a fixed effects panel data model Regional Science and Urban Economics 40 (6) 453–470
[8] Elhorst J P 2012 Dynamic spatial panels: models, methods, and inferences Journal of geographical systems 14 (1) 5–28
[9] Elhorst J P 2014 Spatial econometrics: from cross-sectional data to spatial panels Springer, Heidelberg, New York, Dordrecht, London
[10] Franzese R J, Hays J C 2007 Spatial econometric models of cross-sectional interdependence in political science panel and time-series-cross-section data Political Analysis 15 (2) 140–164
[11] Semerikova E V 2014 Unemployment in West and Easy Germany: spatial analysis of panel data Applied econometrics 35 (3) 107–132
[12] Patuelli R, Arbia G 2016 Spatial Econometric Interaction Modelling. Springer International Publishing (Switzerland)
[13] Vakulenko E 2015 Analysis of the relationship between regional labour markets in Russia using Okun’s model Applied Econometrics 40 (4) 28–48
[14] Demidova O 2015 Spatial effects for the eastern and western regions of Russia: a comparative analysis International Journal of Economic Policy in Emerging Economies 8 (2) 153–168
[15] Semerikova E, Demidova O 2015 Analysis of regional unemployment in Russia and Germany: spatial and econometrics approach Spatial economics 42 (2) 64–85

Table 2. SAR-model coefficients.

| Attribute | Ecology | Education | Housing conditions | Healthcare |
|-----------|---------|-----------|-------------------|------------|
| Coefficient | 0.000506 | 0.000690 | 0.001265 | 0.015226 |
[16] Pavlov Yu V, Koroleva E N 2014 Spatial interactions: assessment based on global and local Moran’s indices Spatial economics 3 95–110

[17] Zhukova A K, Silaev A M, Silaeva M V 2016 Analysis of life expectancy considering spatial dependence in regions: Spatial economics 4-5 (48) 112–128

[18] Semerikova E V, Demidova O A 2016 Interaction of local labor markets in Russia: analysis by means of spatial econometrics models Spatial economics 3 57–80

[19] Gafarova E A 2017 Empirical models of regional economic growth with spatial effects: comparative analysis results The Bulletin of Perm State University 12 (4) 561–574

[20] Kagan E S 2012 Application of the Analysis Hierarchy Method and Fuzzy Set Theory to Evaluate Complex Social-Economic Phenomena Izvestiya of Altai State University 1(1), 160–163

[21] Levina E I 2016 Techniques of assessing demographic processes in a region The Bulletin of Tomsk State University 3 (35) 185–195

[22] Arbia G 2014 A Primer for Spatial Econometrics: With Applications in R. Palgrave Macmillan UK, Basingstoke