ANALYSIS OF RURAL AREAS DEVELOPMENT OF THE REGION USING THE ADL-MODEL

Nikolay Didenko[1], Djamilia Skripnuk[1], Olga Mirolyubova[1], Maria Radion[2]
[1]Peter the Great St. Petersburg Polytechnic University, Russia
[2]Tyumen State University, Russia
didenko.nikolay@mail.ru

Abstract
The paper presents an approach for the analysis of rural areas' socio-economic development. Rural regions include spheres of human activity: institutional, environmental, social, labor, industrial, economic, technological development spheres. Under the development of the rural regions, changes in the variables of different spheres of human activity in the region, increasing the potential for human development, is understood. Human Development Index measures the potential for human development measurement. The model of socio-economic development of the rural regions presented in the article is characterized by both endogenous and exogenous variables of the model. Model of socio-economic development of the rural regions is presented in the form of six econometric equations, each of which is an ADL-model. Each ADL-model contains endogenous and exogenous variables. The article presents the structural form of a system of economical equations in each one equation there is more than one endogenous variable. It shows the transition from the structural form to the form of the system of economical equations, where each equation has only one endogenous variable. The system of equations was developed for the rural regions of Russia. Baseline data were obtained from the database of the State Statistics Committee of Russia. The article provides a solution to the system of equations. The solutions to the model can be used for developing strategic development of each separate rural region or a combination of all the rural regions.
Key words: rural regions, socio-economic development, ADL-model.

Introduction

The development of the rural regions is one of the key targets for the regional politics. However, today there is no single process to measure this process. And there are several reasons for it. First of all, regions differ in terms of socio-economic and technological development (Tritsch & Arvor, 2016). This fact is a major obstacle to choose the best method and to check the indicators of development of the rural regions with different regional qualities. These indicators of the rural development are usually chosen within helix dependence from the set goal to the subject matter. Secondly, there is no unified opinion, which methods to be used when measuring the development of the rural regions.

Chromy *et al*. (2011) considering rural regions as any country area, conclude that the key factors of analysis must include indicators that measure social-economic and institutional environment. Looking at the example of the Czech Republic (Straka & Tuzová, 2016), 14 indicators can be highlighted, mostly the social-economic development indicators.

Some studies employ such indicator as the agricultural land usage indicators. Bonfiglio, Arzeni, & Bodini (2017) measure eco-efficiency of arable farms in rural areas. Authors conclude that farms are more eco-efficient if they are led by young farmers and participate in agri-environmental schemes. Bański & Mazur (2016) classify the rural regions using the cognitive methods. The authors concentrate on the methodological procedure of the new typology, which accounts for three aspects: development dynamics, economic structure and transport-wise accessibility.

Provenzano, Arnone, & Seminara (2016) implanted the Smart Specialization Strategy and the Quintuple Helix Model, highlighting the importance of connecting the innovation process with rural territories. Jeong *et al*. (2016) introduce a hybrid model, as the combined application of the geographic information systems (GIS) and the multi-criteria decision analysis (MCDA) with the fuzzy DEMATEL method in order to make rural housings' suitability map. The sixteen criteria were clustered into four groups, namely physical, natural, environmental and socio-economic group. Weight coefficients are also proposed to be used to distinguish the importance of the factors within the model.

The purpose of the study is to design a mathematical model for the development of rural areas. The mathematical model must contain several interrelated equations. Each equation connects the resulting (endogenous) and influencing (exogenous) variables. The authors intend to use such a model for analyzing and forecasting the indicators of rural development. It is possible to see how endogenous variables behave with the change in exogenous variables.

The local tasks of the study are the following:
- Describe the object of research - a specific rural area;
- Choose the type of mathematical model;
- Describe the method of empirical verification of the model;
- Check the model using the statistics of a particular rural area.

Materials and Methods

We analyze the development of Yalutorovsk municipal area for the period 2000 – 2015. The area is a part of the Tyumen region, one of the typical Russian regions. Data was taken from the Federal State Statistics Service of the Russian Federation (http://www.gks.ru). The choice of the Tyumen region is proved in resent papers by Rudenko (2015) and Tcvetkov *et al.* (2015). The area has around 50 thousand people. The population has been growing since the year 2000. The municipal territory is located in a relatively good place regarding the climate. The area has only one large populated community (a small town) and around 30 small villages. Around 50 percent of the local population is in the populated areas, which is relatively far from the average Russian statistics, which says that 74 percent of the population lives in cities. The third part of population is concentrated in urban settlement and the other parts of population are distributed over the territory of the district. The area is a rich region for natural resources. Analysis of infrastructural industry (electrical, transport, and informational industries) shows us the region as the area with contradictory state of the economic spheres. On the one hand, per capita incomes of the population are higher than in other regions and the average monthly nominal wage is quite high. On the other hand, financial indexes of the region show us that the area is subsidized. Analysis of the industry differentiation of the municipal territory shows us the absence of differentiation of the economy, the dominating of livestock sectors and the dependence of the municipal territory on the meat market situation. Rating of the region in the ranking of subjects of the Russian Federation in terms of investment in fixed assets among the subjects is lower than the average for Russia, but it is not the smallest one.

Despite quite high income of the population, higher education is not so popular. For 1000 people only 159 have higher professional education. This is the 78th place out of 83 in Russia. In terms of population size per doctor, we can say that the health situation in the region is not so bad. However, it is not entirely clear how rural settlements are provided with health services, primarily engaged in livestock farming and located far from a large the urban-type settlement - the center of the district, equipped with hospitals. An analysis of the state of the infrastructure transport industry in the municipal territory shows that the only major population center, the urban-type settlement, is not connected with other rural settlements neither by regular bus routes nor by railway.

Basic transport for this region is a car. In addition, the region has only one power station – gas turbine power station, the total length of power transmission lines of which is about 500 km. (Romashkina,

Didenko, & Skripnuk, 2017). The municipal territory produces less electricity than it is necessary.

The housing sphere of the municipal area has poor conditions according to analysis. There were no dilapidated living quarters existent in the district at the end of 2015, but houses equipped with water supply are only 50% of the entire houses. Houses that are equipped with hot water supply and sewerage constitute only 40%.

Seventy percent of population has its own computer, fifty percent has access to the Internet, many people have a mobile phone, but the information and telecommunication environment of the territory is poorly developed.

The region takes the last place in the rating of innovation in the Russian regions. Looking at the analysis, we can state the following conclusions: there is underdeveloped infrastructure - lack of standard roads, railways, and low power supply and communication services. There is lack of population, and therefore lack of permanent workforce. The expansion of the range of products, reorientation of sales markets, and development of new types of production in order to increase production efficiency is not carried out in the region.

The analyzed territory depends on the supply of goods from other districts of Russia. The transportation of goods is hampered by the poor condition of the transport infrastructure. There is insufficient production efficiency. The analyzed territory has a small percentage of the population with higher education; there is no right number of qualified staff, which creates an obstacle for innovative development.

Development of the region is usually understood as the process of the steady positive development of the area, in which an individual is living and working. The development of all the spheres should be interconnected from within. The spheres in which a person is making a living are usually the following: natural environment, production and manufacturing sphere, social systems, ecological system, public institutions. The scientific and technological progress develops all spheres of life and working conditions of a person. The science-technological progress is the key factor in the change of all the areas in which an individual is living and working. In this case, sustainable development is considered from the point of view of the following three components. The first component is the economical component. It is necessary to use environmentally sound technologies in accordance with the economic component of sustainable development. We must use the energy (or natural materials) effective technologies at all stages of the product life cycle. The second component is the social component. The social component of sustainable development means an increase of

living standards of the population of the territory. The third component is the ecological component. The conservation of ecosystems, careful attitude to natural resources, conservation of flora and fauna is the goal of the ecological component of sustainable development.

Based on the above, we can give a general definition of the sustainable development of the analyzed territory. The sustainable development of the analyzed territory is the policy of using the territory, which characterizes the following conditions: a) a uniform change of the areas that surround a human in the territory – flora and fauna, sphere of production, social institutes, ecological systems; b) it is necessary to assess the change in the spheres of human activity from the point of view of economic, social and environmental perspectives; c) the purpose of changing the areas of the territory is to increase the human development index (HDI).-

The theory of spatial economics and the theory of modeling is the methodological base of the sustainable development of the analyzed territory.

The theory of spatial economics is a modern discipline, but the development of theory for a specific territory that gave a push towards spatial economics development comes from the times of Adam Smith and David Ricardo. The significant contribution towards the development of the scientific categories of the theory of spatial economics was made by Max Weber, J. Thünen, Wolfgang Kasper, Stanislaw Korenik and Katarzyna Mishchak, and a Nobel laureate Paul Krugman (Tabata, Eshima, & Sakai, 2015). If the task of the spatial economics is the coverage of the wide region of the spatial phenomena for the development of the region, then the main task of the theory of the modeling is in giving the research such a technology to create such models, which would be enough to fully record the interesting properties of the space and then it would be quicker and easier to research the data and transfer the results on to the original (Grazi, Bergh, & Rietveld, 2007). Thus, using the mechanisms of the theory of spatial economics and the theory of modeling that include modeling of the development of the analyzed territory, it is possible to analyze the problems that the development of the rural region faces.

The ADL-model (autoregressive distributed lags model) is chosen as the theoretical model for the research, in which a regression equation is used to predict current values of a dependent variable based on both the current values of an explanatory variable and the lagged (past period) values of this explanatory variable. The model is generalized in case there is a multiple number of exogenous variables X. In general, all exogenous variables are included in the model with the same number of the lags, but there might be an exception of a lag of some variables.

ADL-model is the following:

$$y_t^i = \sum_{j=1}^{k^1} a_0 y_{t-j}^i + \sum_{j=1}^{k^2} a_1 x_{1,t-j}^i +$$

$$+ \sum_{j=1}^{k^3} a_2 x_{2,t-j}^i + \sum_{j=1}^{k^4} a_3 x_{3,t-j}^i + E_t \qquad (1)$$

Where, k^1, k^2, k^3, k^4 – quantity of lags of the variables; E_t – reminder that creates the white noise.

The model states that if in any moment of time t takes place a change in the constant variable X, then the change will influence the value of the variable Y in the future moments of time.

Results and Discussion

Primary data. The procedure for working with primary data includes two stages: a) analyzing and choosing both endogenous and exogenous variables that fit with the process that was being analyzed and that were showing the content of the variables; b) collecting the necessary both endogenous and exogenous variables.

Analysis and choice of the necessary endogenous and exogenous variables. The result of the analysis allowed us to choose the following six variables for the region: the proportion of the volume of region in the total volume of production of the region in the year t (y_t^1); the share of the export of the region from the total export of the Russian Federation in the year t (y_t^2); a level of a salary of the population in the year t (y_t^3); a pollution levels of the atmosphere in the analyzed territory in the year t (y_t^4); the volume of the shipped goods from the total volume of the shipped goods of the Russian Federation in the year t (y_t^5); the development level of the infrastructure branches of the region (y_t^6). For each endogenous parameter there are exogenous parameters, including the above stated six endogenous parameters. The task was to choose exogenous parameters for each endogenous parameter from the list of all available variables following the logical of the economic progress that is in the model.

List of exogenous variables: volume of shipped innovative goods, works, services (x_t^1); labor productivity (x_t^2); innovative technologies expenses (x_t^3); doctors salaries per 10000 population (x_t^4); growth rate of the labor productivity (x_t^5); development level of the manufacturing industries (x_t^6); proportion of the social infrastructure industries (x_t^7); energy expenses (x_t^8); average annual population level (x_t^9).

Empirical test model method. Method of empirical test model contained several stages.

(A) The correlation analysis was used to choose the exogenous variables that are closely connected with each endogenous variable. The correlation coefficients and their significance level were determined for pairs of variables.

(B) The analysis of the autocorrelation of the time series both endogenous variable and exogenous variables was performed. Analysis is done to select lags that had the highest impact on the resulting variable.

(C) Analysis of multicollinearity of exogenous variables. To find the multicollinearity of the variables, correlation matrix variables are analyzed, the importance is measured using the standard error criteria and Box-Pearsons Q-criterion. One variable from the pair of the variables was deleted from the analysis if the pair of the variables had a correlation greater than 0.8.

(D) The time series were checked for the presence of stationarity properties. We used Dickey – Fuller unit root test.

(E) The system of equations is represented in the structural form of the model. The structural form of the model is an econometric model in which the relationships between the current endogenous variables and the influencing variables are fixed in the form of equations.

(F) Determining identifiability of the equations of the structural form of the model by the criteria of identifiability. A necessary condition of identifiability is: $D+1 = H$ – equation identifiable; $D+1 < H$ – equation unidentifiable; $D+1 > H$ – equation overidentifiable, where H is a number of endogenical variables of i-equation system; $D-$ is a number of exogenous variables that are in the model, but exogenous variables are not included in the current equation. Sufficient condition of identification: equation is identifiable if by the absent variables (endogenous and exogenous) it is possible from the coefficients from the different equation systems to get matrix, a determinant that does not equal to zero, and the rank of the matrix is not lower than the number of endogenous variables in the model minus one.

The method to estimate the coefficients of the structural model was determined according to the following rules. It is necessary to choose an OLS (ordinary Least Squares) method if the model is identifiable. It is necessary to choose a two step OLS method if the model is overidentifiable. Definition for the overidentifibale equation of theoretical values of endogenous variables is from the left side of the equation. For substituting it into the actual values, we use the standard OLS method for the improved structural form of the overidentified equation.

The coefficients of the equations of the structural model were calculated and the transition from the structural model to the reduced model was carried out. The number of equations in the reduced model is equal to the number of the endogenous variables. The endogenous variable is expressed through the predefined model variables in each equation of the reduced model. The adequacy of the equations in the reduced model was evaluated on the basis of the F-criterion of Fisher. The coefficients of the regression equations in the reduced model were determined by the OLS method.

Empirical testing of the model. Empirical testing of the models' development was carried out on statistical data of agricultural region of Russia. The structural model with selected and substantiated endogenous and exogenous variables is constructed in the form of a system of equations.

From the structural form of the model a transaction was mode to an improved structural form. The reliability of each equation of the improved model is verified by Fisher's F-criterion, the coefficients of the equations were calculated, the reliability of the equation coefficients was estimated using Students' t-criteria.

For the first equation: $F_{tabl} = 9.01 < F_{actual} = 32.274$ => model is statistically correct, equation of regression is reliable for the level of significance $a = 0.05$ and is the following:

$$y_t^1 = (3.287E - 010)x_t^2 + 0.013x_t^6 - 0.054x_t^7 + \\ + (-7.913E - 008)y_{t-1}^3 + (4.607E - 009)x_t^8 + \\ + 0.004 \qquad (2)$$

For the second equation: $F_{tabl} = 6.39 < F_{actual} = 7.689$ => model statistically correct, equation of regression is reliable for the level of significance $a = 0.05$ and is the following:

$$y_t^2 = 0.368x_{t-1}^7 + (1.074E - 009)x_{t-1}^8 + \\ + 5.781y_{t-1}^1 - 0.994y_{t-1}^5 - 0.022 \qquad (3)$$

For the third equation: $F_{tabl} = 9.01 < F_{actual} = 134.535$ => model statistically correct, equation of regression is reliable for the level of significance $a = 0.05$ and is the following:

$$y_t^3 = 0.008x_{t-1}^7 + 1020194.030\,x_{t-1}^6 - 0.028x_{t-1}^8 - \\ - 2.137x_{t-1}^9 + 0.198y_{t-1}^3 + 162825.498 \qquad (4)$$

For the forth equation: $F_{tabl} = 5.59 < F_{actual} = 6.335$ => model statistically correct, equation of regression is reliable for the level of significance $a = 0.05$ and is the following:

$$y_t^4 = -78515.947\, x_t^6 + 313.464 \qquad (5)$$

For the fifth equation: $F_{tabl} = 5.41 < F_{actual} = 11.882$ => model statistically correct, equation of regression is reliable for the level of significance $a = 0.05$ and is the following:

$$y_t^5 = -0.131 x_t^6 + 0.131 x_t^7 + (2.722E - 009)x_t^8 + \\ + 0.007 \qquad (6)$$

For the sixth equation: $F_{tabl} = 9.01 < F_{actual} = 35.274$ => model statistically correct, equation of regression is reliable for the level of significance $a = 0.05$ and is this:

$$y_t^6 = (-3.919E - 008)x_t^2 + 138.472x_t^6 + \\ + 17.397 x_t^7 + (-5.130E - 007)x_t^8 + \\ + (8.183E - 006)y_{t-1}^3 + 0.306 \qquad (7)$$

The necessary conditions for identification of model were considered for each equation. All six equations were overidentified. If they were overidentified, then the whole system was overidentified, and consequently, for determination of the indexes of the equation we used two-step OLS (Ordinary Least Squares). The improved structural shape of the model was used for equations 2,3,4,5,6,7, we calculated theoretical indexes of endogenous variables for all overidentified equations, which was in the left part of the equation, substitution of the initial indexes of exogenous variables in the right part of the equations. After that, the theoretical indexes of the endogenous variables were substituted for the actual values. Herewith, the reliability of the regression equation was estimated on the basis of the Fisher F-criterion, calculated the coefficients of the equation by OLS, we estimated the reliability of the coefficients based on the T-test of the Student. The system of econometric equations in its complete form is shown below:

$$\begin{cases} y_t^1 = (3.287E - 010)x_t^2 + 0.013x_t^6 - 0.054x_t^7 + (-7.913E - 008)y_{t-1}^3 + (4.607E - 009)x_t^8 + 0.004 \\ y_t^2 = 0.368x_{t-1}^7 + (1.074E - 009)x_{t-1}^8 + 5.781y_{t-1}^1 - 0.994y_{t-1}^5 - 0.022 \\ y_t^3 = 0.008x_{t-1}^7 + 1020194.030x_{t-1}^6 - 0.028x_{t-1}^8 - 2.137x_{t-1}^9 + 0.198y_{t-1}^3 + 162825.498 \\ y_t^4 = -78515.947x_t^6 + 313.464 \\ y_t^5 = -0.131x_t^6 + 0.131x_t^7 + (2.722E - 009)x_t^8 + 0.007 \\ y_t^6 = (-3.919E - 008)x_t^2 + 138.472x_t^6 + 17.397x_t^7 + (-5.130E - 007)x_t^8 + (8.183E - 006)y_{t-1}^3 + 0.306 \end{cases} \qquad (8)$$

Conclusions

In this article, we used the theory of Spatial Economics and the simulation theory to analyze the development of the rural region. Each sphere of the rural territory was estimated by indicators, which are endogenous and exogenous variables of the model. For each endogenous indicator of the model there are concrete exogenous indicators. Selected and grounded endogenous and exogenous variables allowed us to formulate a system of equations in the so-called structural form of the model. Sometimes the model of structural form can be transformed into an independent system of equations. In our case, the model of structural form could not be transformed into a model with independent system of equations. The model of structural form was transformed into the so-called model of improved structural form. Finding the solution for the model of improved structural form is shown in the article. The empirical checking of the development model was carried out on the statistical data of the rural region of Russia. The necessary and sufficient conditions for the identification of the model have been analyzed and shown that among the equations of the model there are no unidentifiable equations, all equations are overidentified. A two-step OLS was applied and the possibility of solving the system of econometric equations was shown.

Acknowledgements

References

1. Bański, J., & Mazur, M. (2016). Classification of rural areas in Poland as an instrument of territorial policy. *Land Use Policy*, 54, 1 – 17. DOI: 10.1016/j.landusepol.2016.02.005.
2. Bonfiglio, A., Arzeni, A., & Bodini, A. (2017). Assessing eco-efficiency of arable farms in rural areas. *Agricultural Systems*, 151, 114 – 125. DOI: 10.1016/j.agsy.2016.11.008.

3. Chromy, P., Jancak, V., Marada, M., & Havlicek, T. (2011). Rural Areas - Living Space: Regional Differences in the Perceptions of Representatives of Rural Municipalities in Czechia Regarding Rural Areas. *Geografie*, 116, 23 – 45.

4. Grazi, F., Bergh, J., & Rietveld, P. (2007). Spatial welfare economics versus ecological footprint: modeling agglomeration, externalities and trade. *Environmental and Resource Economics*, 38 (1), 135 – 153. DOI: 10.1007/s10640-006-9067-2.

5. Jeong, J.S., García-Moruno, L., Hernández-Blanco, J., & Sánchez-Ríos, A. (2016). Planning of rural housings in reservoir areas under (mass) tourism based on a fuzzy DEMATEL-GIS/MCDA hybrid and participatory method for Alange, Spain. *Habitat International*, 57, 143 – 153. DOI: 10.1016/j.habitatint.2016.07.008.

6. Provenzano, V., Arnone, M., & Seminara, M.R. (2016). Innovation in the Rural Areas and the Linkage with the Quintuple Helix Model. *Procedia - Social and Behavioral Sciences*, 223, 442 – 447. DOI: 10.1016/j.sbspro.2016.05.269.

7. Romashkina, G., Didenko, N., & Skripnuk, D. (2017). Socioeconomic modernization of Russia and its Arctic regions. *Studies on Russian Economic Development*, 28 (1), 22 – 30. DOI: 10.1134/S1075700717010105.

8. Rudenko, D. (2015). Geographic information systems potential for poverty alleviation in the Tyumen region, Russia. *International Multidisciplinary Scientific GeoConference*, SGEM, 2 (2), 855 – 862. DOI: 10.5593/SGEM2015/B22/S11.108.

9. Straka, J., & Tuzová, M. (2016). Factors Affecting Development of Rural Areas in the Czech Republic: A Literature Review. *Procedia - Social and Behavioral Sciences*, 220, 496 – 505. DOI: 10.1016/j.sbspro.2016.05.525.

10. Tabata, M., Eshima, N., & Sakai, Y. (2015). Existence and computation of solutions to the initial value problem for the replicator equation of evolutionary game defined by the Dixit–Stiglitz–Krugman model in an urban setting: Concentration of workers motivated by disparity in real wages. *Applied Mathematics and Computation*, 254, 419 – 451. DOI: 10.1016/j.amc.2015.01.029.

11. Tcvetkov, P., & Cherepovitsyn, A. (2016). Prospects of CCS Projects Implementation in Russia: Environmental Protection and Economic Opportunities. *Journal of Ecological Engineering*, 17 (2), 24 – 32. DOI: 10.12911/22998993/62282.

12. Tritsch, I., & Arvor, D. (2016). Transition in environmental governance in the Brazilian Amazon: emergence of a new pattern of socio-economic development and deforestation. *Land Use Policy*, 59, 446 – 455. DOI: 10.1016/j.landusepol.2016.09.018.