ELIT
Economic Laboratory Transition
Research Podgorica

Montenegrin Journal of Economics

Afanasyev, M.Y., Alexander V. Kudrov, A.V. *(2020), "Estimates of Economic Complexity in the Structure of the Regional Economy"*, *Montenegrin Journal of Economics*, Vol. 16, No. 4, pp. 43-53.

Estimates of Economic Complexity in the Structure of the Regional Economy

MIKHAIL Y. AFANASYEV[1] and ALEXANDER V. KUDROV[2]

[1] Professor, Central Economics and Mathematics Institute of the Russian Academy of Sciences, Moscow Russia, e-mail: miafan@cemi.rssi.ru
[2] Principal Researcher, Central Economics and Mathematics Institute of the Russian Academy of Sciences, Moscow, Russia; kovlal@inbox.ru

ARTICLE INFO

Received April 27, 2020
Revised from May 26, 2020
Accepted July 12, 2020
Available online december 15, 2020

JEL classification: C12; C51; R15

DOI: 10.14254/1800-5845/2020.16-4.4

Keywords:

Regional economy,
econometric modeling,
hypothesis testing,
economic complexity,
indicators

ABSTRACT

In this article it is considered the problem of grouping regions according to the structure of strong sectors, identifying features of formed groups and assessing their economic complexities. Minimum spanning trees are constructed to identify groups of regions with close structure of strong sectors. It was found the consistency of economic complexity of the region with the economic complexities of its neighbors, with whom it is directly related by an edge on the minimal spanning tree. The results of the regional grouping on the minimal spanning tree do not contradict the results of clustering of regions by GRP sectoral structure, which were previously obtained by the authors. Groups of regions by structure of strong sectors were typologized in accordance with the current classifier of economic sectors.

INTRODUCTION

Modern concepts of economic complexity are related to the diversification of production. Countries exporting more "complex" goods tend to have higher per capita wealth levels than countries exporting simple goods. Moreover, there can be a structural transformation of the economy and a transition from more simple forms of production to more complex ones, which is accompanied by an increase in the level of socio-economic development. Comparatively recently, it was developed a procedure that allows to measure the economic complexity of both sectors and structure of the economy as a whole (see, for example Hartmann, 2017; Hausmann et al., 2006; Hausmann and Rodrik, 2003; Hidalgo and Hausmann, 2009). Further, in contrast to the traditional approach, according to which the concept of revealed comparative advantages is applied to the produced products, see (Balassa, 1965), the emphasis is placed on the study of sectors of the regional economies. Further it is considered the problem of grouping regions according to the structure of strong sectors, identification of features for the formed groups of regions together with the assessment of their economic complexities. The obtained results can be used to solve the problems of project management, see (Makarov, 2010).

Methodology of research. To describe the structure of strong sectors of the regional economy, it is used the author's generalization of the approach to assessment of economic complexity described in (Lyubimov et al.,2017; Hausmann, Hidalgo et al.,2011) based on the use of data on tax revenues by sectors of the economy. These data take into account the production volume of each sector of the economy for export and for domestic consumption.

Precondition 1. Data on tax revenues reflect the proportions of sectors output for the economy in value terms.

The indicator RCA_{cp} of the revealed comparative advantages is determined by the formula

$$RCA_{cp} = \frac{(y_{cp}/\sum_p y_{cp})}{(\sum_c y_{cp}/\sum_c \sum_p y_{cp})}, \quad (1)$$

where y_{cp} – the volume of tax revenues from sector *p* of the regional economy *c*. Accordingly, RCA_{cp} represents the ratio of the share of tax revenues from sector *p* in total tax revenues from all sectors of the economy of the region *c* with the share of tax revenues from sector *p* in all regions in total tax revenues from all sectors of the economies of all regions. In the articles that use the indicator RCA_{cp} to assess the identified comparative advantage in the economies, see (Hausmann and Klinger,2006), it is placed a limit-filter on it from below.

If the value RCA_{cp} exceeds one[1], then, taking into account precondition 1, we can consider that the economy of region *c* has the revealed comparative advantages in the output of the sector *p*. Otherwise, it is considered that the revealed comparative advantages do not exist. Using calculated *RCA*, it is defined matrix *M*, which contains data on the sectors for each economy, which are developed at the level of revealed comparative advantages, determined with the use of expression (1). The rows of this matrix correspond to the regions and the columns correspond to the sectors of the economies. Element $m_{c,p}$ of matrix *M* is equal to 0 if the region c has no identified comparative advantage in the production of the sector p defined by expression (1), and 1 otherwise. That is:

$$m_{c,p} = \begin{cases} 1, if\ RCA_{cp} > 1 \\ 0, if\ RCA_{cp} \le 1 \end{cases} \quad (2)$$

Matrix *M* can be used to obtain characteristics of the level of diversification for the region's economy, identifying strong sectors whose products the region produces at the level of the revealed comparative advantages, as well as to calculate indices that allow for comparative analysis of economic complexities of different regions.

The approach based on the graph theory has been used to form groups regions of closely related according to the structure of strong sectors. On the basis of matrix *M* describing the structure of strong sectors of the Russian Federation regions, we will construct a hierarchy of interrelationships of structures of strong sectors in the form of the so-called "minimum spanning trees" and study its topological properties. It is worth mentioning the articles which concern the “minimal spanning trees” for correlation networks:

- prices of stocks traded in the following markets: US (Onnela,2006; Onnela, Chakraborti et al., 2002), UK (Coelho, Hutzler et al., 2007) and Japan (Jung, Kwon et al., 2008). One of the important results obtained in these works is the discovery of grouping stocks from one sector on a “minimum spanning tree branch”;

[1] Other thresholds higher than one can be used.

– magnetoencephalography data from various parts of the human brain. Revealing and studying of functional modules of a brain with use of magnetoencephalographic metrics allows to define more precisely target areas of a brain in case of necessity of surgery. As a result of application of minimal spanning trees, functional brain modules consisting of the areas performing the same functions are located on the "branches" of the tree were obtained, see (Lee, Kim et al. 2006; Stam, Tewarie et al., 2014);
– indicators in the construction of socio-economic development indicators, see (Aivazian et al., 2019).

Let us denote $G = (V, E)$ - an undirected graph of interrelationships for the regions based on the strong sectors structures each of which is described by a row of matrix *M*, where V is a set of nodes, each of which corresponds to a region from the set of regions under consideration; E is a set of edges. With the use of edges E nodes V are connected. Each edge is characterized by the strength of the relationship or the distance between the corresponding nodes.

More formally, at a fixed point of time for each region $c \in \{c_1, \dots, c_g\}$, where g - the number of considered regions, the vector-row $M_c = (m_{c,1}, \dots, m_{c,f})$ of matrix *M*, describing the structure of strong sectors of the region c (here f – total number of the possible sectors). Let's determine the distance between vectors M_{c_i} and M_{c_j}, $i, j \in \{1, \dots, g\}$ as

$$d\left(M_{c_i}, M_{c_j}\right) = \sqrt{1 - \left(\frac{2\left(M_{c_i}, M_{c_j}\right)}{\sum_p\left(m_{c_i p} + m_{c_j p}\right)}\right)^2}. \quad (3)$$

The value $\frac{2\left(M_{c_i}, M_{c_j}\right)}{\sum_p\left(m_{c_i p} + m_{c_j p}\right)}$ equals to the number of common strong sectors in the total number of strong sectors of the two regions c_i and c_j. If the structures of the strong sectors of the two regions coincide, then $M_{c_i} = M_{c_j}$ and $d\left(M_{c_i}, M_{c_j}\right) = 0$. In this case regions do not differ in the structure of strong sectors and will be classified into one group.

Definition (spanning tree). *Subgraph $G' = (V, E')$ of the graph G is called the spanning tree, if it has all the nodes V which are connected with $|V| - 1$ edges.*

It could be shown that the graph G will be connected if and only if there's a spanning tree for it. Let's assume that graph G is connected. Then there's at least one spanning tree for the graph G. Among all the spanning trees for the graph G we will be interested in some sense the minimal spanning trees:

Definition (minimal spanning tree). *The spanning tree $\tilde{G}'$ for the graph G is called minimal spanning tree if*

$$\tilde{G}' = \operatorname{argmin}_{G' \in H} \sum_{(X_i, X_j) \in E'} d(X_i, X_j)$$

where H – the set of all spanning trees for the graph G.

The use of minimal spanning trees allows to identify the groups of close nodes in the graph G, extracting from matrix *M* the most similar structures of strong sectors. As a result, the close nodes of graph G are arranged in the form of a "branch" of the minimal spanning tree. And although the transition to the minimum spanning tree is accompanied by the loss of some links, it allows to identify the "branches" characterized by a single specificity. There is a number of algorithms for constructing the minimal spanning trees for the graph G. In this study, the Kruscal algorithm is used, in more details see (Kruskal, 1956).

The concept of economic complexity of a region is considered as a characteristic reflecting the level of its technological development, which in turn is determined by strong sectors in the structure of its economy. Similarly, the economic complexity of a sector depends on the level of technological development of those regions where this sector is presented in the structure as a strong one. Let us give a more formal definition of economic complexity, which corresponds to the procedure of its calculation presented in (Hausmann et al., 2011):

a) the economic complexity of the region (ECI_c) or of the sector (ECI_p) is a latent characteristic;
b) the economic complexity of the region is proportional to the average level of economic complexity of strong sectors in the structure of its economy. Namely:
c) $ECI_c = a_1 \sum_p w_{c,p} ECI_p$, where $w_{c,p} = \frac{m_{c,p}}{k_{c,0}}$, $k_{c,0} = \sum_p m_{c,p}$,

where a_1- positive constant;
d) the economic complexity of the sector is proportional to the average level of economic complexities of the regions in whose structure this sector is strong:
e) $ECI_p = a_2 \sum_c w^*_{p,c} ECI_c$, where $w^*_{p,c} = \frac{m_{c,p}}{k_{d,0}}$, $k_{d,0} = \sum_c m_{c,p}$,

where a_2- positive constant;

Let's introduce some additional notations:
$\mathbf{c} = (ECI_{c_1}, ECI_{c_2}, \dots, ECI_g)$ - vector of economic complexities for the regions;
$\mathbf{p} = \left(ECI_{p_1}, ECI_{p_2}, \dots, ECI_{p_f}\right)$ - vector of economic complexities for the sectors;
$\mathbf{W_1} = (w_{c,p})$, $\mathbf{W_2} = (w^*_{p,c})$ – weight-matrices.

Write down properties (b) and (c) in the following matrix form: $\mathbf{c} = a_1 \mathbf{W_1 p}$, $\mathbf{p} = a_2 \mathbf{W_2 c}$. From that it follows:

$$\mathbf{c} = a_1 a_2 \mathbf{W_1 W_2 c},$$
$$\mathbf{p} = a_1 a_2 \mathbf{W_2 W_1 p}.$$

Thus, the economic complexity of a region is defined as the correspondent coordinate of the eigenvector for the matrix $\mathbf{W_1 W_2}$, and the economic complexity of a sector – correspondent coordinate of the eigenvector for the matrix $\mathbf{W_2 W_1}$. In the article (Hausmann, Rodrik,2003) it is proposed to use the standardized second principal component of the matrices $\mathbf{W_1 W_2}$ and $\mathbf{W_2 W_1}$ as the estimates of economic complexities of regions and sectors, respectively. It should be noted that the coordinates of the first principal component for these matrices consist of the same values, as they are stochastic, in more details see (Hausmann,Rodrik,2003; Kemp-Benedict,2014). We would also like to remind that if $\mathbf{x}$ is the eigenvector for matrix $\mathbf{A}$ corresponding to its eigenvalue λ, then the vector $r\mathbf{x}$, where r is any non-zero real number, is also the eigenvector of matrix $\mathbf{A}$ corresponding to its eigenvalue λ.

Now let us briefly describe the procedure of calculating economic complexity:

– Using the data matrix $M = (m_{c,p})$ let's calculate the characteristics of diversification ($k_{c,0}$) and uniqueness of the sector ($k_{p,0}$):

– $k_{c,0} = \sum_p m_{c,p}$ and $k_{p,0} = \sum_c m_{c,p}$.

– Let's calculate the matrix $\mathbf{W_1 W_2}$, whose elements characterize the degree of similarity for sets of strong sectors in the structure of economies of the corresponding pair of considered regions. The used similarity metric is a weighted sum for all common sectors where weights provide greater contribution of those sectors to which corresponds a higher level of uniqueness ($k_{p,0}$). It is easy to notice that the matrix element $\mathbf{W_1 W_2}$ at the intersection of row c and column c' is calculated by the formula:

$$\frac{1}{k_{c,0}} \sum_p \frac{m_{c,p} m_{c',p}}{k_{p,0}}, \text{ where } c, c' \text{- regions.}$$

Similarly, let's calculate the matrix $\mathbf{W_2 W_1}$, in which (p, p')-element is calculated by the formula:

$$\frac{1}{k_{p,0}}\sum_c \frac{m_{c,p}m_{c,p'}}{k_{c,0}}, \text{ where } p, p' \text{- regions.}$$

– Finally, the region's economic complexity is calculated as a standardized value of the second principal component of matrix $\mathbf{W_1 W_2}$. The direction of the second principal component is chosen so that its correlation of it coordinates with the diversification characteristic of $k_{c,0}$ is positive, see (Inoua, 2016). Namely:

$$\mathbf{c} = \begin{pmatrix} \frac{f_1 - \bar{f}}{d} \\ \frac{f_2 - \bar{f}}{d} \\ \dots \end{pmatrix},$$

where vector $\mathbf{f} = (f_1, \dots)$ – second principal component of the matrix $\mathbf{W_1 W_2}$, for which correlation with $(k_{c_1,0}, \dots, k_{c_g,0})$ is positive, $\bar{\mathrm{f}}$ – mean for the coordinates of vector $\mathbf{f}$, d – standard deviation calculated for the coordinates of vector $\mathbf{f}$.

Also, the economic complexity of sectors is calculated as a standardized value of the second principal component of the matrix $\mathbf{W_2 W_1}$, the direction of which is chosen so that its correlation with $k_{p,0}$ is positive. Namely:

$$\mathbf{p} = \begin{pmatrix} \frac{g_1 - \bar{g}}{e} \\ \frac{g_2 - \bar{g}}{e} \\ \dots \end{pmatrix},$$

where vector $\mathbf{g} = (g_1, \dots, g_f)$ – second principal component for the matrix $\mathbf{W_2 W_1}$, which is correlated positively with $(k_{p_1,0}, \dots, k_{p_f,0})$, $\bar{\mathrm{g}}$ – mean for the coordinates of vector $\mathbf{g}$, e - standard deviation calculated for the coordinates of vector $\mathbf{g}$.

Research results. For the formation of the matrix *M* constructed on the basis of the strong sectors structure was used the data on tax revenues in 82 sectors of the Russian Federation regional economies.[2] The number of strong sectors in the economy of each region is indicated in column 4 of Table P1, see Appendix section. Regions with the most diversified economies (with a large number of strong sectors): Tver Region (42), Chuvash Republic (40), Moscow Region (39), Novosibirsk Region (39), Vladimir Region (37), Lipetsk Region (36). The regions with the least diversified economy are as follows: Orenburg Region (6), Tyumen Region (8), Astrakhan Region (9), Tomsk Region (10) and Sakha Republic (11).

In Figure 1 it is presented the minimum spanning tree for the correlation graph $G = (V, E)$ constructed on the basis of matrix *M* characterizing the structure of strong sectors of regions. Each node of the tree corresponds to some region of the Russian Federation. The nodes are numbered in the order used in column 1 of Table P1, see Appendix. The presence of edge on the graph corresponds to the similarity in metric (3) of the structures of strong sectors for two correspondent regions, which are represented by nodes. On each branch of the tree there is one or several groups of regions that are very close in the structure of strong sectors. On the basis of visualization of the minimum spanning tree and expert analysis of strong sectors matrix *M*, thirteen groups of regions presented in Table P1 have been identified, see Appendix. Column 6 of this table shows the group number for the considered regions that are close in metric (3) in terms of strong sectors structure. The identified groups have the following properties: each group corresponds to the structure of the minimum spanning tree; for each group it was identified one or more sectors, which are strong only in the structure of all regions in this group. Codes[3] and names of these sectors are given in column 7 of Table P1, see Appendix. These sectors are a distinctive feature of the corresponding group of regions. Column 7 of Table P1 also contains codes of strong sec-

[2] Data on tax receipts https://www.nalog.ru/rn77/related_activities/statistics_and_analytics/forms/8826515/
[3] Specified according to the classifier of economic sectors
https://www.nalog.ru/rn77/related_activities/statistics_and_analytics/forms/8826515/

tors for all regions of the group, which at the same time could be strong sectors in some other groups. And also codes (in italics) of sectors that are strong for all regions of the group, except for one. This information characterizes the structure of the regional economies included in each group in a fairly complete way.

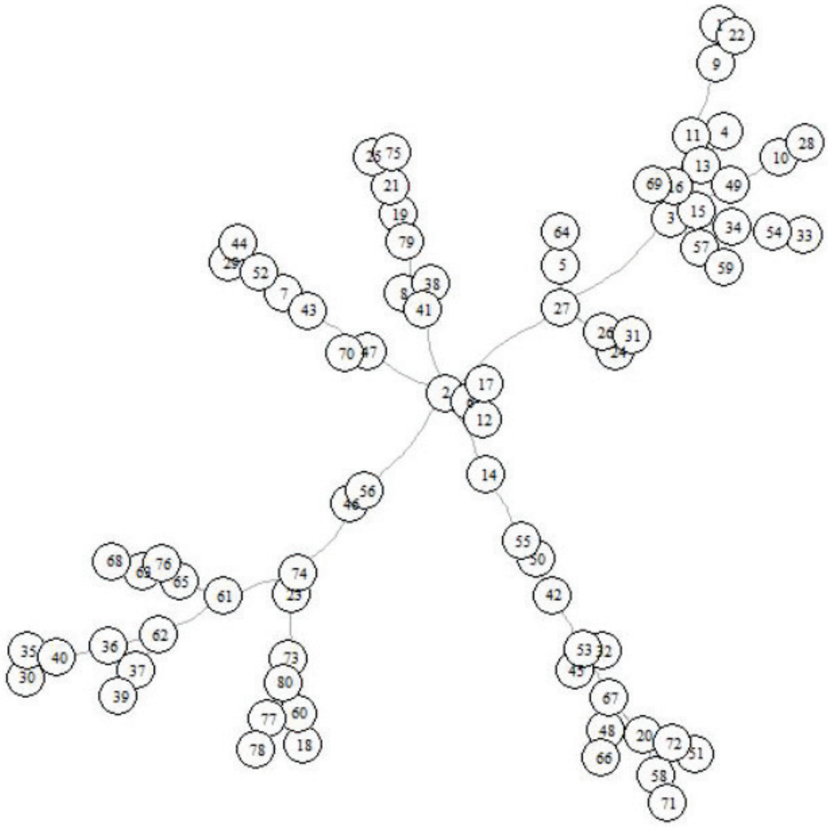

Figure 1. Minimal spanning tree for grouping regions by structure of the strong sectors

The index of economic complexity is presented in column 5 of Table P1, see Appendix. For each group of regions which are close by the structure of strong sectors, estimate of economic complexity are obtained as an average value of indices of regions included in the group. At the same time, for stability of group estimations, each group did not take into account the highest and lowest values. In Table P1 (see Appendix), groups of regions with close in structure of strong sectors regions are ordered in descending order of group economic complexity estimates.

Network effects model for the index of economic complexity. Let's denote through $P = \left(p_{i,j} | i, j \in \{c_1, \ldots, c_g\}\right)$ the adjacency matrix defining the minimum spanning tree $\tilde{G}' = (\tilde{V}', \tilde{E}')$. Specifically:

$$p_{i,j} = \begin{cases} 1, \text{if there is an edge between nodes } i \text{ and } j \text{ of graph } G \\ 0, \text{if there isn't an edge between nodes } i \text{ and } j \text{ of graph } G \end{cases}$$

where $i, j \in \{c_1, \ldots, c_g\}$.

We will also put that $p_{i,i} = 0$ for all i. Thus, all diagonal elements of matrix P are equal to zero. Matrix P represents the structure of graph $\tilde{G}' = (\tilde{V}', \tilde{E}')$.

In order to check if there is consistency in economic complexities of neighboring nodes of the minimum spanning tree according to the structure of strong sectors, $\tilde{G}' = (\tilde{V}', \tilde{E}')$, it is proposed to use a linear network autocorrelation model. This model is defined by the following equation:

$$\mathbf{c} = \alpha\mathbf{I} + \beta QP\mathbf{c} + \boldsymbol{\epsilon},\ \boldsymbol{\epsilon} \sim N(\mathbf{0}, \sigma^2\mathbf{I}), \qquad (4)$$

where α, β –constants; $\mathbf{I}$ –unit matrix, dimension of which coincides with the size of vector $\mathbf{c}$, which equals to the number of regions under consideration;

$Q = diag\left(\frac{1}{\sum_{i\in\{c_1,\dots,c_g\}} p_{c_1,c_i}}, \dots, \frac{1}{\sum_{i\in\{c_1,\dots,c_g\}} p_{c_g,c_i}}\right)$ – diagonal matrix.

Using component $QP\mathbf{c}$ in the right part of the equation allows to reflect the consistency of economic complexity of the region with the economic complexity of its neighbors, with whom it is directly related by an edge on the graph $\tilde{G}' = (\tilde{V}', \tilde{E}')$. The parameter β is called the network effect parameter, which reflects the strength of network interaction. The absence of a network effect corresponds to the fact that $\beta = 0$.

Table 1. *Linear network autocorrelation model for graph* $\tilde{G}' = (\tilde{V}', \tilde{E}')$.

	Estimate	*St.dev.*	*t-statistics*	*p-value*
α	0,09	0,05	1,74	0,09
β	1,00	0,05	18,26	0,00

R-squared	0,81
Adj. R-squared	0,81

As it can be seen from the results of parameter estimates presented in Table 1, the model (4) is characterized by a rather high value of R-squared which equals to 0.81, and β-coefficient is statistically significant with a rather high value *t*-statistic which equals to 18.26. The value of statistical estimate of β coefficient equals almost one - it means that the value of economic complexity of a region is approximately equals to the average of economic complexities of those regions that are directly connected by edge on the minimal spanning tree $\tilde{G}' = (\tilde{V}', \tilde{E}')$.

Discussion of the results. In Figure 2 the regions are represented in the space "number of strong sectors (x-axis) - economic complexity (y- axis)". Increase in the number of strong sectors from 6 (minimum value) to 14 is accompanied by growth of economic complexity. For the number of strong sectors from 15 to 30, estimates of economic complexity approach the maximum value, but for some regions relatively low economic complexities are observed that do not correspond to the number of their strong sectors. In case when the number of strong sectors is greater than 30, estimates of economic complexity are consistently high. There is an effect of diversification of economy. Belgorod region (24 strong sectors), Kursk region (22 strong sectors) and Vladimir region (37 strong sectors) have the highest estimates of economic complexity. Relatively low estimates of economic complexity are observed for the following regions: Orenburg Region (6 strong sectors, estimate -4.569), Tyumen Region (8, -2.833), Sakha Republic (11, -2.710), Tomsk Region (10, -2.509), Komi Republic (14, -2.469).

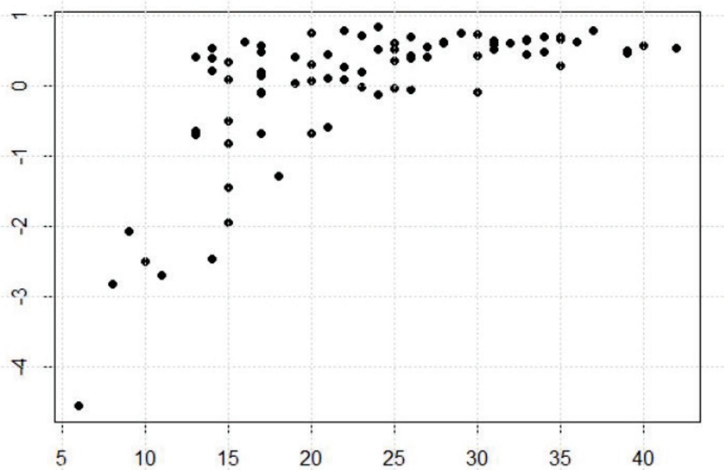

Figure 2. *On the x- axis it is the number of complex sectors of the region, on the y- axis - estimation of economic complexity*

In describing the specifics of groups of regions which are close in metrics (3) calculated on data for the structure of strong sectors, we will compare them with clusters of regions presented in (Aivazian et al. 2016). Based on the author's approach to data on GRP sectoral structure, regions were divided into five clusters: 1) “evenly developed”, 2) “extractive”, 3) “manufacturing”, 4) “agricultural” and 5) “developing”. Cluster number is given for each region in column 3 of Table P1, see Appendix. Relationship between the composition of groups of regions, which are close by the structure of strong sectors and the composition of clusters can be traced by comparative analysis of the groups and cluster numbers, shown in columns 6 and 3 of Table P1, respectively. All 12 “manufacturing” regions were included in groups 1-7 with the highest estimates of economic complexity. All 11 'extractive' regions were included in groups 9-13 with the lowest estimates of economic complexity. The majority (6 out of 8) of "developing" regions were included in Group 8 with higher economic complexity than "extracting" regions but lower than "manufacturing" regions. Estimates of economic complexity of "agricultural" regions are comparable to those of "manufacturing" regions and are superior to those of "developing" and "extractive" regions. Thus "evenly developed" regions are distributed on all thirteen groups.

CONCLUSION

There is a tendency to increase estimates of economic complexity of regions with an increasing number of strong sectors. It was found the consistency of economic complexity of the region with the economic complexities of its neighbors, with whom it is directly related by an edge on the minimal spanning tree.

Results of regions clustering by structure of strong sectors do not contradict results of regions clustering by GRP sectoral structure obtained earlier by the authors. Moreover, "manufacturing" and most "agricultural" regions are included into groups with the highest group estimates of economic complexity. Estimates of economic complexity of most “developing” regions are lower than those of “processing” and “agricultural” regions, but higher than those of “extracting” regions. Grouping of regions according to the structure of strong sectors allows to specify peculiarities of the regional economies previously included in the cluster of "evenly developed" regions.

The results obtained concretize the role of the economic complexity concept in production theory. The applied value of the results is determined by the possibilities of their use in solving project management tasks.

REFERENCES

Aivazian, S.A., Afanasiev, M.Y., Kudrov, A.V. (2016), “Models of production potential and evaluation technological efficiency for the Russian Federation regions with account of production structure”, *Economics and Mathematical Methods*, Vol. 52, No. 1, pp. 28–44 (in Russian).

Aivazian, S.A., Afanasiev, M.Y., Kudrov, A.V. (2019), “Integral indicator for quality of life conditions”, *Digital-Economy*, Vol. 1, No. 5, pp. 43-56 (in Russian).
Lyubimov, I.L., Gvozdeva, M.A., Kazakova, M.V., Nesterova, K.V. (2017), “Economic Complexity of Russian Regions and their Potential to Diversify”, *NEA Journal*, No. 2 (34), с. 94–122 (in Russian).
Makarov, V.L. (2010), *Social Clusterism. Russian challenge*, Biznes Atlas, Moskow (in Russian).
Balassa, B. (1965), “Lafayrade Liberalization and “Revealed” Comparative Advantage”, *The Manchester School*, Vol. 33, pp. 99—123.
Coelho, R., Hutzler, S., Repetowicz, P., Richmond, P. (2007), “Sector analysis for a FTSE portfolio of stocks”, *Physica A*, Vol. 373, pp. 615–626.
Hartmann, D. (2017), “Linking Economic Complexity, Institutions, and Income Inequality”, *World Development*, Vol. 93, pp. 75–93.
Hausmann, R. et al. (2011), *The Atlas of Economic Complexity: Mapping Pathsto Prosperity. Cambridge: Center for International Development*, Harvard University, MIT.
Hausmann, R., Hwang, J., Rodrik, D. (2006), „What you export matters”, *Journal of Economic Growth*, Vol. 12, No. 1, pp. 1–25.
Hausmann, R., Klinger, B. (2006), “Structural Transformation and Patterns of Comparative Advantage in the Product Space”, *CID Working Paper,* No. 128.
Hausmann, R., Rodrik, D. (2003), “Economic development as selfdiscovery”, *Journal of Development Economics*, Vol. 72, No. 2, pp. 603–633.
Hidalgo, C.A., Hausmann, R. (2009), “The building blocks of economic complexity”, *Proceedings of the National Academy of Sciences*, Vol. 106, No. 26, pp. 10570–10575.
Jung, W.-S. et al. (2008), “Group dynamics of the Japanese market”, *Physica A*, Vol. 387, pp. 537–542.
Kemp-Benedict, E. (2014), “An interpretation and critique of the method of reflections”, *MPRA Paper*, No. 60705, https://mpra.ub.uni-muenchen.de/60705/1/MPRA_paper_60705.pdf.

Kruskal, J.B. (1956), “On the Shortest Spanning Subtree of a Graph and the Traveling Salesman Problem”, *Proceedings of the American Mathematical Society*, Vol. 7, No. 1, pp. 48-50.
Lee, U., Kim, S., Jung, K.Y. (2006), “Classification of epilepsy types through global network analysis of scalp electroencephalograms”, *Physical Review E,* Issue 73, Vol. 4, pp. 19-20.
Onnela, J.-P.(2006), *Complex Networks in the Study of Financial and Social Systems*, Helsinki University of Technology, Helsinki.
Onnela, J.-P., Chakraborti, A., Kaski, K., Kertesz, J. (2002), “Dynamic asset trees and portfolio analysis”, *The European Physical Journal B - Condensed Matter and Complex Systems*, No. 30, pp. 285-288.
Stam, C.J., Tewarie, P., Van Dellen, E., van Straaten, E.C., Hillebrand, A., Van Mieqhem, P. (2014), 2The trees and the forest: characterization of complex brain networks with minimum spanning trees”, *International Journal of Psychophysiology,* Vol. 92, pp. 129–138.
Inoua, S. (2016), *A Simple Measure of Economic Complexity*, https://arxiv.org/ftp/arxiv/papers/1601/1601.05012.pdf

APPENDIX

Table 1 Results of grouping and estimation of economic complexity.
Name of columns:

(1) – № of the region in the order used by Rosstat,
(2) – the name of the region,
(3) – the number of region's cluster by GRP structure,
(4) – the number of strong sectors in the region,
(5) – the economic complexity index,
(6) – the number of the region’s group by strong sector structure,
(7) – codes and names of common strong sectors of all regions of this group only, codes of sectors of all regions of the group, codes (in italics) of common sectors of all but one regions of the group.

(1)	(2)	(3)	(4)	(5)	(6)	(7)
12	Ryazan region	1	16	0,614	1	1125 leather production; 1095, 1203, 1250; *1100, 1110, 1255, 1261, 1400.*
38	Karachay-Cherkessia Republic	1	27	0,407		
2	Bryansk region	1	31	0,591		
8	Kursk region	1	22	0,785		
17	Yaroslavl region	3	25	0,514		
6	Kaluga region	3	29	0,751		
41	Stavropol Krai	4	23	0,707		
1	Belgorod region	1	24	0,830	2	1170 metal fabrication; 1175 iron, steel and ferroalloy production; 1135 paper production; 1020, 1100, 1110, 1165, 1250; *1090. 1255, 1261, 1420.*
22	Vologda region	3	25	0,596		
9	Lipetsk region	3	36	0,628		
26	Novgorod region	3	32	0,595	3	1090 foodstuffs production; 1020, 1315; *1025, 1100, 1250, 1261, 1263, 1280, 1285, 1290.*
24	Leningrad region	1	14	0,526		
5	Ivanovo region	1	28	0,617		
31	Krasnodar Region	4	27	0,552		
64	Altai Krai	4	33	0,653		
27	Pskov region	4	35	0,687		
33	Volgograd region	1	17	0,475	4	1160 rubber and plastics manufacturing; 1200 metal casting; 1020, 1165, 1261, 1305; *1400, 1410, 1420.*
13	Smolensk region	1	31	0,639		
11	Oryol region	1	30	0,735		
54	Saratov region	1	21	0,101		
16	Tula region	3	34	0,690		
59	Chelyabinsk region	3	35	0,284		
57	Sverdlovsk region	3	30	0,429		
3	Vladimir region	3	37	0,783		
4	Voronezh region	4	34	0,483		
34	Rostov region	4	33	0,643		
28	Saint Petersburg	1	23	0,197	5	1120 textile manufacturing; 1140 printing and media copying; 1110, 1203, 1220, 1430;
10	Moscow region	1	39	0,464		
69	Novosibirsk region	1	39	0,492		
15	Tver region	1	42	0,531		
49	Kirov region	1	35	0,654		
44	Republic of Mordovia	1	20	0,754	6	1100 dairy production; 1020, 1130; *1400, 1410, 1420, 1430, 1440.*
43	Mari El Republic	1	31	0,522		
47	Chuvash Republic	1	40	0,574		
7	Kostroma region	1	33	0,446		
70	Omsk region	3	17	0,131		
52	Penza region	4	26	0,693		
29	Republic of Adygeya	4	22	0,077		
55	Ulyanovsk region	1	25	0,359	7	1205 machinery and equipment production; 1220; *1165.*
42	Republic of Bashkortostan	3	17	-0,689		

50	Nizhny Novgorod region	3	24	0,519		
14	Tambov region	4	28	0,610		
37	Kabardino-Balkarian Republic	1	17	0,574	8	1345 mailing and courier activities; 1261, 1400, 1410, 1420; *1250, 1430, 1440.*
62	Republic of Tyva	5	17	0,187		
35	Dagestan Republic	5	19	0,413		
39	Republic of North Ossetia-Alania	5	14	0,202		
30	Republic of Kalmykia	5	14	0,387		
36	Republic of Ingushetia	5	15	0,082		
40	Chechen Republic	5	13	0,409		
76	Amur region	1	17	-0,092	9	1046 coal and anthracite mining and processing; 1047 brown coal mining and processing; 1245 electricity generation, transmission and distribution; *1090, 1255, 1270, 1290, 1305, 1315, 1400, 1410.*
65	Transbaikal region	1	19	0,027		
63	Republic of Khakassia	1	22	0,266		
68	Kemerovo region	2	20	0,294		
25	Murmansk region	1	17	-0,110	10	1030 fishing, fish farming; 1081 mining (except coal, oil, gas, iron ores and nonferrous metal ores); 1305, 1315, 1340; *1255, 1261, 1400, 1410, 1420.*
75	Khabarovsk Territory	1	21	-0,586		
19	Republic of Karelia	1	26	0,424		
21	Arkhangelsk region	2	20	0,058		
79	Jewish Autonomous Region	5	21	0,446		
74	Primorsky Krai	1	26	-0,050	11	1215 motor vehicle and trailer manufacturing; 1025; *1090, 1095, 1130, 1261, 1290.*
23	Kaliningrad region	1	15	0,342		
46	Udmurt Republic	2	15	-0,496		
56	Kurgan region	4	26	0,385		
61	Buryatia Republic	4	25	-0,042		
18	Moscow	1	24	-0,122	12	1398 administrative activities and related services; 1330 air and space activities; 1270 construction site; 1340; *1284, 1400.*
73	Kamchatka region	1	23	0,194		
77	Magadan region	1	23	-0,023		
80	Chukotka Autonomous Okrug	2	13	-0,645		
78	Sakhalin region	2	18	-1,298		
60	Altai Republic	5	30	-0,092		
48	Perm Krai	1	20	-0,678	13	1055 crude oil and natural gas production; 1084 providing services in the field of oil and natural gas production.
32	Astrakhan region	1	9	-2,076		
51	Orenburg region	2	6	-4,569		
67	Irkutsk region	1	15	-1,456		
66	Krasnoyarsk Territory	1	15	-1,955		
53	Samara region	1	15	-0,818		
20	Komi Republic	2	14	-2,469		
58	Tyumen region	2	8	-2,833		
71	Tomsk region	2	10	-2,509		
72	Sakha Republic (Yakutia)	2	11	-2,710		
45	Republic of Tatarstan	2	13	-0,707		