On diversification taking into account innovation activity and resource availability of regions

M.Yu. Afanasiev*, *A.V.* Kudrov, and *M.A.* Lysenkova

Central Economics and Mathematics Institute of the Russian Academy of Sciences, Moscow, Russia

Abstract. An approach to assessing the possibility of diversification of the regional economy based on the formation of recommendations for the development of sectors to the level of strong ones is presented. The approach is based on regression analysis using an extended economic basis, including characteristics of regional differentiation and innovation activity. On this basis, sectors can be identified, the development of which depends on the innovation activity of the regions. A condition is described that makes it possible to assess the sufficiency of the resource provision of the sector in the region for its development to the level of a strong one. As a result of the approbation of the proposed approach, 20 sectors have been identified, the volume of production of which depends on the innovative activity of the regions aimed at creating international patent applications. For the Construction sector, considered as an example, 11 regions have been identified in which this sector has sufficient resource provision for development to the level of a strong one. The possibility of turning this sector into a strong one in other regions may be associated with an increase in their innovation activity and government support.

1. Introduction

Currently, one of the priorities of state policy is the transition of the economy to an innovative path of development. The researchers note that today the economy should not only provide a platform for the implementation of entrepreneurial initiatives, but also demonstrate a receptivity to innovation and encourage innovative initiatives, research and development. As well as the economic potential of the region, the potential of innovation activity is a factor in the diversification of the regional economy. The theory of diversification and empirical estimates are presented in [1-3]. According to this theory, companies benefit from facing a heterogeneous environment consisting of different industries, as new ideas come from the external environment. In this paper, the process of diversification of the regional economy is associated with the emergence of new developed sectors in it.

* Corresponding author: mi.afan@yandexi.ru

2 Research methodology

Assessment of the impact of innovation activity in the region on the development of the economic sector. To assess the impact of the region's innovation activity on the development of the economic sector, the components of the economic basis are used, including characteristics of regional differentiation and indices of innovation activity. The description of the economic basis $\{L, te, s^1, s^2\}$ and the methodology of its application for assessing socio–economic development at the regional level are presented in [4, 5]. The description of the innovation activity indices used below, based on the stochastic boundary concept, is given in [6]. An expanded economic basis $\{L, te, s^1, s^2, INN\}$ is being formed, including the *INN* innovation activity index. This basis reflects not only the economic structure of the regional economy, but also the specifics of the innovation activity of the regions, focused on a specific result of innovation activity. If the index of innovation activity statistically depends on some components of the economic basis, then to prevent the effect of multicollinearity, it is advisable to use a modification of the index, cleared of the influence of these components. A regression analysis of the production volumes of each sector of the economy is carried out using an expanded economic basis. Let's build regressions of the form

$$\ln y_{ij} = const_i + \beta 1_i lnL_j + \beta 2_i te_j + \beta 3_i s^1{}_j + \beta 4_i s^2{}_j + \beta 5_i INN_j + \varepsilon_{i,j} \quad (1)$$

Here y_{ij} — the volume of production of sector *i* in region *j*; L_j— the scale of the economy of region *j* (the Rosstat indicator "number of economically active population" is used as a characteristic of the scale of the economy); te_j— assessment of the technical efficiency of regional production [7]; $s^1{}_j$— the index of industry specialization (the first main component of the GRP structure); $s^2{}_j$— the index of industrialization (the second main component of the GRP structure). When constructing the main components, the author's methodology and Rosstat indicators for the industry structure of GRP were used [8]. *INN* is an index of innovation activity (one of the author's indexes is used here, built on the basis of the stochastic boundary concept based on data on international patent applications (TEMPZ), patent applications (TEPZ), granted patents (TEVP), newly developed production technologies (TETTCH) [6]. $\varepsilon_{i,j}$— regression error. From the set of sectors, those for which the parameter score is $\beta 5_i$positive and significant at 95% level are distinguished. The volume of production of each of these sectors depends on the level of innovation activity of the regions, determined by the *INN* index.

The structure of strong sectors. To describe the structure of strong sectors of the economy, regional data on production volumes in a fairly wide range of sectors are used. To begin with, we will determine the indicator RCA_{cp} of the identified comparative advantages [9]:

$$RCA_{cp} = \frac{\left(y_{cp} / \sum_p y_{cp}\right)}{\left(\sum_c y_{cp} / \sum_{c,p} y_{cp}\right)} \quad (2)$$

where y_{cp} is the volume of production of sector p of the economy of the region c.

The indicator RCA_{cp} is the ratio of the share of production from sector p in the total volume of production from all sectors of the economy of region c to the share of production of sector p for all regions in the volume of production from all sectors of the economy of all regions. In accordance with [10], to identify comparative advantages in economies, an indicator RCA_{cp} is used for which a condition of the type of restriction from below is checked. Namely, if the value RCA_{cp} exceeds one, it is assumed that the economy of region

c has identified comparative advantages in the output of sector p. Otherwise, it is assumed that the identified comparative advantages do not exist. More formally:

$$\left(a_{c,p}\right)=\begin{cases} 1, \text{ if } RCA_{cp} \geq 1 \\ 0, \text{ if } RCA_{cp} < 1 \end{cases}$$

The matrix $\left(a_{c,p}\right)$ contains data on the sectors of the economy that are developed in different regions at the level of the identified comparative advantages determined using the expression (2). The rows of this matrix correspond to regions, the columns correspond to sectors of the economy. Next , we will call the vector $\left(a_{c,p_1},\ldots,a_{c,pm}\right)$ a structure of strong sectors of the region 's economy c . Note that for any c there exists a p for which $RCA_{cp}>1$

Resource availability of the sector in the region. The assessment of the level of resource availability of the sector in the region is determined by the level of compliance of the actual volume of production of the sector with the expected, due to the characteristics of differentiation of the region. The regions in which the expected output of the sector is higher than the actual one are identified. In such regions, the transformation of the sector into a strong one is possible due to the unfulfilled potential of economic development. This applies to all sectors. Including those whose production volume does not depend on the level of innovation activity of the region.

If the actual output of the sector is higher than expected, then the sector, having already realized the growth potential in the region, still has not become strong. In this case, the development of the sector to a strong level can be based on the growth of innovation activity in the region. This applies to sectors whose production volume depends on the level of innovation activity of the region. Different metrics can be used to compare the actual output of the sector with the expected output in a particular region and to assess resource availability. For example, the resource availability of the sector $p_j$in the region $c_i$can be assessed based on the concept of identified comparative advantages. With sufficient resources, the indicator $E\_RCA_{c_ip_j}$ of the identified comparative advantages corresponding to the expected output of the sector $p_j$in the region $c_i$should be at least 1 in order for the sector to become strong. This means that the inequality must be satisfied

$$(E\_y_{c_ip_j} \ /((E\_y_{c_ip_j} \ + \sum_{p\neq p_j} y_{cp}))/(\sum_c y_{cp}/\sum_{c;} y_{cp)} \geq 1 \quad (3)$$

where $E\_y_{c_ip_j} \ = exp\left\{lny_{c_ip_j}-\varepsilon_{i,j}\right\}$

Inequality (3) is equivalent to inequality

$$\varepsilon_{i,j} \leq \ln\left(\frac{(1-u_{p_j})RCA_{c_ip_j}}{(1-u_{p_j})RCA_{c_ip_j}}\right. \quad (4)$$

where $u_{p_j}=\frac{\sum_c y_{cp_j}}{\sum_{cp} y_{cp}}$, and $RCA_{c_ip_j}$is determined by the formula (2).

Note that the right side of inequality (4) is a negative value. This follows from the inequality $RCA_{c_ip_j}<1$which is true since the sectorp_j is not strong in the region c_i. Let's denote it $\varepsilon *_{i,j}$. Thus, if the regression error $\varepsilon_{i,j}(1)$ is less than a negative value$\varepsilon *_{i,j}$, then the sector $p_j$has sufficient resource provision in the region $c_i$in the sense that with the expected volume of production it will become strong. Otherwise, we believe that the resource provision of the region $c_i$is not enough to turn the sector $p_j$into a strong one.

The choice of a priority direction for the diversification of the region's economy is associated with the choice of a sector for its development in the region to the level of a strong one. The rationale for the choice may be the solution of a multi-criteria optimization problem taking into account a number of characteristics for each sector from a set of sectors that are

not strong in the region. Including assessments of the impact of the region's innovation activity on the development of the sector and the resource availability of the sector in the region.

3 Research results

Table 1 shows the correlation matrix of the components of the economic basis and the indices of innovation activity. The correlation analysis of the four components of the economic basis and the four indices of innovation activity shows: all components of the economic basis can be considered mutually independent; the indices of innovation activity can be considered mutually independent (with the exception of the TEPZ and TEVP indices, the dependence of which is due to their specifics); each index of innovation activity is independent or weakly dependent on the economic basis.

Table 1. Correlation matrix of the components of the basis and indices of innovation activity according to 2019 data.

	L	te	s^1	s^2	TEMPZ	TEPZ	TEPV	TETTCH
L	1	0.17	-0.13	0.19	0.23	0.207	0.12	0.14
te	0.17	1	0.20	0.23	0.12	-0.20	-0.19	0.21
s^1	-0.13	0.20	1	-1E-9	-0.22	-0.30	-0.35	-0.05
s^2	0.19	0.23	-1E-9	1	0.39	0.05	0.14	0.16
TEMPZ	0.23	0.12	-0.22	0.39	1	0.04	0.06	0.40
TEPZ	0.21	-0.20	-0.31	0.05	0.04	1	0.87	0.15
TEVP	0.12	-0.19	-0.35	0.14	0.06	0.87	1	0.13
TETTCH	0.14	0.21	-0.05	0.16	0.40	0.15	0.13	1

Table 2 presents the results of a regression analysis of production volumes by sector on the characteristics of the economic basis, expanded, as an example, due to the TEMPZ innovation activity index, based on data on international patent applications. As estimates of production volumes, data on tax revenues by economic sectors can be used (Data on tax revenues by economic sectors https://www.nalog.ru/rn77/related_activities/statistics_and_analytics/forms/8826515 /), which makes it possible to characterize the structures of regional economies, including sectors focused on both external and internal markets.

Table 2. Sectors whose development depends on the innovation activity of the regions according to 2019 data.

Sector name	***const***	$\boldsymbol{\beta 1_j}$	$\boldsymbol{\beta 2_j}$	$\boldsymbol{\beta 3_j}$	$\boldsymbol{\beta 4_j}$	$\boldsymbol{\beta 5_j}$
(1)	(2)	(3)	(4)	(5)	(6)	(7)
Support activities for petroleum and natural gas extraction	-6.31 (-1.60)	2.28 (4.;23)	0.06 (0.14)	2.43 (5.42)	-0.65 (-1.30)	1.16 (2.20)
Manufacture of other non-metallic mineral products	4.87 (6.08)	1.18 (10.54)	0.02 (0.20)	-0.18 (-1.87)	0.43 (3.90)	0.22 (2.20)
Other productions	4.54 (5.13)	1.22 (9.86)	-0.12 (-1.18)	0.12 (1.09)	0.43 (3.51)	0.28 (2.37)

Continuation of table 2.

Manufacture of gas, distribution of gaseous fuels through mains	7.70 (9.74)	0.75 (6.89)	0.09 (0.95)	0.14 (1.40)	0.28 (2.85)	0.22 (2.25)
Waste collection, treatment and disposal activities, materials recovery	7.94 (17.12)	1.11 (17.15)	0.11 (1.95)	0.16 (2.88)	0.10 (1.58)	0.23 (3.84)
Construction	7.21 (12.77)	1.13 (14.42)	0.09 (1.32)	0.22 (3.23)	0.15 (1.99)	0.27 (3.76)
Wholesale and retail trade, repair of motor vehicles and motorcycles	7.25 (16.59)	1.05 (17.28)	0.16 (3.16)	0.04 (0.84)	0.02 (0.36)	0.15 (2.80)
Railway transport activities	1.23 (0.50)	1.24 (3.55)	0.30 (1.10)	1.29 (4.86)	-0.23 (-0.72)	0.59 (2.12)
Transport via pipeline	5.29 (5.56)	1.18 (8.91)	0.29 (2.53)	0.40 (3.34)	-0.004 (-0.03)	0.26 (2.11)
Air and space transport	4.89 (8.75)	1.09 (14.01)	0.11 (1.67)	0.17 (2.41)	0.16 (2.18)	0.15 (2.10)
Postal and courier activities	5.87 (9.21)	1.12 (12.68)	0.14 (1.81)	0.14 (1.78)	0.08 (0.90)	0.23 (2.82)
Hotels and similar accommodation	5.59 (11.64)	1.06 (15.83)	0.12 (2.13)	0.16 (2.74)	0.02 (0.30)	0.14 (2.29)
Telecommunications	-2.06 (-2.40)	1.78 (14.89)	0.11 (1.08)	0.12 (1.19)	0.22 (1.90)	0.28 (2.59)
Financial and insurance activities	4.55 (7.75)	1.36 (16.62)	0.08 (1.19)	0.03 (0.50)	0.19 (2.42)	0.17 (2.30)
Financial service activities, except insurance and pension funding	4.46 (6.80)	1.41 (15.46)	0.07 (0.94)	0.43 (5.49)	0.18 (2.21)	0.32 (3.94)
Real estate activities	7.56 (31.71)	0.98 (29.70)	0.08 (2.90)	0.24 (8.11)	-0.004 (-0.10)	0.09 (3.21)
Professional, scientific and technical activities	7.80 (24.76)	0.92 (20.96)	0.05 (1.44)	0.20 (5.22)	0.04 (1.10)	0.07 (1.90)
Administrative and support service activities	5.67 (12.97)	1.02 (16.73)	0.16 (3.05)	0.17 (3.17)	0.06 (1.13)	0.11 (2.09)
Education	4.27 (10.31)	1.09 (18.92)	0.06 (1.35)	0.16 (3.24)	0.12 (2.19)	0.17 (3.17)
Human health and social work activities	2.53 (4.96)	1.13 (15.95)	0.08 (1.40)	0.15 (2.36)	0.07 (1.03)	0.18 (2.75)

Column (1) of Table 2 shows the names of sectors whose production volumes in the regions depend on the value of the TEMPZ innovation activity index. Column (2) shows the estimate of the constant in the regression, in parentheses t-statistics. Column (3) shows an estimate of the regression coefficient for the logarithm of the economically active population and t-statistics. Column (4) shows an estimate of the regression coefficient for the index of technical efficiency of regional production and t-statistics. Column (5) shows an estimate of the regression coefficient for the first main component of the GRP structure and t-statistics. Column (6) shows an estimate of the regression coefficient for the second main component of the GRP structure and t-statistics. Column (7) shows an estimate of the regression coefficient for the TEMPZ innovation activity index and t-statistics. The coefficient of determination for each of the 20 sectors of the economy is quite high. This means that the

basis used for the characteristics of regional differentiation, expanded by the index$TEMPZ_i$, explains quite well the specifics of the production volumes of the sectors.

At this stage of the study, 20 sectors of the economy have been identified, the development of which depends on the innovative activity of the region when creating international patent applications. Regions, forming international patent applications and demonstrating activity in this area, influence the development of each of these 20 sectors. It follows from the simulation results that the economic development potential of each of these sectors is associated with the growth of the scale of the regional economy, specialization or industrialization of the region, and increased technical efficiency. Depending on which components of the economic basis there are significant estimates of the coefficients in the regression (1). Another way is connected with the realization of the potential of innovation activity. If we replace the TEMPZ index with another index of innovation activity, we will get a list of sectors whose output volumes depend on the innovation activity of the region when creating the corresponding result of innovation activity.

Column (2) of Table 3 shows the number of strong sectors in the structure of the region's economy, that is, an assessment of economic diversification. The most diversified (with more than 35 strong sectors) economies of the regions are: Tver Region – 42; Chuvash Republic - 40; Moscow Region – 39; Novosibirsk Region - 39; Vladimir Region - 37; Lipetsk region – 36. The least diversified (with the number of strong sectors less than 10) economies of the regions: Astrakhan region – 9; Tyumen region – 8; Orenburg region – 6.

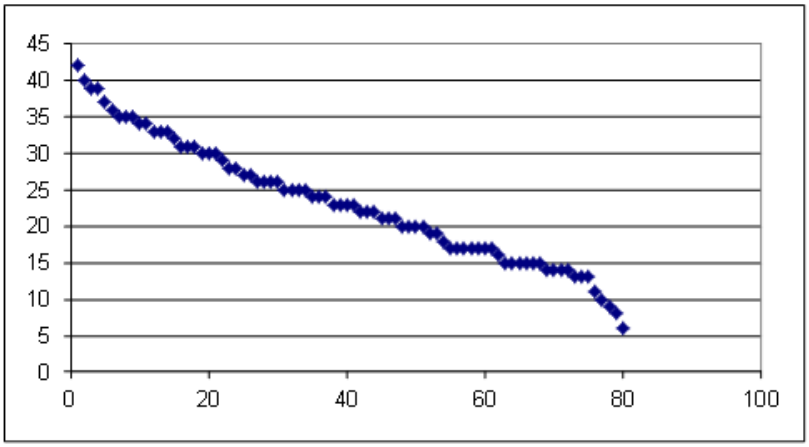

Fig. 1. Distribution of the number of strong sectors by region.

Figure 1 shows the distribution of strong sectors by region. On the abscissa axis – the number of the region, on the ordinate axis – the number of strong sectors, ordered in descending order. As an example, column (5) of Table 3 provides estimates of the identified comparative advantages of the Construction sector for various regions. It is indicated in which regions the Construction sector has identified comparative advantages and is strong (value 1), and in which it is not (value 0). The estimates obtained indicate that the Construction sector, according to 2019 data, is strong in the economy of 33 regions. In these regions, the growth of the output of this sector will no longer lead to diversification of the structure of strong sectors of the economies. The Construction sector is not strong for 47 regions. For these regions, economic diversification is possible due to the growth of the output of this sector and its transformation into a strong sector. According to the data of 2019, the Construction sector is one of the 20 sectors whose production volume depends on the innovation activity of the region.

Let's consider the estimates of regression errors (1) of the volume of production of the Construction sector in order to verify the fulfillment of the condition of resource availability of the sector in regions where the sector is not strong. Column (3) of Table 3 shows the values of the regression error$\varepsilon_{i,j}$ (1). Column (4) shows the threshold value $\varepsilon *_{i,j}$of the right side of the inequality (4) for checking the fulfillment of the condition of resource availability of the sector in the region. Note that for all regions in which the sector is not strong, the threshold value is a negative value. If the regression error value in column (3) for the region is less than the threshold value in column (4), then the Construction sector has sufficient resource availability in the region in the sense that it will become strong with the expected volume of production. In this case, the development of the sector in the region to the level of a strong one can rely on the economic potential of the region. According to the data of 2019, 11 such regions were revealed. They are marked with a sign (*) in the first column of Table 3.

Table 3. Construction sector: estimates according to 2019 data.

Name of the region	**Number of strong sectors**	**Regression error**	**Threshold value**	**1 is a strong sector, otherwise 0**
(1)	(2)	(3)	(4)	(5)
Belgorod region*	24	-0.75	-0.152	0
Bryansk region	31	0.304	0.443	1
Vladimir region	37	0.043	-0.316	0
Voronezh region	34	0.279	-0.040	0
Ivanovo region	28	0.412	-0.071	0
Kaluga region	29	-0.177	0.258	1
Kostroma region	33	-0.323	-0.354	0
Kursk region*	22	-1.087	-0.013	0
Lipetsk region	36	-0.044	0.224	1
Moscow region	39	1.265	0.028	1
Oryol region	30	0.134	-0.009	0
Ryazan region	16	0.347	-0.772	0
Smolensk region	31	1.523	-0.245	0
Tambov region	28	-0.525	0.708	1
Tver region	42	0.09	0.084	1
Tula region	34	0.109	0.663	1
Yaroslavl region	25	0.493	-0.612	0
Moscow	24	1.83	0.253	1
Republic of Karelia	26	-0.302	0.209	1
Komi Republic	14	1.398	-0.647	0
Arkhangelsk region	20	-0.154	-0.493	0
Vologda region	25	0.882	0.152	1

Continuation of table 3.

Kaliningrad region	15	0.536	-0.157	0
Leningrad region	14	0.75	-0.303	0
Murmansk region	17	0.603	-0.766	0
Novgorod region	32	0.268	0.343	1
Pskov region	35	0.015	0.204	1
St. Petersburg	23	0.752	0.476	1
Republic of Adygea	22	-0.062	0.108	1
Krasnodar Territory	27	0.406	0.224	1
Astrakhan region	9	0.364	-0.691	0
Volgograd region	17	0.129	-0.326	0
Rostov region	33	-0.331	0.111	1
Republic of Dagestan	19	-1.482	0.378	1
Republic of Ingushetia	15	-2.532	1.057	1
Kabardino-Balkarian Republic*	17	-1.043	-0.900	0
Karachay-Cherkess Republic	27	-0.079	0.084	1
Republic of North Ossetia – Alania*	14	-0.704	-0.107	0
Chechen Republic*	13	-10	-8.399	0
Stavropol Territory	23	0.232	-0.254	0
Republic of Bashkortostan	17	-0.312	0.076	1
Republic of Mari El*	31	-0.28	-0.190	0
Republic of Mordovia	15	-0.2	0.328	1
Republic of Tatarstan	13	-0.112	-0.376	0
Udmurt Republic	15	-0.182	-0.723	0
Chuvash Republic	40	-0.057	0.403	1
Perm Krai	20	0.179	-0.428	0
Kirov region	35	0.153	0.024	1
Nizhny Novgorod region	24	0.31	-0.202	0
Orenburg region	6	-0.315	-1.157	0
Penza region*	26	-0.110	-0.108	0
Samara region	20	-0.2	-0.486	0
Saratov region	21	0.219	-0.556	0
Ulyanovsk region	25	-0.363	-0.442	0
Kurgan region	26	-0.091	-0.277	0
Sverdlovsk region	30	0.797	-0.145	0
Tyumen region*	8	-0.897	-0.017	0

Continuation of table 3.

Chelyabinsk region	35	0.089	-0.500	0
Altai Republic	30	0.64	0.921	1
Republic of Buryatia	25	-0.771	0.150	1
Republic of Tyva	17	-0.783	0.219	1
Republic of Khakassia	22	-0.673	0.179	1
Altai Territory	33	0.063	-0.325	0
Zabaykalsky Krai	19	-1.027	0.318	1
Krasnoyarsk Territory*	15	-0.648	-0.380	0
Irkutsk region	15	-0.343	-0.685	0
Kemerovo region*	20	-0.406	-0.341	0
Novosibirsk region	39	0.6	-0.003	0
Omsk region	17	0.164	-0.566	0
Tomsk region	10	0.147	-1.144	0
Republic of Sakha (Yakutia)*	11	-0.756	-0.059	0
Kamchatka region	23	1.136	0.023	1
Primorsky Krai	26	0.397	-0.043	0
Khabarovsk region	21	0.876	-0.073	0
Amur region	17	-0.217	0.863	1
Magadan region	23	1.429	0.664	1
Sakhalin region	18	0.263	0.546	1
Jewish Autonomous region	21	-0.67	-8.399	0
Chukotka Autonomous Area	13	1.103	0.515	1

4 Conclusion

An approach to the formation of recommendations for the development of regions and sectors of the economy, taking into account innovation activity, is proposed. The approach is based on the regression analysis method using an extended economic basis. The approbation of the approach confirmed the possibility of identifying a set of economic sectors whose production volume in the region depends on innovation activity. According to the data of 2019, 20 sectors have been identified, the volume of production of which depends on the innovative activity of the regions aimed at creating international patent applications.

A condition has been formalized under which the sector, having provided the expected volume of production corresponding to the characteristics of the differentiation of the region, becomes strong. It is shown that for each sector of the regional economy, regions can be identified that have sufficient resources to turn the sector into a strong sector based on the realization of the potential for economic growth. For the Construction sector according to 2019 data, 11 such regions have been identified.

The work was supported by the Russian Foundation for Basic Research (RFBR project 20-010-00223)

References

1. U. Blien, K. Wolf, Labour Economics **13(4)**, 445 (2006)
2. M. Fuchs, Empirical Economics **40(1)**, 177 (2011)
3. A. Illy, M. Schwartz, C. Hornych, M. Rosenfeld, J. of Econ. and Soc.Geography **102 (5)**, 582 (2011)
4. S.A. Aivazian, M.Yu. Afanasyev, A.V. Kudrov, Montenegro J. of Economics **14(3)**, 7 (2018)
5. S.A. Aivazian, M.Yu. Afanasyev, A.V. Kudrov Model Assisted Statistics and Applications **15(4)**, 311 (2020) DOI 10.3233/MAS-200502
6. M. Lysenkova, M. Afanasiev, SHS Web of Conferences. Vol. 93 of 3rd International Scientific Conference on New Industrialization and Digitalization, 05002 (2020). DOI: 10.1051/shsconf/20219305002
7. S.A. Aivazian, M.Yu. Afanasyev, A.V. Kudrov, Economics and Mathematical Methods, **54(1)**, 43 (2018)
8. S.A. Aivazian, M.Yu. Afanasyev, A.V. Kudrov, Applied Econometrics, publishing house Synergy (Moscow), **41(1)**, 24 (2016).
9. B. Balassa, “Lafayrade Liberalization and “Revealed” Comparative Advantage”, The Man-chester School **33**, 99 (1965)
10. R. Hausmann, B. Klinger, *Structural Transformation and Patterns of Comparative Advantage in the Product Space*, 128 (CID Working Paper, 2006)