DOI: 10.17323/2587-814X.2023.4.25.40

The impact of economic complexity and industry specialization on the gross regional product of Russian regions

Alexander V. Kudrov
E-mail: kovlal@inbox.ru

Central Economic and Mathematical Institute, Russian Academy of Sciences
Address: 47, Nakhimovsky Prospect, Moscow 117418, Russia

Abstract

The economic complexity index defines the basis of the modern theory of economic complexity and reflects the level of knowledge embedded in the production structure of the economy. This study examines the direct relationship between the economic complexity index and gross regional product (GRP) while taking into account other factors of the GRP production function in its generalized representation. As a result, we can isolate the impact of the economic complexity index from other phenomena. The non-linear nature of the relationship between economic complexity and GRP is revealed, and the direct relationship is manifested only at sufficiently high values of economic complexity, exceeding a certain threshold, which is found endogenously using econometric methods. In addition, the paper studies the relationship between economic complexity and indices of sectoral specialization. We found that there is a direct relationship between economic complexity and the extractive industry index and no relationship with the level of development of manufacturing industry. We obtained a clarification of the generalized production function of GRP, in which the threshold effect of the influence of economic complexity manifested itself as a factor of nonlinear dependence describing the elasticity of labor: a high level of economic complexity provides greater labor productivity. Overall, the results of the study of the dependence of GRP on economic complexity lead to the conclusion that increasing economic complexity can be an effective way to stimulate economic growth and development, but only starting from a certain threshold level. This suggests that an economy must reach a minimum level of diversity and complexity in its industrial activities before it can experience the productivity gains necessary for substantial GRP growth.

Keywords: economic complexity index, sectoral specialization, generalized production function, direct relationships, nonparametric regression, nonlinear regression, returns to scale

Citation: Kudrov A.V. (2023) The impact of economic complexity and industry specialization on the gross regional product of Russian regions. *Business Informatics*, vol. 17, no. 4, pp. 25–40.
DOI: 10.17323/2587-814X.2023.4.25.40

Introduction

For a relatively long time, economists have agreed that a country's ability to create and distribute income depends on its productive structure, as discussed in papers such as [1–3]. Paper [4] presents a study of the convergence of productivity levels across US states and finds that higher productivity levels are associated with a more diverse and complex production structure.

However, quantifying the productive structure of a country is difficult. Various approaches have been attempted, such as the concentration index, which measures the share of agriculture, manufacturing or services in the economy, as well as aggregate measures of diversity and concentration [5]. Other approaches measure diversification by dividing sectors into related and unrelated sectors [6–8]. However, these methods have their limitations, including the possibility of some bias, since large countries or regions tend to be more diversified. In addition, they do not take into account the interconnections between different economic activities, complexity and sophistication of production activities.

These shortcomings are resolved by considering the identified comparative advantages and constructing an economic complexity index [9–11]. The economic complexity index allows us to obtain estimates of the complexity of economic structures, taking into account both the diversity and uniqueness of sectors. This allows us to reflect both the breadth and depth of the economic structure.

One of the most important aspects of economic complexity is sectoral network structure, which measures the degree to which different sectors of the economy are interconnected through production processes. This interconnectedness is believed to facilitate the transfer of knowledge, technology and other resources between sectors and support economic growth. Large values of the economic complexity index indicate that the structure of the economy is dominated by interconnected sectors. For example, sectors with long production cycles, such as electronics and engineering, require higher levels of coordination and knowledge and therefore have high levels of economic complexity. In contrast, economic structures dominated by primary and agricultural sectors have low values of economic complexity.

The paper [12] presents calculations of the economic complexity index for countries and shows how it can be used to forecast economic growth and identify potential areas for diversification and development of the economies of countries.

The relationship between economic complexity and gross domestic product (GDP) is of great interest to economists, since GDP is a widely accepted indicator of regional production and economic development.

The authors of [11] have shown that countries with a more complex production structure tend to have higher levels of economic growth and higher GDP per capita, which in turn are associated with lower poverty rates and better social welfare [13]. Hence, one may conclude that development policies should aim to create conditions that will stimulate growth in economic sophistication (for a more detailed discussion see [10]).

In recent years, statistical studies have used the economic complexity index as an explanatory factor for economic growth, knowledge level, human capital, inequality and other socio-economic indicators [12, 14, 15]. However, the relationship between economic complexity and socioeconomic indicators is not always unambiguous, and there are other factors that may influence this relationship. As will be shown in this paper, often the assumption or conclusions that there is a direct relationship of GRP with economic complexity is erroneous, because, as a rule, other basic indicators of the economy and science are not considered.

In [16], a generalized GRP production function was obtained, in which regional output depends on the number of employed persons (L), the value of fixed assets (K) with their elasticity coefficients, which are given by the sectoral structure of GRP, and the number of researchers (P) (distinguished as an additional production factor with a constant elasticity coefficient). These factors of the production function will be considered as the main characteristics of the economy.

The purpose of this paper is to test two hypotheses using data on the regions of the Russian Federation. First, we will examine whether there is a direct relationship between the index of economic complexity and GRP. Second, we will examine whether there is a direct relationship between economic complexity and the factors of the generalized production function. For this purpose, we will use the methodology of finding direct relationships and hypothesis family testing [16, 17].

1. Data

Let us consider GRP for the year 2019 and the main factors of the generalized GRP production function from [16]:

- gross regional product for the year 2019 [18];
- fixed assets at the end of 2019 [18];
- average annual number of employed persons for 2019 [18];
- indices of extractive (S_1) and manufacturing (S_2) industries for 2019 (see below);
- number of researchers in 2019 [18].

Let us explain in more detail about the indices of extractive (S_1) and manufacturing (S_2) industries, which characterize the sectoral features of the region. These indices were constructed by the author based on the data of GRP sectoral structure using component analysis with rotation and reflect the sectoral specialization of the regions under consideration (*Fig. 1*). The data of the GRP sectoral structure included the main six industries that determine the nature of the economy of the Russian regions [18]:

- agriculture, forestry, hunting, fishing and fish farming;
- mining, oil and gas;
- manufacturing;
- wholesale and retail;
- real estate operations;
- public administration and military security; social security.

Further, these indicators were expressed through two factors by the method of principal components with rotation (*Fig. 1*), which account for more than 80% of the explained variance in the data for the year 2019. For other close years, very similar results are observed, indicating a very slow change in the structure of regional GRP.

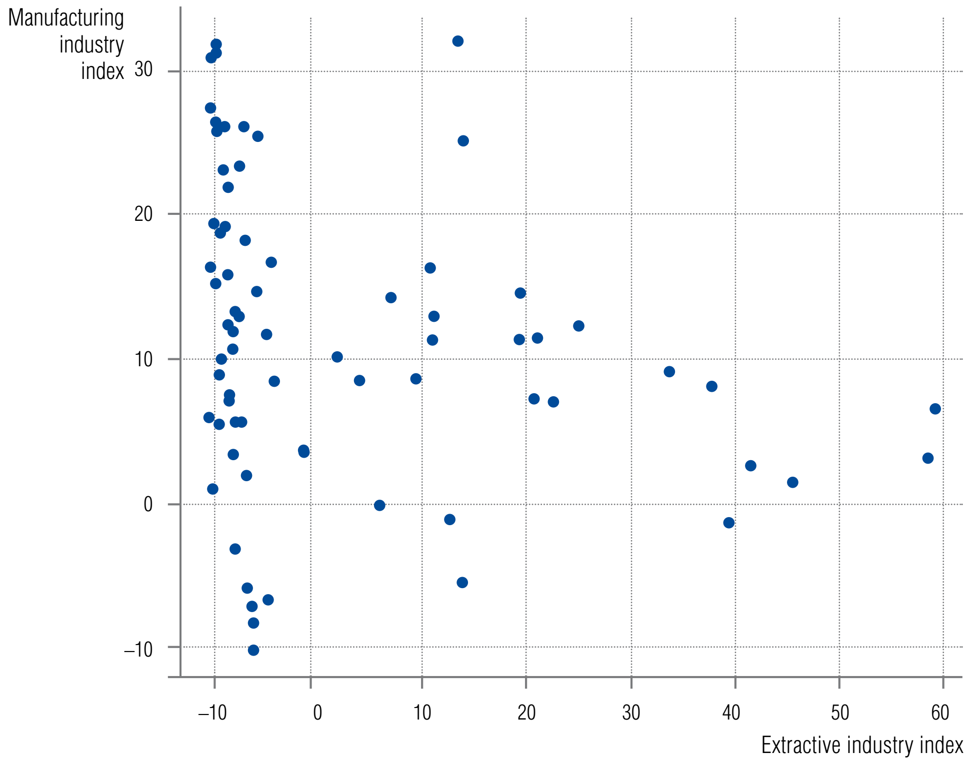

Fig. 1. Distribution of Russian regions in the space of sectoral orientation indices (listed in ascending order; based on data for 2019).

At the same time, we considered the data on tax revenues for 82 sectors of the Russian Federation regions [19], which reflect the production volumes of each sector of the economy for export and for domestic consumption. Based on these data, an index of economic complexity was constructed for 2019–2020 [20]. *Figure 2* shows the estimates we obtained of the economic complexity index. Note that the index of economic complexity, in fact, is equivalent to the generalized eigenvector of the matrix "region–region," the elements of which characterize the nested structures of economies.

The region of economic complexity values can be roughly divided into ranges, within which locally the points are well approximated by linear rank dependencies:

- **Range-1:** regions with a predominance of unique sectors in the economic structure. As a rule, these regions are characterized by specialization in the extractive industry. For them, the average value of the extractive industry index (+13.64). A sufficiently high average value of the manufacturing industry index (+11.05) indicates the presence of regions with a mixed-type structure, where manufacturing sectors are also sufficiently represented.
- **Range-2:** regions with a weakly diversified mix of strong sectors and non-unique sectors. These regions include regions with emerging economies. Average value of the extractive industry index (+7.16), average value of the manufacturing industry index (+7.45).
- **Range-3:** regions with highly diversified structures of strong sectors and long value chains. This includes regions characterized by the presence of long value chains and specialization in manufacturing. The average value of the extractive industry index (–3.87) indicates the absence of extractive industry

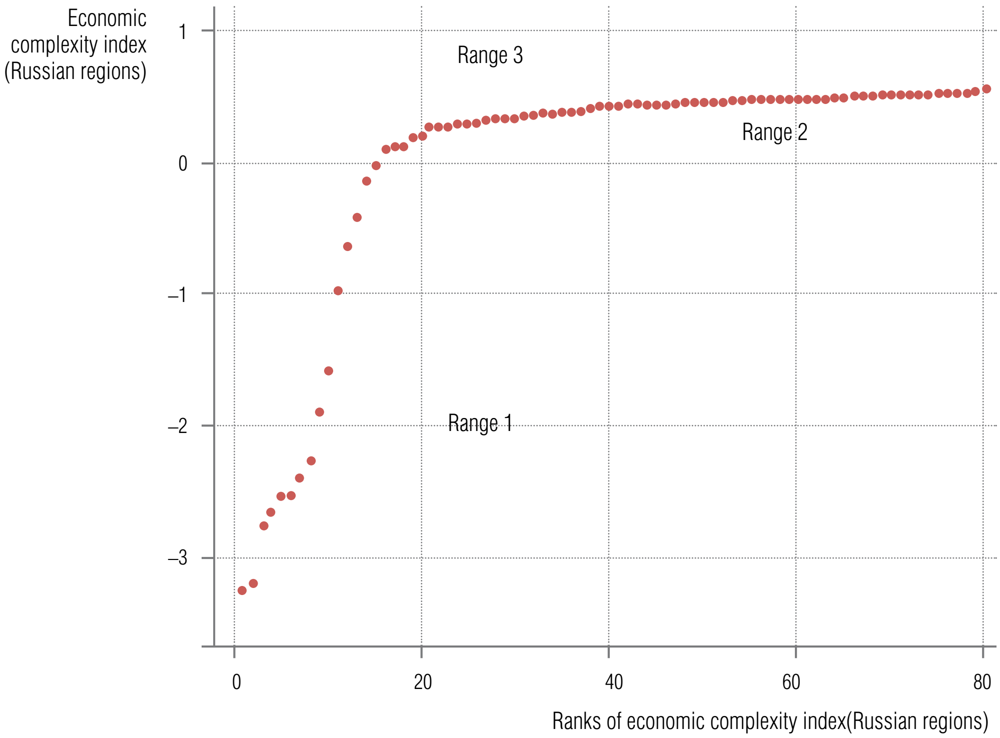

Fig. 2. Estimates of economic complexity by regions of the Russian Federation (listed by ascending order; based on data for 2019).

in most of these regions. The average value of the manufacturing industry index is (+17.68).

It should be noted that the smallest dispersion of economic complexity values is observed for points from Range 1, and the dispersion of points from Range 3 is the largest. The minimum average value of the extractive industry index is for regions with economic complexity in Range 1, and the maximum for Range 3.

2. Research methodology

The proposed methodology aims to obtain an analytical expression describing the impact of economic complexity on gross regional product (GRP). The methodology consists of several steps:

1. Identification of explanatory variables directly related to GRP. The first step is to identify the variables that are directly related to GRP (hereinafter referred to as Y_i – GRP of the i-th region). This is done using the technique of the so-called "causal analysis" or analysis of the direct relationships structure [16]. Let us explain this concept. If in a set of random variables (including both resultant and explanatory variables) $\boldsymbol{Z} = (Z_1, Z_2, ..., Z_n)$, the conditional distribution of value Z_i from all others is determined *only by their part* $Z_j, Z_k, ..., Z_l$ (those not included in the condition can take any values). Let us denote by lowercase letters $z_1, ..., z_n$ the realized values of the corresponding random variables $Z_1, Z_2, ..., Z_n$. Then the definition of direct relationships can be written as:

$$P(z_i \mid z_1, ..., z_n) = P(z_i \mid z_j, z_k, ..., z_l) \;\; z_1, ..., z_n), \quad (1)$$

for all $z_1, ..., z_n$,

and the variables $Z_j, Z_k, ..., Z_l$ are called *directly related* to the variable Z_i. In the continuous case, the partial correlations of Z_i with the directly related (and only with them!) are not zero. Namely, if Z_i and Z_j are directly correlated, then different from zero will be:

$$corr\left(Z_i, Z_j | \boldsymbol{Z}_{-(i,j)}\right) = \\ = corr\left(Z_i - Pr_{Z_i}\left(\boldsymbol{Z}_{-(i,j)}\right), Z_j - Pr_{Zj}\left(\boldsymbol{Z}_{-(i,j)}\right)\right), \qquad (2)$$

where $\boldsymbol{Z}_{-(i,j)}$ – the set of variables excluding Z_i and Z_j;

$Pr_{Z_i}\left(\boldsymbol{Z}_{-(i,j)}\right)$ and $Pr_{Zj}\left(\boldsymbol{Z}_{-(i,j)}\right)$ – projection of Z_i and Z_j onto the linear subspace $sp\left(\boldsymbol{Z}_{-(i,j)}\right)$

2. Identification of the form of the relationship between GRP and economic complexity: monotonic or non-monotonic? To identify the non-monotonic relationship between GRP and economic complexity, the estimation of the non-parametric Nadaraya–Watson kernel regression $g_\tau(x)$ is used [21]:

$$g_\tau(x) = \frac{\sum_{i=1}^{N} w_i(x) \log(Y_i)}{\sum_{i=1}^{N} w_i(x)}, \qquad (3)$$

where $w_i(x) = k\left(\frac{x - X_i}{\pi}\right)$, $k(y) = \frac{\exp\left(-\frac{y^2}{2}\right)}{\sqrt{2\pi}}$;

$\log(Y_i)$ – logarithm of GRP for i-th region;

X_i – ranks of economic complexity for i-th region;

$k(y)$ – kernel of nonparametric regression (3) with parameter τ, τ– window width in nonparametric kernel regression (3).

Note that the window width τ was estimated using the so-called leave-one-out estimate cross validation, see [22] for details:

$$\tau_{opt} = \underset{\tau}{\operatorname{argmin}} \sum_{i=1}^{N} \left(\log\left(Y_i\right) - g_{\tau,(i)}\left(X_i\right)\right)^2, \qquad (4)$$

where (i) means that point i is not considered when computing the nonparametric estimate at point X_i. The use of cross validation with one point left out is particularly useful when the data size is small, as it allows the model to be trained on almost the entire data set. However, for large data, this cross-validation approach can be computationally expensive because the model has to be retrained for each individual data point.

3. Construction of nonlinear regression dependence of GRP on directly related explanatory variables. After the variables that are directly related to GRP (denoted by Y_i – GRP of the i-th region, $i = 1, ..., N$) are identified, the form of non-monotonic dependence of GRP and economic complexity is determined, a nonlinear regression dependence on these variables is constructed:

$$Y_i = f(x_i, \boldsymbol{\theta}^*) + \varepsilon_i, \; i = 1, ..., N, \qquad (5)$$

where $f: \mathbb{R}^k \to \mathbb{R}$ a nonlinear function of the explanatory and directly related to Y_i variables, $x_i \in \mathbb{R}^k$;

$\boldsymbol{\theta}^* \in \mathbb{R}^p$ – vector of true values of the unknown parameters;

(ε_i) are assumed to be independent identically distributed random variables (not necessarily normally distributed) with $E(\varepsilon_i) = 0$ and $Var(\varepsilon_i) = \sigma^2$.

Under the assumption that the function $f(\cdot)$ is known, the parameter vector $\boldsymbol{\theta}$ of model (5) is estimated as the solution to the following problem:

$$\hat{\boldsymbol{\theta}} = \underset{\theta}{\operatorname{argmin}} \sum_{i=1}^{N} \left(Y_i - f\left(\boldsymbol{x}_i, \boldsymbol{\theta}\right)\right)^2. \qquad (6)$$

Finding a solution to this problem is done by numerical methods using the Levenberg–Marquardt algorithm [23, 24].

According to the results presented in [23, 24], under sufficiently large n and appropriate regularity assumptions (such as twice continuously differentiable $f(\boldsymbol{x}_i, \boldsymbol{\theta})$ with respect to $\boldsymbol{\theta}$), the LSE-estimator $\hat{\boldsymbol{\theta}}$ has an asymptotically normal distribution:

$$\hat{\boldsymbol{\theta}} \sim N_p\left(\boldsymbol{\theta}^*, \sigma^2\left[\left(F\left(\boldsymbol{\theta}^*\right)\right)^T F\left(\boldsymbol{\theta}^*\right)\right]^{-1}\right), \qquad (7)$$

where $F\left(\boldsymbol{\theta}^*\right) = \left[\left.\frac{\partial f\left(\boldsymbol{x}_i, \boldsymbol{\theta}\right)}{\partial \boldsymbol{\theta}_j}\right|_{\boldsymbol{\theta}=\boldsymbol{\theta}^*}\right]_{i,j} \in \mathbb{R}^{N \times p}$.

Thus, this methodology combines several statistical techniques including causal analysis, non-parametric estimation and non-linear regression to establish the relationship between economic complexity and GRP.

3. Research results

Let us analyze the mutual relations between economic complexity and the above-mentioned characteristics of science, the economy in the regions of the Russian Federation. For this purpose, the matrix of partial correlations was estimated using the data for the year 2019 and the hypotheses about the absence of direct relationship between each variable and economic complexity were consistently tested (*Table 1*).

Table 1 (right part) presents the results of testing the family of hypotheses about equality of partial correlations to zero; for a more detailed description of the testing procedure of the hypotheses considered [16]. Units indicate cases when there are no direct relationships between economic complexity and the corresponding variable.

As can be seen from *Table 1*, economic complexity is not related to the index of manufacturing industry, but it is related to the index of extractive industry. *Figure 3* provides visual confirmation for that.

Among all the variables considered, economic complexity has a statistically significant direct relationship with the extractive industry index. The existence of this relationship indicates that the scenario of a transition from an extractive-based economy (e.g., mining or oil, gas) to a more diversified one (in particular, oriented towards long value chains) is associated with an increase in the level of economic complexity.

The lack of a relationship between the economic complexity index and the manufacturing index may imply that the mere presence of manufacturing in the economy is not sufficient to increase its complexity. This may be the case if manufacturing is concentrated in a few low-complexity industries or if other sectors of the economy remain underdeveloped.

As shown in *Table 1*, the partial correlation for GRP and the economic complexity index is insignificant, while in the case of the partial correlation for ranks, there is a statistically significant relationship for these variables (hypothesis accepted at 5% level). This indicates the existence of a non-linear relationship between the index of economic complexity and GRP.

Let us take a closer look at the form of dependence of the logarithm of GRP on the ranks of economic complexity (*Fig. 4*).

Table 1.

Statistical estimates of partial correlations with economic complexity (for original variables and their ranks) for 2019. Results the family of hypothesis testing for equality of partial correlations to zero

	Partial correlation with economic complexity	H0: partial correlation with economic complexity is zero	Partial correlation with economic complexity for ranks	H0: partial correlation with economic complexity for ranks is zero
Fixed assets	0.16	1	0.20	1
Average annual number of employed persons	–0.30	0	–0.06	1
GRP	–0.14	1	–0.24	0
Researchers	0.03	1	–0.05	1
Extractive industry index	–0.53	0	–0.61	0
Manufacturing industry index	–0.12	1	–0.01	1

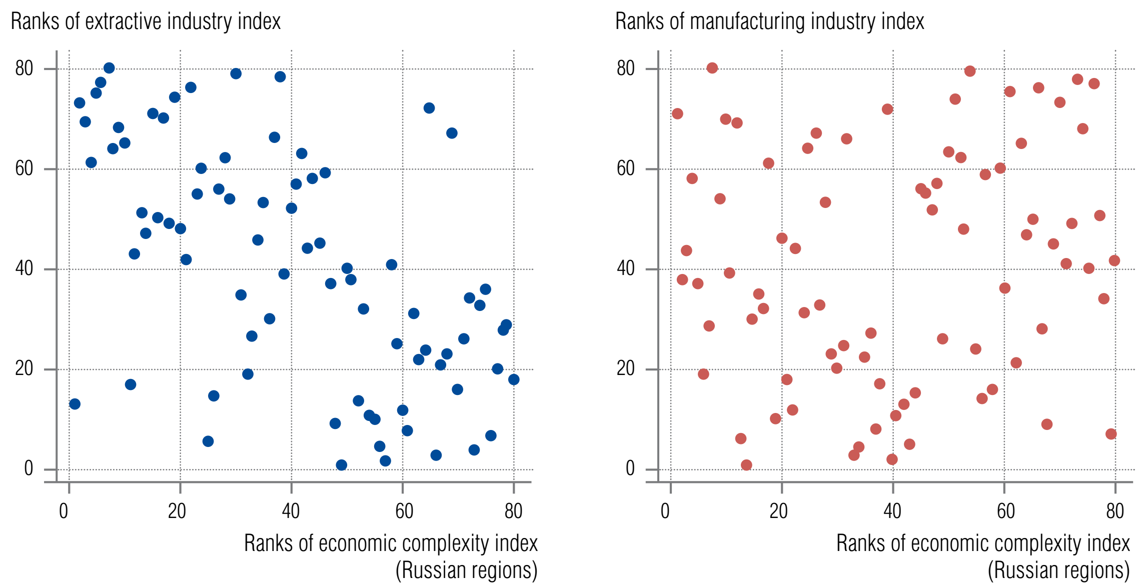

Fig. 3. Right: index of extractive industry (ranks) and economic complexity (ranks).
Left: index of manufacturing industry (ranks) and economic complexity (ranks).

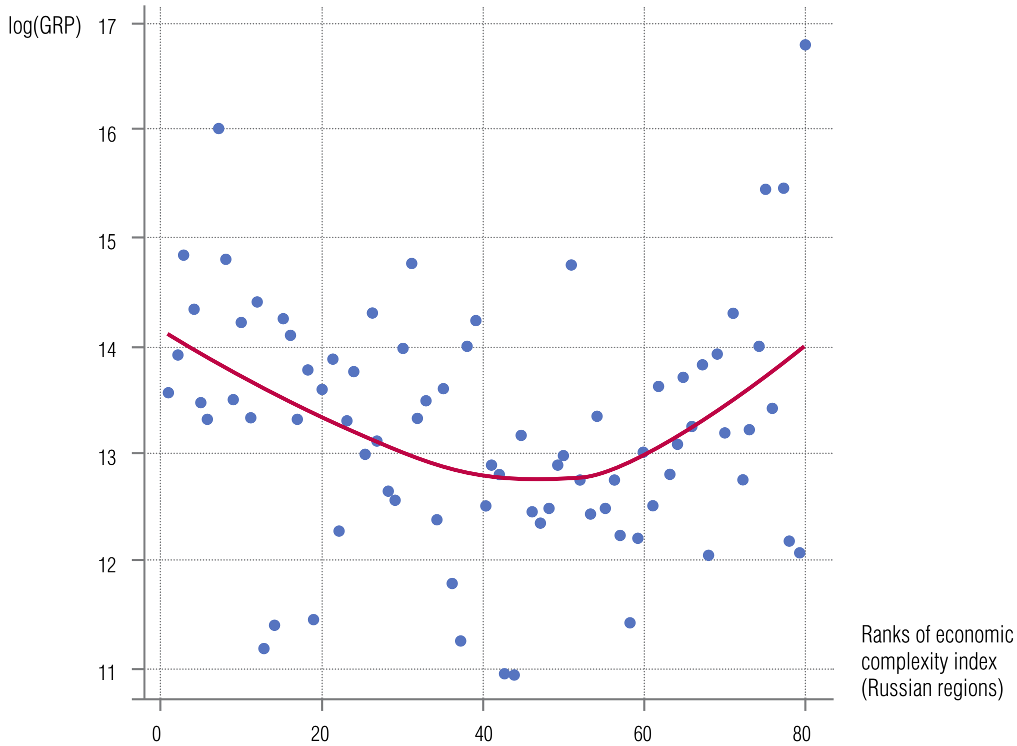

Fig. 4. GRP and ranks of economic complexity for the year 2019;
nonparametric Nadaraya–Watson regression (3).

Note that the U-shaped dependence in *Fig. 4*, obtained using nonparametric Nadaraya–Watson regression, suggests that there cannot be a simple linear relationship between economic complexity and GRP.

A U-shaped relationship between economic complexity and GRP implies that both very low and very high levels of economic complexity correspond to high GRP, while medium levels of economic complexity correspond to lower values of GRP. Thus, we can distinguish the following types of regional economies:

1. Low economic complexity, high GRP: regional economies tend to be rich in natural resources and their GRPs are heavily concentrated in extractive industries such as oil, gas or mining. Despite the low complexity of their economies (as they are mainly focused on one or a few sectors), these regions can have high GRPs due to the high market value of their resources and intensified mining.

2. High complexity, high GRP: regional economies with high economic complexity tend to have a wide range of developed and interconnected industries that include high-tech industries. These regions are typically characterized by high levels of industrialization, investment in human capital and technological innovation.

3. Medium level of complexity, lower GRP: regional economies that are in the process of transitioning to a more diversified and complex economy. There is a lack of developed capacity to efficiently produce more complex goods and services.

Thus, according to *Fig. 4*, we can distinguish two possible paths to higher GRP: (i) through natural resource extraction or (ii) through the development of a more sophisticated industrialized economy. Each pathway has its own advantages and challenges. For example, resource-rich regions may achieve high GRP quickly, but they may face instability due to fluctuations in commodity prices and may have difficulty diversifying their economies.

Due to the non-monotonicity of the correspondence between the logarithm of GRP and economic complexity, we take as a threshold for economic complexity the argument x_{opt}, at which the minimum of the constructed nonparametric Nadaraya–Watson regression (3) $g_{\tau_{opt}}(x)$ is reached (*Fig. 4*):

$$x_{opt} = \underset{x}{\operatorname{argmin}}\, g_{\tau_{opt}}(x) = 46.$$

The rank x_{opt} corresponds to an economic complexity value of 0.45.

Let's estimate the threshold impact of economic complexity on GRP:

$$cor(\text{GRP}, ECI \mid ECI \geqslant 0.45, X_{-(\text{GRP}, ECI)}) = 0.79,$$

$$cor(\text{GRP}, ECI \mid ECI < 0.45, X_{-(\text{GRP}, ECI)}) = -0.18,$$

where $X_{-(\text{GRP}, ECI)}$ all considered indicators of science and economy except GRP and ECI.

Thus, only at values of economic complexity exceeding 0.45, there is a direct relationship between GRP and the economic complexity index.

Based on the identified threshold direct relationship between economic complexity and GRP, the representation of the extended production function for GRP was summarized:

$$Y = c \cdot K^{\beta_1(S_1, S_2)} L^{\beta_2(S_1, S_2, T)} P^{\gamma} + \epsilon, \tag{8}$$

where

$$\beta_1(S_1, S_2) = \frac{\mu_1 e^{(\mu_2 \cdot S_1 + \mu_3 \cdot S_2)}}{1 + \mu_1 e^{(\mu_2 \cdot S_1 + \mu_3 \cdot S_2)}}, \beta_2(S_1, S_2, T) =$$

$$= \frac{\lambda_1 e^{\left(\lambda_2 \cdot S_1 + \lambda_3 \cdot S_2 + \lambda_4 \cdot T^2\right)}}{1 + \lambda_1 e^{\left(\lambda_2 \cdot S_1 + \lambda_3 \cdot S_2 + \lambda_4 \cdot T^2\right)}};$$

$$T = \begin{cases} ECI, \text{ if } ECI \geq 0.45 \\ 0, \text{ otherwise}; \end{cases}$$

$c, \gamma, \lambda_1, \lambda_2, \lambda_3, \lambda_4, \mu_1, \mu_2, \mu_3$ – constants;

Y – gross regional product in 2019;

K – fixed assets at the end of 2019;

L – average annual employment for 2019;

P – number of researchers for 2019;

ECI – economic complexity index calculated from data for 2019;

S_1 and S_2 – indices of extraction and manufacturing, respectively, calculated for 2019;

ϵ – errors of model (8).

Note that the expression found for GRP is estimated with greater accuracy, namely $R^2 = 0.982$, which is greater than in [16].

Endogeneity in model (8) occurs when the error ϵ is statistically dependent on one or more explanatory variables among K, L, P, S_1, S_2, T. Namely:

$$E(\epsilon \mid K, L, P, S_1, S_2, T) \neq 0.$$

It is well known that the presence of endogeneity leads to bias and invalidity of the LSE-estimators of the model parameters, leading to incorrect conclusions about the statistical significance of the relationships. To test the hypothesis about the absence of endogeneity is equivalent to the hypothesis testing for independence of each explanatory variable and errors in model (8). To test independence, we use the Hilbert–Schmidt independence criterion [25]. In contrast to the statistics in the Hausman test for independence of explanatory variables and the residuals of the model, which is assumed to be linear [26], the Hilbert–Schmidt test for independence allows for the presence of nonlinearity. A high value of the Hilbert–Schmidt independence criterion for a pair of variables indicates their dependence, while a low value corresponds to independence. Assuming that the null hypothesis is independence of the pair of variables under consideration, *Table 3* presents the results of the test.

As can be seen from the results of *Table 3*, the hypothesis that errors in model (8) are independent of explanatory variables is not rejected.

In order to make sure that the observed lack of relationship between errors and explanatory variables is not due to confounding variables, we test for conditional independence. For this purpose, we can also use the Hilbert–Schmidt independence criterion (*Table 4*).

Thus, according to the results presented in *Table 4*, the hypothesis of conditional independence of errors and explanatory variables in model (8) is also confirmed.

Table 2.

Parameter estimates of model (1) and their statistical significance

	Estimate	**Sd. error**	***t*-value**	***p*-value**	
C	6.77	0.42	4.53	0.00	***
μ_1	1.79	0.21	2.72	0.01	**
μ_2 (extractive industry index; fixed funds)	0.01	0.00	3.53	0.00	***
μ_3 (manufacturing industry index; fixed funds)	–0.02	0.01	–3.68	0.00	***
λ_1	0.33	0.26	–4.35	0.00	***
λ_2 (extractive industry index; employed)	–0.01	0.01	–1.96	0.05	*
λ_3 (manufacturing industry index; employed)	0.05	0.01	3.83	0.00	***
λ_4 (economic complexity)	3.34	1.16	2.89	0.01	**
γ (researchers)	0.05	0.02	2.81	0.01	**

Denotations: *** – p-value at less than 0.001 level, ** – p-value at less than 0.01 level, * – p-value at less than 0.05 level.

Table 3.

Testing the hypothesis of independence of the errors of model (8) and its explanatory variables

Pairs of variables	Hilbert–Schmidt independence criterion	p-value	Existence of independence
ϵ, T	0.0000033	0.13	independent
ϵ, K	0.0000364	0.8	independent
ϵ, L	0.0000366	0.79	independent
ϵ, P	0.0000257	0.97	independent
ϵ, S_1	0.000193	0.27	independent
ϵ, S_2	0.000237	0.23	independent

Table 4.

Testing hypotheses about conditional independence of errors of model (8) and its explanatory variables

Pairs of variables | condition	Hilbert–Schmidt independence criterion	p-value	Existence of independence
$(\epsilon, T \| T, L, P, S_1, S_2)$	0.00000154	0.11	independent
$(\epsilon, K \| T, L, P, S_1, S_2)$	0.0000503	0.98	independent
$(\epsilon, L \| T, K, P, S_1, S_2)$	0.000258	0.45	independent
$(\epsilon, P \| T, K, L, S_1, S_2)$	0.000149	0.89	independent
$(\epsilon, S_1 \| T, K, L, P, S_2)$	0.000368	0.77	independent
$(\epsilon, S_2 \| T, K, L, P, S_1)$	0.000657	0.18	independent

The presence of a statistically significant positive parameter λ_4 at truncated economic complexity suggests the possibility of the effect of "spillover" of innovations. Regions with a more complex production structure tend to have greater diversification, which creates opportunities for inter-sectoral diffusion of knowledge and technology, which in turn can lead to increased innovation and productivity growth. In addition, a region that produces a variety of products and has interconnected production processes has more opportunities to exploit economies of scale.

Figure 5 illustrates that increasing returns to scale are characteristic of regions with high elasticity of labor and low elasticity of capital.

As can be seen from *Fig. 5*, the growth of labor elasticity is accompanied by a decrease in the elasticity of capital and vice versa. This indicates a shift in the production function due to sectoral differences of regional economies.

Figure 6 shows that the presence of diminishing returns to scale is characteristic of regions with a high concentration of extractive industries in the structure of the regional economy. Declining returns to scale mean that a proportional increase in labor and capital leads to a less than proportional increase in output. This may be because extractive industries (e.g. mining, oil and gas) are often capital intensive and may face problems such as resource depletion, environmental regulations or high operating costs.

Figure 7 shows that increasing returns to scale are characteristic of regions with high manufacturing concentration and high economic complexity. Sufficiently large values of economic complexity, exceeding the threshold of 0.45, correspond to large returns to scale.

Since the economic complexity index characterizes the concentration of related sectors in the structure of an economy, as an economy becomes more complex, networking or relatedness facilitates the sharing of best practices and collaboration on innovation, and hence contributes to higher productivity. In a complex economy characterized by intricate inter-sectoral linkages and advanced production, the wealth of diverse knowledge and skills tends to be high. Sectoral relatedness allows this knowledge to be transferred between sectors, contributing to overall productivity. With high levels of sectoral relatedness, innovations and technological advances are more easily diffused across related sectors.

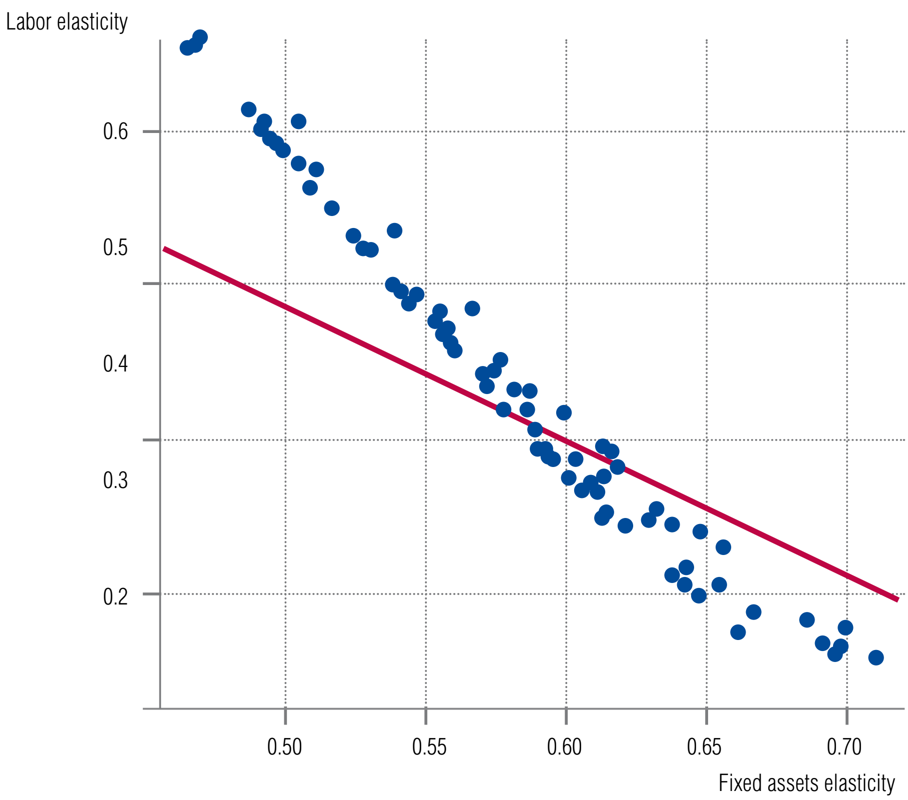

Fig. 5. For each region of Russia according to model (1):
$\beta_1(S_1, S_2)$ – elasticity of fixed assets, x-axis; $\beta_2(S_1, S_2, T_1,)$ – the elasticity of labor, y-axis.
Straight line: $x + y + \gamma = 1$.

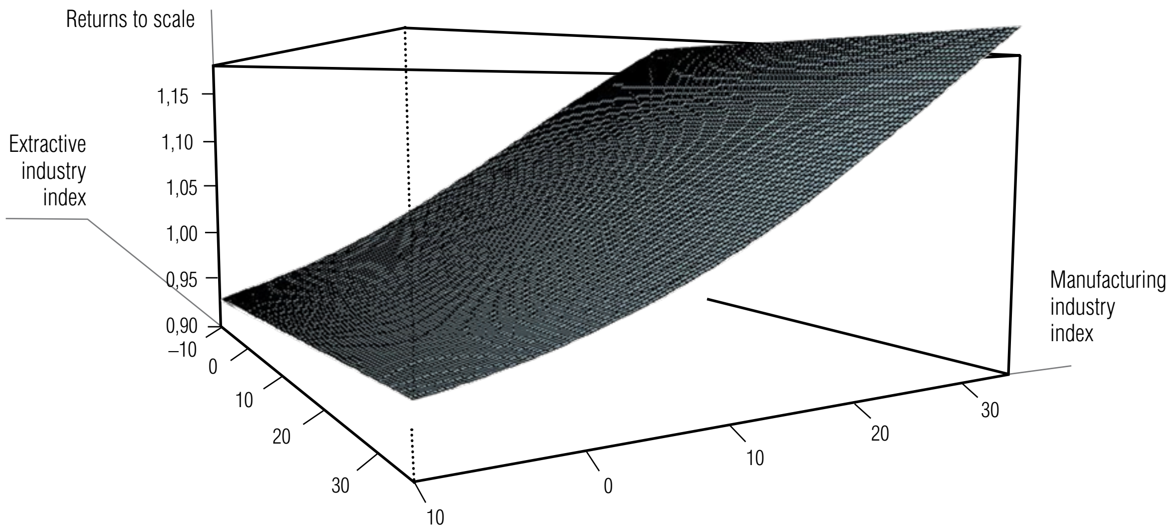

Fig. 6. Extractive industry index (x-axis); manufacturing industry index (y-axis); returns to scale (z-axis) calculated as $\beta_1(S_1, S_2) + \beta_2(S_1, S_2, T_1,) + \gamma$.

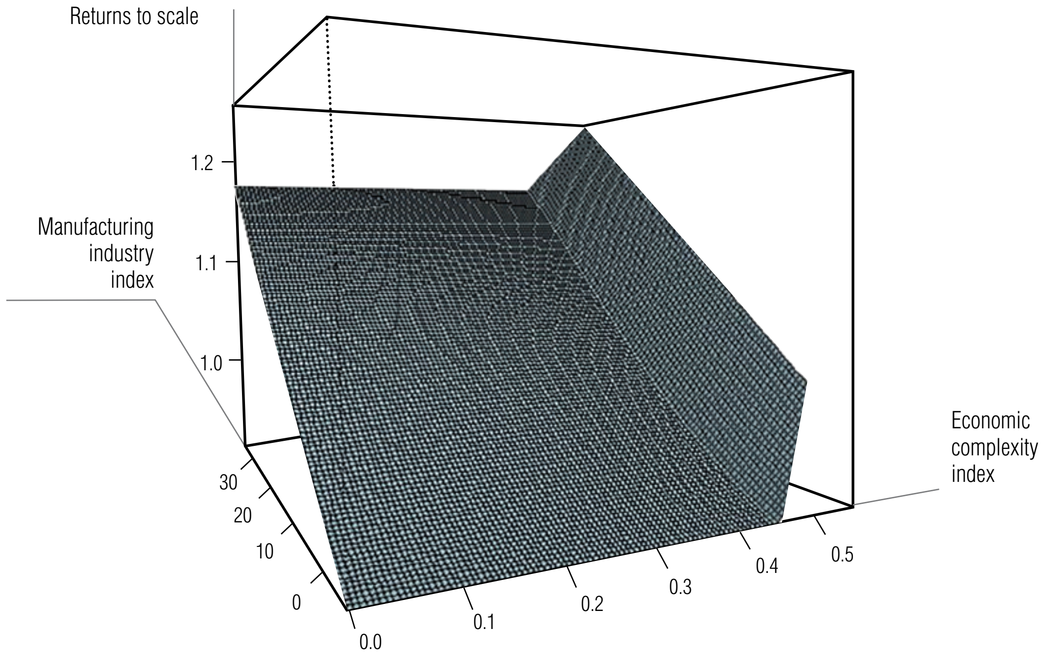

Fig. 7. Economic complexity index (x-axis); manufacturing index (u-axis); returns to scale (z-axis) calculated as $\beta_1(S_1, S_2) + \beta_2(S_1, S_2, T_1,) + \gamma$.

The formation of related sectors encourages co-development, where sectors do not grow in isolation, but through the joint development of technology, skills and knowledge. Such interconnected growth can further increase productivity through synergies between different sectors.

Thus, regions with more diverse and complex production structures with specialization in manufacturing are better positioned to benefit from economies of scale and respond to economic change. As noted earlier, regions with more complex economic structures tend to have more diversified economies, making them more adaptable to volatile economic conditions.

Conclusion

The most important results of the econometric study of the impact of economic complexity on the GRP of the Russian Federation regions, performed through the consistent use of three statistical methods (partial-correlation analysis of identifying direct relationships between variables, the Nadaraya–Watson method of nonparametric regression estimation, and the least squares method for nonlinear production functions) based on the statistical data for the year 2019, are as follows:

- There is no direct statistical relationship between the manufacturing industry index and economic complexity for the regions of the Russian Federation. This means that the emergence of new manufacturing sectors or the expansion of previously existing ones is not necessarily accompanied by an increase in economic complexity.
- The extractive industry index has a direct relationship with the economic complexity index. An increase in the extractive industry index corresponds to a decrease in the economic complexity index.
- Statistical evaluation of non-parametric Nadaraya–Watson regression showed that there is a non-linear relationship between GRP and the economic complexity index.
- Having ranked the regions by the level of economic complexity and excluding the influence of other variables in the sample, the rank number and the corresponding level of economic complexity were found, above which there is a direct relationship between GRP and economic complexity, and below which there is no such relationship.
- Statistical estimates of the parameters of the considered generalized production function show that the elasticity of fixed assets depends in a statistically significantly manner on the indices of sectoral specialization, while the elasticity of labor depends both on the indices of sectoral specialization and economic complexity. For values of economic complexity above a certain threshold, high economic complexity corresponds to higher labor elasticity. This indicates that regions with a more complex, diverse and interconnected production structure have higher productivity and, consequently, have more opportunities for efficient use of their labor resources.
- Increasing returns to scale are evident only in regions where manufacturing industries predominate and there is a sufficiently high level of economic sophistication. Regional economies with a high concentration of extractive industries are characterized by decreasing returns to scale, which potentially limits their growth.
- Manufacturing can provide more opportunities for productivity and value addition than extractive industries.

In general, the results of the study of the relationship between GDP and economic complexity emphasize the importance of taking into account economic complexity as an explanatory variable of the production function for regional GRP in its generalized form. Stimulating increases in economic complexity can be an effective way to promote economic growth and productivity, but this effect is only evident when the level of economic complexity is high enough. By increasing the diversity and economic complexity of their production structures, regions can increase productivity, competitiveness and economic stability, leading to higher levels of GRP and sustainable economic growth.

The importance of the composition of economic sectors and the balance between labor and capital in shaping output and growth should also be emphasized.

Regions concentrated in sectors with high labor elasticities are characterized by increasing returns to scale and hence potentially higher economic growth. Conversely, regions with a high concentration of extractive industries may experience declining returns to scale, potentially limiting their growth. This underscores the importance of policies aimed at increasing productivity and diversification away from extractive industries for sustainable economic growth.

The methodology presented in this paper for quantifying the impact of economic complexity on gross regional product (GRP) can be useful in the decision-making process of locating new production facilities, distribution centers or branches of enterprises. Understanding the impact of economic complexity on GRP can help identify economically stable and sufficiently diversified regions with more favorable business conditions. However, regions with a complex economic structure are also characterized by a higher potential level of competition.

Higher GRP usually correlates with higher purchasing power of consumers. Therefore, the results presented may help businesses to identify potentially lucrative regional markets for their products or services. However, it is important to note that, while useful, this methodology is one tool and should be used in conjunction with other data sources and market research to make comprehensive decisions. ■

References

1. Hirschman A.O. (1958) *The strategy of economic development*. New Haven: Yale Univ. Press.
2. Rosenstein-Rodan P.N. (1943) Problems of industrialization of eastern and southeastern Europe. *The Economic Journal*, vol. 53(210/211), pp. 202–211.
3. Teece D., Rumelt R., Dosi G., Winter S. (1994) Understanding corporate coherence: Theory and evidence. *Journal of Economic Behavior & Organization*, vol. 23(1), pp. 1–30.
4. Bernard A.B., Jones C.I. (1996) *Productivity and convergence among U.S. States*. National Bureau of Economic Research.
5. Imbs J., Wacziarg R. (2003) Stages of diversification. *The American Economic Review*, vol. 93(1), pp. 63–86.
6. Boschma R., Iammarino S. (2009) Related variety, trade linkages, and regional growth in Italy. *Economic Geography*, vol. 85(3), pp. 289–311.
7. Frenken K., Oort F.V., Verburg T. (2007) Related variety, unrelated variety and regional economic growth. *Regional Studies*, vol. 41(5), pp. 685–697.
8. Frenken K., Saviotti P. (2008) Export variety and the economic performance of countries. *Journal of Evolutionary Economics*, vol. 18(2), pp. 201–218.
9. Hausmann R., Hidalgo C.A., Bustos S., Coscia M., Chung S., Jimenez J., Simoes A. (2007) The building blocks of economic complexity. *PLoS One*, vol. 2(1), e268.
10. Hausmann R., Hidalgo C.A. (2017) *Atlas of economic complexity: Mapping paths to prosperity*. MIT Press.
11. Hidalgo C.A., Hausmann R. (2009) The building blocks of economic complexity. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 367(1897), pp. 1817–1825.
12. Hausmann R., Hidalgo C.A., Bustos S., Coscia M., Simoes A., Yildirim M.A., Hwang J. (2014) *The atlas of economic complexity: Mapping paths to prosperity*. MIT Press.
13. Ravallion M. (2004) Competing concepts of inequality in the globalization debate. *Brookings Trade Forum Globalization, Poverty, and Inequality*, pp. 1–38. Available at: https://www.jstor.org/stable/25063189 (accessed 14 February 2023).

14. Hartmann D., Guevara M.R., Jara-Figueroa C., Aristarán M., Hidalgo C. (2017) Linking economic complexity, institutions, and income inequality. *World Development*, vol. 93, pp. 75–93.

15. Hidalgo C.A., Klinger B., Barabási A.L., Hausmann R. (2007) The product space conditions the development of nations. *Science*, vol. 317(5837), pp. 482–487.

16. Gavrilets Y.N., Kudrov A.V., Tarakanova I.V. (2018) Analysis of the internal structure of economic growth potential. *Bulletin of CEMI RAS*, vol. 1, no. 1 (in Russian).

17. Gavrilets Y., Kudrov A., Tarakanova I. (2022) Statistical analysis and modeling of regional economy and science relationship. *Ekonomika i matematicheskie metody*, vol. 58(4), pp. 56–70. https://doi.org/10.31857/S042473880023019-9 (in Russian).

18. *Regions of Russia. Socio-economic indicators, 2017–2019*. Moscow: Rosstat (in Russian).

19. *2019–2020 tax base and accrual structure reports for taxes and levies*. Available at: https://www.nalog.gov.ru/rn77/related_activities/statistics_and_analytics/forms/ (accessed 14 February 2023) (in Russian).

20. Afanasiev M., Kudrov A. (2021) Economic complexity and embedding of regional economies' structures. *Ekonomika i matematicheskie metody*, vol. 57(3), pp. 67–78. https://doi.org/10.31857/S042473880016410-0 (in Russian).

21. Bierens H.J. (1994) *Topics in advanced econometrics estimation, testing, and specification of cross-section and time series models*. Cambridge University Press.

22. Bishop C.M. (2006) *Pattern recognition and machine learning*. Springer.

23. Ratkowsky D. (1993) Principles of nonlinear regression modeling. *Journal of Industrial Microbiology*, vol. 12, pp. 195–199.

24. Seber G.A.F., Wild C.J. (2003) *Nonlinear Regression*. New York: Wiley–Interscience.

25. Gretton A., Fukumizu K., Harchaoui Z., Sriperumbudur B.K. (2009) A kernel statistical test of independence. *Advances in Neural Information Processing Systems*, pp. 1067–1074.

26. Hausman J.A. (1978) Specification tests in econometrics. *Econometrica*, vol. 46(6), pp. 1251–1271.

About the author

Alexander V. Kudrov

Cand. Sci. (Phys.-Math.);

Leading Researcher, Central Economic and Mathematical Institute of the Russian Academy of Sciences, 47, Nakhimovsky Prospect, Moscow 117418, Russia;

E-mail: kovlal@inbox.ru

ORCID: 0000-0003-2495-5496