Universities’ efficiency and regional economic growth: empirical evidence from Russia

Tommaso Agasisti

*Politecnico di Milano School of Management, Italy*

Aleksei Egorov\*, Daria Zinchenko and Oleg Leshukov

*National Research University Higher School of Economics, Russia*

\* Corresponding author: aegorov@hse.ru

**Abstract.** This paper analyses the link between efficiency of regional higher education systems in Russia and rates of regional economic development. The efficiency scores are calculated at institutional level using a double-bootstrap data envelopment analysis procedure taking into account different internal characteristics of universities that may affect its production process, and then aggregated at regional level. We formulate a regional economic endogenous growth model that considers the efficiency of regional higher education systems as one of the explanatory variables by means of a robust GMM estimator. The model also includes spatial interactions between regional economies and between regional higher education systems in neighboring regions. The findings highlight a positive, substantial and statistically significant effect of HEIs’ efficiency on the regional economic growth rate We also found negative spillover effects indicating that efficient regional higher education systems may extract resources from the regions in the neighborhood.

**JEL Classification.** I25, I21, E02

**Keywords.** Regional economic development; efficiency in higher education; knowledge spillovers; economic growth

**1. Introduction**

 Higher education institutions are often considered as economic agents and analyzed in terms of their economic activity. Universities may be engaged in processes of social and economic development of territories where they operate and, consequently, they might contribute to economic growth (Belenzon and Schankerman, 2013; Pinheiro et al., 2012; Varga, 1997). There is a large number of empirical studies that show positive and causal relationships between development level of higher education systems and rates of economic development (Valero and Van Reenen, 2016). Moreover, most of the existing economic growth theories consider human capital as one of the most important determinants of economic development (Hanushek, 2016).

Human capital stock is usually measured as the number of years of schooling and this proxy-variable is used in economic growth functions (Barro and Lee, 2013). Positive empirical relationship between average years of schooling in the country or region and economic growth rates reflects the fact that all levels of education (school, university, etc.) may have positive impact of economic growth. However, higher education is especially important in this context, since this level of education provides specific set of skills that is needed for generating new ideas and innovations. In such setting, universities are considered not as a burden for the state budgets, but as investment in human capital development that can bring positive returns in the future. Such perception of higher education has its policy implications. Development of regional or national higher education systems is often considered as a policy instrument that can lead to positive economic outcomes.

Being treated as economic agents, universities have become the subject of debate on the efficiency of their activities (Kosor, 2013). If universities can generate economic impact and public funds are invested in them in order to intensify this impact, society can require effective allocation of this funds. As economic agents, higher education institutions have a goal of maximizing their outputs (teaching, research and “third mission”) using limited amount of resources. The debate on economic impact of universities and debate one the efficiency of their activities are very close to each other and this relationship is very important for elaboration of public policy in the field of higher education.

Russia is an appropriate case where these issues can be investigated. Current federal public policy in Russia in the field of higher education tries to take into account both engagement of universities in processes of social and economic development and efficiency of their activities. In 2012 the “Annual Monitoring of Efficiency of Higher Education Institutions” was launched by the Ministry of Education and Science of Russia. Using this policy tool, the Ministry aims at detecting the universities that are inefficient and assumes managerial decisions on such cases that may include reorganization. In response to the challenge of the limited links between higher education institutions and regional administrations and enterprises, a special federal program (so-called Flagship universities program) was launched in 2015 for increasing universities’ efficiency in terms of ensuring a positive impact on regional economic development.

Current empirical studies that analyze relationship between development level of education system and economic development in Russia (Egorov, et al., 2017) give an evidence that scale and quantity of higher education system matters for economic development. However, the discussion regarding the role of universities’ efficiency is rather limited. The aim of this paper is to explore the link between the efficiency level of regional higher education systems in Russia and the rates of economic growth of regions where these systems operate.

The analysis is based on the endogenous economic growth theory and consists of three methodological steps. First, the efficiency scores of particular universities are estimated using double-bootstrap data envelopment analysis (DEA) following Simar and Wilson (2007). Second, the results obtained at the first step are aggregated to the regional level and regional higher education systems efficiency scores are obtained. Third, a model for regional economic growth model is proposed treating the efficiency of regional higher education system as one of the explanatory variables. Specifications of the models take into account the structure of regional economies and spatial effects both in gross regional product growth rates and efficiency of regional higher education systems. Such specification allows testing the hypothesis about existence of spillover effects on a geographical basis. The model is estimated by means of a robust GMM system that handles the problem of potential endogeneity between universities’ efficiency and economic performance.

The paper is organized as follows. Section §2 discusses the background and particularities of regional development in Russia and specific features of Russian higher education system. Section §3 presents existing literature on economic impact of universities and the concept of efficiency in higher education. Section §4 contains the underlying theoretical framework of this research, while section §5 illustrates the methodology of the study and describes the dataset. Section §6 discusses the main results, and section §7 contains some policy implications and concluding remarks

**2. Background of regional development and higher education in Russia**

*2.1 Higher education in Russia*

The Russian system of higher education has undergone an unprecedented worldwide reform over the past 30 years and the landscape of university’s system is still being transformed. This process is primarily forced by the collapse of the USSR and the transition to a market economy (Froumin et. al., 2014).

In Soviet times, universities were the part of unified system of national economy and they were obligated to integration into national supply chains. The curriculum base, specialization, size and even the territorial location of universities was regulated by the central government. The higher education system was centralized and subject to rigid control, in adherence to the state's political agenda (Johnson, 2008). The principle of mandatory job placement of university graduates into specific workplaces across the entire country allowed to plan the recruitment for certain specialties demanded by the Soviet economy.

In the conditions of the emergence after 1991 a new nationhood, market mechanisms and new branches of the economy a higher education system was forced to adapt to new social and economic realities. Such changes implied the following consequences:

* the emergence of a private sector of higher education (the opening of private universities and the right to public universities to open tuition fees admission);
* liberalization of education programs and increase in the university autonomy (Froumin and Leshukov, 2015);
* massification of higher education and a sharp increase in the number of universities. The number of head universities had increased [1.5 times](http://context.reverso.net/%D0%BF%D0%B5%D1%80%D0%B5%D0%B2%D0%BE%D0%B4/%D0%B0%D0%BD%D0%B3%D0%BB%D0%B8%D0%B9%D1%81%D0%BA%D0%B8%D0%B9-%D1%80%D1%83%D1%81%D1%81%D0%BA%D0%B8%D0%B9/1.5%2Btimes) during the past 20 years (see Figure 1). Also the growth of the higher education sector was ensured by the opening of branches - as of 2017 there are 840 branches of universities in Russia (597 public and 243 private). This led to an increase of higher education coverage - as of 2016, the share of young people aged 17-25 years attending higher education programs is 31.8%.
* adaptation of universities to the needs of the market, primarily at the local level. The termination of model of mandatory job placement of graduates forced the universities to adjust to the needs of the regional population (the main educational “boom” occurred in the economic, management and law programs).

Most of these changes proceed in chaotic way, not accompanied by coordination at the national level. Therefore, since the mid-2000s period of rational public reforms in the sphere of higher education has been undertaken which determined the basis of the current system of higher education. The main declared tasks of the public policy in the field were maintenance control over universities and improving the quality of educational and scientific activities.

**[Insert figure 1 here]**

However, the issue of governance of large and heterogeneous country as the Russian Federation remains particularly urgent. Thus, the regions of Russia are significantly heterogeneous in terms of the scale of higher education systems. A third of all head universities are concentrated in two regions - in Moscow with the Moscow region (185 universities) and St. Petersburg (66 universities), while in the half of other regions the number of universities does not exceed five organizations.

 The Soviet legacy defined the current principle of centralization of university governance - 90% of universities are regulated by federal government which defines the Russian education governance system as the one of the most centralized in the world (Morsy et al., 2018). The financial and regulatory possibilities for the participation of regional authorities in the development of universities are substantially restricted (Froumin and Leshukov, 2015). Therefore, the issue of assessing the contribution of universities to the economic development of the regions of the Russian Federation in conditions of power centralization and minimal opportunities for direct interaction between universities and regions becomes especially topical from an academic and institutional viewpoint.

*2.2 Regional development in Russia*

 Russia is a federation that is rather natural form for such large and heterogeneous country. Today Russia Federation consists of 85 regions that are equal among themselves in relations with federal government in accordance with the Constitution. There are different legal types of regions in Russia (“oblast”, “republic”, “krai” etc.), but all of them have their own legislative, executive and judicial branches.

Historically, the main character of Russian regional development was the high level of regional differentiation. Regions were very different in terms of their socio-economic characteristics. Due to large territory of the country Russian population is distributed unevenly, regions located in the European part of Russia make up only 30% of the total area, while the share of population that resides there is equal to 73% (Sardadvar and Vakulenko, 2016). Only the three regions (Moscow, St. Petersburg and Tyumen region) generate 35% of the entire GRP of the country.

These differences became stronger during the transformation period after the collapse of Soviet Union (Froumin and Leshukov, 2018). In order to deal with this challenge in the beginning of 2000 federal authorities started the regional development policy aimed at support of poor performing regions. Federal center concentrated substantial amount of resources and then distributed it in favor of regions with budget deficit. In other words, underdeveloped regions began to be supported by highly economically developed ones.

At the same time, the current state policy is aimed at supporting regions on a competitive basis. Particularly, special economic zones are established in such regions in order to make them more attractive for investors through tax exemptions. Another government program suggests establishment of clusters in particular regions that bring together and intensify the interaction between industry, scientific organizations, universities, etc. This interaction is supposed to strength competitiveness of the territory where it forms. These reforms force the regions to be proactive in their development, while maintaining a high level of the control from the center (Froumin and Leshukov, 2018). But general policy keeps a high level of control and management of regional development on the federal level.

 These policies have positive impact, however, the regional inequality in Russia remains at the high level. For instance, in 2015 the gross regional product growth rates ranged between -6.2% to 6.9%; Gross regional product per capita – between 1.1 ml rub and 0.09 ml rub; share of workforce with higher education – between 47.8% and 22.1%; share of students in age cohort – between 14.8% and 46.6%. Despite this differentiation, Russian regions are not closed from the economic prospective, they actively interact with each other in order to find drivers for their development. There are empirical studies that confirm the existing of spillover effects (Demidova and Ivanov, 2016); (Egorov et al., 2017).

Current regional development policy pay a special attention to higher education sector as a source of regional growth. Universities are considered as organizations that can attract resources to the region and provide its competitiveness. This paradigm was reflected, particularly, in special government program – formation of “flagship university” aimed to facilitate regional development by their education and research activities, “third” mission. Current regional development context in Russia raises the questions of how university can contribute to regional development and how this contribution can be increased. These questions are addressed by this paper.

**3. Literature on economic impact of universities and the concept of efficiency in higher education**

*3.1 Existing evidence on the role of higher education for economic growth*

There are different approaches to the description of how university may be engaged in processes of social and economic development ranging from triple helix (Etzkowitz, 1993) to different econometric approaches based on the theoretical macroeconomic (Romer, 1986) and regional development (Capello, 2011) models. This paper considers universities as economic agents and analyzes it in terms of macroeconomic models of economic growth.

There are three main approaches to evaluation of universities’ economic impact in the literature. The first one is the *traditional economic approach* (Elliott et al., 1998) suggesting that universities make different economic transactions within regional economy – they pay salaries, taxes, purchase goods and services, provide job positions in the local labor market. These transactions increase aggregate demand in regional economy and, consequently, transform it into gross regional product with multiplier effects. The second approach is the *skill-based framework* of university contribution (Bluestone, 1993). Within this framework universities are considered as generators of human capital for regional labor market. This line of reasoning considers two main channels. First of all, human capital theory (Becker, 1964) suggests that workers with higher level of education tend to earn higher salary (Mincer, 1974), so people with higher education tend to have higher level of expenditures in regional economy. Secondly, more educated labor force is usually considered as more productive and influence the gross regional product growth rates directly. Finally, the third approach to HEIs’ economic impact evaluation considers universities as *drivers of regional innovation systems* (Huggins and Johnston, 2009), that can cooperate with innovative companies, stimulate entrepreneurship, contribute to the development of new economic activities.

This paper is based on macroeconomic theory on economic growth and analyses higher education as one of the factor of economic development. The theory of economic growth has been developed and extended over a long period of time. Robert Solow (1956; 1957) provided the basic modern framework for economic growth modelling. He elaborated the long-run economic growth model, which included the factor of technological progress in addition to the standard determinants of economic growth, i.e. physical capital accumulation and labor-force increases. The technological progress variable was determined exogenously and contained those part of economic growth rate which could not be explained by the increase of labor-force and physical capital. This neoclassical growth model suggested that physical capital accumulation and labor-force increases constituted the foundation of economic growth, but the growth of workforce productivity and capital increased as a result of this technological progress. Therefore, human capital is not a driver of economic growth in the neoclassical paradigm of growth modelling: its increase does not lead to a greater long-run growth rate.

 The new theories of economic growth proposed since the late 1960s, on the contrary, confirmed that human capital had positive impact on the country’s economic growth rate of the country for long-term development. This perspective is identified as endogenous growth models. This theoretical paradigm was discussed in the seminal contribution by Romer (1986). The more detailed analysis of human capital as economic growth factor in the context of endogenous growth theory was provided in the models of Lucas (1988) and was continued in the study of Mankiw, Romer and Weil (1992).

The most recent models indicate the larger importance of tertiary education particularly for countries near the technological frontier (i.e. the most economically-developed countries), where the growth requires new innovations (Vandenbussche et al., 2006; Aghion et al., 2009).

 In the economic literature, several approaches exist to measuring human capital: school attainment and results of standardized achievement tests of students (cognitive skills) (Hanushek, 2016). The concept to quantify human capital in terms of school attainment was early suggested by Jacob Mincer (1974). This approach has some limitations. The first one is the prerequisite that one year of schooling correlates with the same amount of learning in all countries. The second limitation is that schooling is considered as the only source of human capital (Hanushek, 2016). The second approach to measuring human capital via assessments of skill levels was first applied in Hanushek and Kimko (2000) and was extended in Hanushek and Woessmann (2007, 2015). Such a position measures human capital directly and reflects net indicator of the knowledge capital. Hanushek (2016) demonstrated that while considered measures of human capital are included simultaneously in the model, school attainment was not significantly related to growth. This can be attributed to direct relationship between skills and years of schooling.

 The concept of endogenous economic growth model is constantly reviewed and extended. In the model elaborated by Romer (1986) technological progress is determined by knowledge accumulation, while Lucas (1988) considers education and human capital accumulation as drivers of productivity growth. Under the latter approach of endogenous economic growth model human capital stock and the development level of higher education institutions (HEIs) can determine the technological progress. The fundamental idea is that development and expansion of higher education contributes to economic growth. Thus, a new branch of researches on the economic impact of universities emerged.

 Under the latter approach of endogenous economic growth model a variety of proxy-variables for development level of HEIs has been proposed. Usually, the development level of HEIs are simply measured by student population, the number of universities and the share of student in the appropriate age-cohort (Valero and Van Reenen, 2016).

*3.2 Why the efficiency of universities can have an impact on economic growth*

The variables listed in previous section are quantitative measures of development level of HEIs and take into account teaching activities of HEI’s. However there exist more qualitative measure of development level of HEIs. It is measured through universities’ efficiency estimates and considers indicators of the at least two traditional missions of university. This approach was proposed by Barra and Zotti (2017). These authors assumed to use the university efficiency together with simple measure of human capital development in the specification of endogenous economic growth model. The main author's hypothesis is that the presence of efficient university in a certain area has positive impact on its economic growth. This approach to modeling growth including university efficiency was extended in Agasisti et al. (2017). It was proposed to use more complete measure of efficiency and to take into account the “third mission” (Laredo, 2007) or “knowledge transfer” (Bekkers and Freitas, 2008) of HEIs[[1]](#footnote-1).

**4. Theoretical framework**

 Based on the related literature review the following theoretical framework describing how efficiency of higher education system may be related with rate of regional economic development was formulated.

 Firstly, since we consider university as economic agent, we can assume that university can be characterized by the production function of the form (1):

 Where composite result of university’s activities (teaching, research, third mission etc.); resources available for university (financial resources, competence of academic staff etc.) that are used in order to generate composite output.

 Production function of the form (1) specifies the production possibility frontier meaning that (- output of the university ). When the university is said to be perfectly efficient. It other words, higher efficiency indicates position of the unit closer to production possibility frontier and, consequently, relatively higher output of the unit with the fixed amount of resources (2):

(2)

 Where is a set of available resources for university (), is an efficiency of university *i*. In such setting the position of the university relative to the frontier (1) depends on the amount of available resources and on how efficiently these resources are used.

 Next, we can consider aggregate output of all universities in the region (3):

(3)

 Where is an output of university *i, N* – number of universities in the region. is positively related to the individual efficiencies of universities in the region and to the amount of available resources.

 Without loss of generality, we can assume that the regional economy can be described of the production function of the analogous type (4):

 Where gross regional product; capital stock in the economy; labour resources available for university; technological progress that is endogenously determined inside the economy in case of endogenous economic growth model.

 Since parameter A is endogenously determined, we may assume that it depends on different internal parameters of regional economy, including performance of regional higher education system that is positively related to efficiency (5)

 Where parameter reflecting technological progress; composite result of university activities; all other factors that potentially may determine the level of technological progress.

 Due to the link (3) we may assume that regional higher education system efficiency can be directly associated with economic growth rates.

 In such setting, we can incorporate university production function and regional economy production function (6):

Where all notations are the same as in equations (1)-(5).

 There are different intuitive explanations for the relationship between universities’ efficiency and rates of regional economic growth, as determined in research papers (Agasisti et al., 2017), who considered the following channels: (i) relatively efficient university may employ less staff in order to achieve the same result, and this additional labor force may find alternative occupation in regional economy and generate additional gross regional product. (ii) If the universities are perceived by society as efficient organizations, they may be more successful in building relationships and collaborations with other organizations in the region. (iii) Efficient universities create incentives for becoming more efficient for other institutions with which university interacts. (iv) More efficient universities can product relatively larger amount of output with the same resources.

 All these explanations are valid also for Russian economy. However, in our theoretical setting, the most important explanation is the forth one. Since we have assumed that universities’ efficiency level may affect the level of technological progress that consequently determine the rate of GRP growth, higher output of higher education institutions can translate into more intense influence on technological progress and on rates of economic growth.

 One more important consideration in our theoretical setting is spatial interactions. There is an argument that the borders of regional economy sometimes do not correspond to the administrative borders of real regional economy. That is why it is important to analyze spatial effects that account for this particularity of regional economic development in Russia. The same argument may be formulated for regional higher education system, so the model should also include spatial effects for efficiency levels. In other terms, we expect that geographical proximity with efficient peers would stimulate the efficiency of universities, somehow as in a yardstick competition setting.

 Based on this theoretical framework we can formulate the following research hypothesis:

(H1) Efficiency of regional higher education system is an important determinant of regional economic growth;

(H2) Positive spillovers exist: rates of regional economic growth are determined by the efficiency of higher education systems in neighboring regions.

**5. Methodology and data selection**

In order to achieve the aim of the research – analyze the link between efficiency level of regional higher education systems and rates of regional economic growth and to test the hypotheses H1 and H2 we employ a methodology that consists of three steps.

*5.1 Efficiency estimation on institutional level*

We employ a two-stage Data Envelopment Analysis (DEA) and the bootstrap procedures suggested by Simar and Wilson (2007). Simar and Wilson (1999) demonstrated that DEA scores obtained in the first stage are biased, and the environmental variables from the second stage are correlated to output and input variables. This fact recommends using of bootstrap procedure to overcome this problem. The two-stage DEA assumes that environmental variables might affect university outputs and proposes to re-estimate the DEA model with adjusted outputs for environmental variables through the bootstrap procedure.

DEA analysis involves a selection of orientation (input or output) and the type of returns to scale. We considered output-oriented model with assumption of constant returns to scale (CRS). CRS assumption was considered as suitable one, as universities cannot increase the effects rapidly ([Nazarko](https://scholar.google.ru/citations?user=fPc2DEAAAAAJ&hl=ru&oi=sra) and S[aparauskas](https://scholar.google.ru/citations?user=S3jnP1QAAAAJ&hl=ru&oi=sra), 2014). Output-oriented model evaluates how much outputs can be increased holding inputs fixed (Agasisti and Perez-Esparrells, 2010). It seems to be reasonable assumption in the context of efficiency in higher education due to the legislative fixing of education production costs.

The linear programming model assuming output-oriented framework and constant return to scale must be solved for the Decision Making Unit[[2]](#footnote-2) (DMU), which transforms an input vector into an output vector :

Here refers to the value of efficiency of the DMU and satisfies . measures the efficiency of the unit as the distance to the frontier, which is the linear combination of the best practice units. implies that the DMU is below the best practice frontier (inefficient), while indicates that the DMU lies on the estimated frontier (efficient). is a vector of intensity variables (Simar and Wilson, 2007).

 Following Simar and Wilson (2007) approach, the second stage model is constructed as censored regression:

 (8)

 Where represents the efficiency score from solving equation (7), is the vector of environmental variables that can influence the efficiency through the vector of parameters and indicates the statistical noise. Equation (8) is estimated only for .

More detailed description of the algorithm used and the obtained results is presented in the Annex 1.

 According to Afonso and St. Aubyn (2006) the results, obtained with the two-stage DEA approach, are likely to be biased in a small sample. Firstly, this bias is explained by serial correlation of . Secondly, environmental variables are correlated with the error term . In order to overcome these problems Simar and Wilson (2007) elaborated bootstrap method, which used in this article.

 In order to formulate the efficiency model it is necessary to make some assumptions regarding production process in the university and about the input and output sets. At this step, we have to consider universities as multi-product organizations (Baumol at al., 1982) that utilize different inputs in order to produce different outputs. Referring to the existing literature, we assume universities’ production technology with three inputs. The first input is the financial resources of the university measured by the income of universities from all sources in constant prices (normalized by the number of academic staff). This variable is rather common in the research concerning universities production functions (such input is used, for example, in Agasisti and Johnes, 2009; Agasisti and Perez-Esparrells, 2010). The second input is the relative weight of academic staff with advanced degrees (Candidate of Sciences, i.e. Russian analogue to PhD degree, and Doctor of Sciences) in total number of academic staff in universities (excluding part-time staff and independent contracts). This variable measures the human resources available for higher education institutions in order to carry-out teaching and research activities. Existing research in this field usually use the total number of academic staff as an input that measures the universities’ human capital resources (Agasisti and Johnes, 2009; Agasisti and Perez-Esparrells, 2010; Agasisti and Pohl, 2012; Wolszczak-Derlacz and Parteka, 2011). However, using the share of academic staff with advanced degrees we want to capture not the quantity, but the quality of human resources available for the university. The third input that we employ is an average unified state exam score that reflects the quality of entrants. This variable is rather important for university, however, often it is not clear is this indicator an output or an input in the production process (Abankina et al., 2013). If we treat this variable as input, we assume that more prepared students are important resource for the university. If we consider it as output, the underlying assumption is that this variable reflects the ability of university to attract the most talented students indicating the reputation of the university. We follow (Johnes, 2006; Barra and Zotti, 2017; Agasisti et al., 2017) and consider an average entrance exam score as an input variable.

 The set of outputs consists of three variables reflecting three different activities of higher education institutions – teaching, research and engagement in collaborations with industrial partners as a proxy for third mission. The first output is the total number of publications indexed in Web of Science, Scopus and Russian science citation index (normalized by the total number of academic staff). This variable reflects the scientific productivity of universities’ academic staff (Parteka and Wolszczak-Derlacz, 2013; St. Aubyn et al., 2009). The second output is the total income from grants obtained for applied research and developments carried out by university. This variable reflects the intensity of engagement of university in collaborations with industrial partners and measures partially the money spent by companies on applied researches conducted by universities. In the Russian context, this factor seems to be a good proxy for third mission that reflects the cooperation between universities and industries. Finally, the third output is the total number of students per academic staff. The most widespread indicator of universities’ teaching activity is the number of graduates (Agasisti and Johnes, 2009; Bonaccorsi et al., 2007; Agasisti and Pohl, 2012). Due to the data availability constraints, we cannot use the same indicator – as it is not disclosed by official statistics. However, taking into account the fact that dropout ratio in Russia is very close to zero, the correlation between total number of students in university and the total number of graduates is very high. So, in the particular case of Russia, we can replace the total number of graduates by the total number of students without losing much information.

 In order to take into account different internal characteristics of higher education institutions that potentially may affect the production process inside the university we used a set of environmental (exogenous) variables to determine the efficiency scores correctly. Different papers that examine the efficiency in higher education sector stress that there are different factors besides inputs and outputs that may affect production process adopted by university. For instance, some of the environmental variables are population size of the city where university is located, percentage of students with need-based financial aids (for example, see Nazarko and Saparauskas , 2014); year of university’s foundation (Agasisti et al., 2017; Wolszczak-Derlacz and Parteka, 2011); dummy variable reflecting the presence of medical school in the university (Agasisti et al., 2017; Wolszczak-Derlacz and Parteka, 2011); real GDP per capita in the regional where university is arranged, number of different faculties, the share of core funding revenues in total revenues, the share of women in academic staff (for example, see Wolszczak-Derlacz and Parteka, 2011); number of years since a technology transfer office opened in the university, the percentage of dropouts, funding received from central government (for example, see Agasisti et al., 2017).

 Given the inputs and outputs of the efficiency model and the specification of economic growth model we employ five exogenous variables. The first two variables reflect the structure of student body of the university: the share of master’s students in the total number of students in the university; the share of full-time students in the university. Such indicators of students’ body structure, undoubtedly, influence the university strategy and the structure of production process. For instance, if the most students in the university are part-time, the university utilize fundamentally different educational model with different structure of costs and resources. The third exogenous variable is the dummy variable reflecting that university is located in the capital city of the region. Underlying assumption here is that university located in the capital city is usually oriented towards students from the whole region. Moreover, capital cities are usually more attractive for living, so compared to the universities in non-capital cities, these institutions may be more attractive for students from other cities and regions, and this heterogeneous level of attractiveness might in turn affects efficiency. The level of competition on the regional higher education market is also important factor that may determine the level of efficiency (Leshukov et al., 2015). Universities that operate in highly competitive environment tend to consolidate their resources and perform better (see also the conceptual discussion in Agasisti, 2009). The general measure of competition that we used as exogenous variable in efficiency model is the share of university’s students in the total number of students in the region. Finally, we use special dummy variable that indicates the presence of medical faculty in the university. Such field of study as medicine may have strong influence on the technology used by university. For example, universities with medical faculty usually have their own hospitals; the structure of education process differ from the study programs in other fields. These particularities may also affect the production technology and the associated efficiency, thus it should be taken into account.

*5.2 Efficiency of regional higher education systems*

In addition to considering environmental variables under the assessment of the efficiency scores of HEIs one more issue here is to measure the efficiency on the level of the whole regional higher education system. Due to the particularities of regional structure in Russia and data availability constraints, in this study we cannot associate one university with particular sub-regional territory like in (Agasisti et al., 2017), so it was decided to aggregate efficiency score obtained on the institutional level to the regional level. Efficiencies calculated at the second step report weighted averages of university performances by the total number of students of local universities in certain regions according to formula (9):

 number of region

number of university in the region i

 quantity of universities in the region i

efficiency score

 total number of students in university

At the end of this process of aggregation, we obtain synthetic indicator that measures the average level of universities’ efficiency for any given region, and the subsequent analysis of economic growth is conducted at regional level.

*5.3 Economic growth modeling*

The third and the final step of our methodology is to estimate economic growth model with regional higher education system efficiency score among explanatory variables. In order to construct this model we employ endogenous economic growth model (Romer, 1986; Lucas, 1988; Rebelo, 1991). Within the framework of this theoretical model we can expect that education in general and efficiency of higher education system in particular may cause long-term economic growth (Hanushek, 2016), as discussed in sections §2 and §3.

The specification of the regional economic growth model is represented by the formula (10):

 Where gross regional product growth rate; log of gross regional product in previous period; investment growth rate; population growth rate; share of public sector in gross regional product; share of commercial minerals extraction in gross regional product; share of industries in growth regional product; share of employed population with higher education; efficiency level of regional higher education system; efficiency spatial lag; gross regional product growth rates spatial lag; individual region-specific effects; time effects; errors.

In order to construct spatial interactions we used the simple inverse distance matrix. This choice is based on the existing evidence that estimates and inferences in spatial regression models are not sensitive to the spatial weight matrix choice (LeSage and Pace, 2014).

 The regional economic growth model contains standard variables that are assumed to be related with gross regional product (GRP) growth rates and used in most research devoted to economic growth modelling based on the endogenous economic growth theory. These variables are investment growth rate; employed population growth rates; the share of employed population with higher education that reflects accumulated stock of human capital in the region; log of gross regional product in previous period that captures convergence effect (Sala-i-Martin, 1994).

We employ also a set of variables that capture the structure of regional economies. These variables are the share of public sector (education, public administration, healthcare) in gross regional product, share of commercial minerals extraction in GRP and the share of industries in GRP. There are some alternative ways of taking into account regional economy structure in economic growth modelling. For instance, Kufenko (2015) used average commercial minerals extraction per capita. Our model accounts not only for commercial minerals extraction, but also for public sector and industries, so the optimal way to take into account the economy’s structure is to consider the shares of these sectors in gross regional product.

The third set of the variables included in the model is needed to resolve the research hypotheses of this study. These variables are efficiency of regional higher education system; spatial lag of the efficiency that accounts for spillover effects (Agasisti et al., 2017); growth spatial interaction that accounts for positive spatial correlation in regional growth rates in Russia (Demidova, 2015).

*5.4 Data and descriptive analysis*

 The source of the data for efficiency evaluation is the Annual Monitoring of Efficiency of Higher Education Institutions[[3]](#footnote-3) conducted by Russian Ministry of Education and Science. This monitoring was launched in 2012, so today the data for the period from 2012 to 2016 is available. However, we used only data for the period from 2012 to 2015 since some macroeconomic characteristics on the regional level that are needed at the third step of methodology for economic growth model estimation were not available at the time that study was conducted.

Only public and head universities of Russia were included in the analysis. Such limitation imposed to the sample is needed for reducing the level of university heterogeneity in terms of their production functions. The limitation does not reduce its representativeness, since non-public universities account for the 5-year average 18% of the student population[[4]](#footnote-4). Given all constraints, the sample represents 449 universities that located in 77 regions of Russia and have data for each year within the period from 2012 to 2015.

Since the outlier-problem and missing values should be taken into account in the analysis of efficiency, preliminary data processing was implemented. In order to deal with missing values we have implemented imputation procedure based on classification and regression trees algorithm. The description of this algorithm is available at the Annex 3. In order to eliminate outliers, we used capping correction – upper outliers were replaced by the values that correspond to the quantile 0.975; the lower outliers were replaced by the values that correspond to the quantile 0.025. Detailed description of the capping procedure for missing data imputation and capping procedure for outliers are presented in the Annex 3.

Our starting point will be a descriptive analysis of institutions’ characteristics needed to assess the university efficiency level. Table 1 presents the key descriptive statistics on the institutions in our sample. Consider the set of variables that are treated as universities resources. The first variable shows the total income of educational organization from all sources per number of Faculty members. The value of this variable was increasing during the period 2012-2014 (9% growth in 2014 compare to 2012 with the largest increment of 7% in 2014 compared to 2013). In 2015 the amount of income per academic staff showed significant reduction by 17% in comparison with 2014. This reduction is largely due to the financial crisis in Russia that occurred in 2014–2015. The next institutions’ characteristic is the relative weight of academic staff with advanced degrees. This indicator is characterized by stable growth dynamics – it increased from 64% in 2012 to 70% in 2015. The third input variable is the average entrance exam score, which does not pronounce some dynamics and has rather stable values.

The first output indicator is the number of publications in science journals indexed in RSCI, Web of Science and Scopus, per capita of academic staff. This indicator has increased significantly during the period under review: the indicator was 58% higher in 2015 in comparison with 2012. This dramatic rise can be mainly attributed to initiating of Performance Based Funding scheme in 2015. However, the universities were notified about the programme requirements several years in advance. The general principle of this funding scheme is that the number of publicly funded student slots available for the university is determined according to the formula, and scientific productivity is one of the components in this formula. In other words, the amount of publications increased due to additional stimulus for higher education institutions for higher scientific productivity. The next output variable is the total quantity of R&D per one employee of academic staff. The value of this indicator has decreased for the period 2012-2014. There was the largest reduction of this variable value (18%) during the period between 2014 and 2015. This reduction was determined by negative macroeconomic shock in Russian economy in this period of time and consequently by decrease of private funds for R&D implementation. The last output indicator is the students to academic staff ratio. This indicator is rather stable in time and the value averages is about 10 students to academic staff.

The share of masters’ students, that is the first environmental variable, has risen by 2.9%: and has reached 7.3% in total number of students in 2015. This growth was primarily due to increase of number of state-funded places in master’s programs. The share of full-time students in total number of students virtually unchanged for the past five years. The absence of a clear dynamics was related to rather stable proportion of budget slots for full-time students and

In overall, we can observe multidirectional dynamics of considered institutional characteristics so we can expect ambiguous dynamics of universities’ efficiency during the period.

 **[Insert table 1 here]**

The source of the data for regional economic growth model estimation is Russian Federal State Statistics[[5]](#footnote-5). The sample includes 77 regions that have efficiency scores of their regional higher education systems. We used data for most of the variables for the period from 2012 to 2015, but the variable for GRP growth rate covers the period from 2011 to 2015. The extension of the period for this variable is needed to use more lags of the dependent variable in model identification using sys-GMM approach. In order to deal with outlier-problem we used the same capping procedure as in the case of dataset for efficiency estimation. Detailed description of the capping procedure is presented in the Annex 3.

 The descriptive statistics for the variables used for regional economic growth model estimation are presented in the Table 2. Average GRP growth rate was declining during the period from 2012-2015. Due to economic crises, in 2015 Russian regions on average experienced economic contraction (-0.2% in 2015). This decrease in GRP growth rate was followed by decline in the investment growth rate. The investment growth rate decreased by 12.5 percentage points – from 2.6 % in 2012 to -9.89 % in 2015. The most rapid decline (4.3 p.p.) of this variable stood out between 2014 and 2015 years. The dynamics of the share of employed population with higher education corresponds to the non-standard mechanism according to which Russian labor market adopts to macroeconomic shocks. The distinctiveness of this mechanism lies in the dominance of price adaptation over quantitative one. In other words Russian labor market is characterized by combination of strong adaption in terms of wages and weak adaptation in terms of employment rate. Great flexibility of wages mitigates negative shocks thereby protecting employment and stemming unemployment growth (Gimpelson and Kapeliushnikov, 2015). Another variable that in spite of negative economic shocks was growing is efficiency level of regional higher education systems. This indicator was 7% higher in 2015 compared to 2012. There was the largest increase of this variable value (6%) during the period between 2012 and 2013. The others variables used in economic growth model estimation had rather stable values and does not have some dynamics.

**[Insert table 2 here]**

*5.5 Dealing with endogeneity and other econometric details*

 In economic growth model represented by equation (10) we may suspect the problem of endogeneity i.e. the correlation between some regressors and the error term. Ignoring this problem, we obtain biased and not consistent parameters estimates that lead to incorrect interpretation of modelling results. Particularly, we may assume here that not efficiency level of regional higher education system influences the GRP gross rates, but relatively efficiency regional higher education systems tend to be formed in the regions with high rates of economic development. The most widespread ways of overcoming this problem is instrumental variables estimator. However, the choice of the relevant instruments is rather difficult problem since these instruments should not be correlated with the error term and should explain substantial share of the variance of endogenous covariates.

 In order to deal with this problem in equation (8) we employ GMM dynamic panel data estimator (sys-GMM) (Arellano and Bover, 1995). Sys-GMM estimator acts like instrumental variables approach and instruments endogenous variables by their lags. Employment of this technique gives us an evidence that we explore causal relationship between efficiency of regional higher education system and gross regional product growth rates. The basic argument here is that efficiency in year t affects the efficiency of year t+n but not directly economic growth in the same year t+n. In order to check the reliability of the model standard statistical tests that are Hansen-Sargan test for overidentifying restriction and second-order autocorrelation test. The estimation strategy can be summarized as follows. At the first step the model is estimated by taking first differences (Wooldridge, 2002). Secondly, lagged levels of the dependent variables are added in order to estimate two-step sys-GMM model (Roodman, 2006). In order to obtain robust results we use corrected standard errors (Windmeijer, 2005). We use all available 4-years panel lags as instruments in sys-GMM regression.

**6. Illustration of main results**

*6.1 Efficiency estimation*

 As anticipated, in order to obtain efficiency scores on institutional level taking into account exogenous factors double-bootstrap data envelopment procedure was implemented. Detailed description of double-bootstrap procedure, as well as intermediate calculation results are presented in the Annex 2. Double-bootstrap procedure uses Farrell’s efficiency concept (Farrell, 1957) to analyze the efficiency level of higher education institutions. According to this concept, it fulfills the condition of , with a value of one indicating a university belongs to the production frontier and identified as efficient, while values between one and infinity correspond to inefficient universities located below the best practice frontier. Within the economic growth model estimation we prefer to work with the so-called Shephard measures which are simply the inverse of the Farrell ones (Bogetoft and Otto, 2010). The Shephard concept satisfies the condition , with unity representing an efficiency observations, and implies that values between zero and one indicate an inefficient observations. Change from Farrell efficiency measure to Shephard measure is needed to remain commensurate the coefficients of DEA and SFA scores in the models of economic growth. The distributions of the DEA efficiency scores obtained on the institutional level are presented on the Figure 3.

**[Insert figure 3 here]**

From Figure 3 and Table 3 it can be noted that efficiency values for each considered year are above 0.67 fluctuating around 0.74. Distributions of efficiency scores have normal form and are characterized by rather higher standard deviation; therefore, we can note that data envelopment analysis efficiency score discriminate universities in the sample rather well. The key descriptive statistics of efficiency estimates on institutional and regional levels are presented in the Table 3; the average efficiency scores for different regions are also graphically depicted on the Figure 4.

**[Insert table 3 and figure 4 here]**

Standard deviations presented in Table 3 shows that after aggregation on the regional level the standard deviations of our efficiency scores become lower, however data envelopment analysis efficiency scores still discriminate universities in the sample rather well in terms of their efficiency. Both at regional and institutional levels efficiency of higher education in Russia was growing in the period from 2012 to 2015. The average annual growth rate of efficiency at institutional level was 3 percent and at the regional level was 1 percent.

 Table 3 demonstrates also that the standard deviation of efficiency scores, that characterizes the level of heterogeneity of regional higher education system in terms of efficiency, is not stable over time. We can assume here that heterogeneity or homogeneity of regional higher education system in terms of efficiency of particular universities may also affect the rates of regional economic development, so we include also the standard deviation of within-region efficiency score into the model for explaining economic growth, in addition to the average regional efficiency.

*6.2. Economic growth model estimation*

Economic growth model was estimated using software Stata 13 and the package xtabond2 (Roodman, 2006). Four different specifications were considered, and the results are reported in Table 6.

 The first, baseline model (model 1) includes standard variables that are usually used in economic growth modelling (Kufenko, 2015). These variables are the rate of investment growth; rate of population growth; share of employed population with higher education as a measure of accumulated human capital in the region; logarithm of gross regional product in previous period in order to capture convergence effects; variables that reflect the structure of regional economy – share of public sector in GRP, share of commercial minerals extraction in GRP, share of industries in GRP. The parameters of this model, except the share of industries, are statistically significant and the signs of estimated parameters corresponds to underlying theoretical assumptions: gross regional product growth rate is positively related to employed population growth rate, growth rate in previous period, share of employed population with higher education. As expected, GRP growth rate is negatively related with total amount of GRP in previous period confirming the existence of convergence in growth rates: poor regions tend to growth faster than the rich ones.

 The second model (model 2) contains additional explanatory variable that is DEA efficiency score of regional higher education systems. This model confirms the main hypothesis of this study and demonstrates that regional higher education system efficiency is an important determinant of regional economic development growth rates. Positive relationship between GRP growth and regional higher education system efficiency is stable and can be observed in all subsequent specifications of the model (models 3 and 4). Also, we have implemented special robustness check in order to obtain additional evidence that this positive relationship exists. We used stochastic frontier efficiency score instead of data envelopment analysis efficiency score in economic growth model. The detailed description of SFA efficiency estimation and results of regional economic growth model with SFA efficiency scores are presented in the Annex 2.

The Model 3 additionally contains the indicator that measures the standard deviation of efficiency scores on institutional level within the region. The inclusion of this variable in the model allows answering the question about what is better for the region: to have universities with approximately the same level of efficiency or to have very different universities in terms of their efficiency. The results presented in the Table 4 suggests that estimated parameter of this variable is positive and statistically significant meaning that more heterogeneous in terms of efficiency regional higher education system is better for regional economic development. Such result suggests that it is better to have heterogeneous higher education system with efficient and inefficient universities for regional development. The regions with high differentiation of universities in terms of their efficiency levels on average demonstrate higher rates of economic development. Possible explanation for this funding is that universities that are different in terms of efficiency contribute to different aspects of regional development. Leading universities contribute to innovation development, while “mass higher education” institutions prepare workers for local enterprises. However, standard deviation of efficiency scores may reflect the size of regional higher education system (standard deviation is higher for large higher education systems) and the revealed dependence may partially reflect the dependence between the size of higher education system and rates of regional development.

**[Insert table 4 here]**

 Finally, Model 4 aims at checking the hypothesis about spillover effects. Model 4 shows that parameter of efficiency spatial interaction is statistically significant and has negative sign meaning that we have negative spillover effects. Such finding may be explained by the fact that regional higher education systems in Russia tend to compete with each other. Each university tries to attract the brightest students and the best academic staff from the regions in the neighborhood. If in one region regional higher education system is very strong and efficiency and in the neighboring regions is not, this efficient higher education system will extract resources, first of all human resources, from the other regions. That is why if the region A is located near the regions with efficient higher education systems, the competitiveness of the region A in terms of competition with other regions will be relatively lower. The parameter of economic growth spatial interaction, as expected, is positive and statistically significant. As was mentioned, existing evidence suggests that spatial correlation in regional economic growth rates in Russia is positive, so our results are in line with the findings of previous study in the field of economic growth modelling.

**7. Concluding remarks**

Universities are multi-product organizations (Baumol et al., 1982), which determines the multiplicity of impact channels. Usually three different types of universities’ contribution to economic development are highlighted – general economic approach, suggesting that universities as economic agents generate additional aggregate demand in regional economy (Elliot et al., 1988); skill-based approach that analyses higher education contribution in terms of human capital reproduction (Bluestone, 1993); innovation approach that considers universities as an integrators of regional innovation ecosystems.

There is an evidence that the economic impact of universities in Russia in terms of highlighted approaches is positively related to the scale of higher education system (Egorov et al., 2017). In other words, the quantity of higher education matter for regional economic development. This study is a first attempt to shed the light on the question whether the efficiency level of regional higher education systems is positively related to regional economic growth rates in Russia. Using the framework which considers an efficiency level as a good instrument to capture the impact of universities on the community that was suggested in (Agasisti et al., 2017) and (Barra and Zotti, 2016) , we estimate higher education efficiency on the institutional level, then aggregate these estimates in order to obtain aggregate measure of regional higher education system efficiency and, finally, construct regional economic growth model that treats regional higher education system efficiency as one of the explanatory variables. In order to evaluate the efficiency we assume that there are different exogenous factors that are out of universities’ managers control and employ 2-stage DEA procedure proposed in (Simar and Wislon, 2007). For causal inference we employ sys-GMM approach for economic growth models identification. Also, we employ spatial econometrics techniques in order to analyze spillover effects (positive economic impact of universities on the neighboring regions). In order to check the stability of the results we implement special robustness check.

Estimated economic growth models shows that data envelopment analysis and stochastic frontier efficiency scores corrected for exogenous factors are statistically significantly related to gross regional product growth rates. We did not find any spillover effects that were assumed initially. Moreover, we find statistically significant and negative spillover effects. The explanation behind this finding is that strong and efficient regional higher education systems may extract resources (predominantly human resources) from the regions located in the neighborhood. Particularly, such regional higher education systems are more attractive for students and scholars from other regions. Such findings suggest that universities can ensure competitiveness of the regions where they are located in relation to other regions. In such highly centralized higher education system as Russian one (approximately 90% of all state-owned universities are governed by federal authorities), this fact can be considered as significant incentive for regional authorities to collaborate with higher education sector more actively.

The main policy implication of this study is that public policy in higher education sector has to be concentrated not only on the quantity of higher education and its availability, but on the efficiency level as well. At the same time public policy should take into account another side of this mechanism – since efficient higher education systems extract resources from the neighbors, the regions that have weak higher education systems may be damaged. In order to overcome this constraint, it is necessary to develop a network of strong regional universities. Currently Russian higher education system is characterized by the high level of geographical concentration of strong higher education institutions. The first steps towards more even distribution of high-quality universities were made within the framework of “Flagship universities program” developed by Russian Government in 2016.

**References**

Abankina, I., Aleskerov, F., Belousova, V., Zinkovsky, K., & Petrushchenko, S. (2013). Evaluating Performance of Universities Using Data Envelopment Analysis. *Educational studies*, (2), 15-48.

Afonso, A., & Aubyn, M. S. (2006). Cross-country efficiency of secondary education provision: A semi-parametric analysis with non-discretionary inputs. *Economic modelling*, *23*(3), 476-491.

Agasisti, T. (2009). Market forces and competition in university systems: theoretical reflections and empirical evidence from Italy. *International Review of Applied Economics*, 23(4), 463-483.

Agasisti, T., & Johnes, G. (2009). Beyond frontiers: comparing the efficiency of higher education decision‐making units across more than one country. *Education Economics*, *17*(1), 59-79.

Agasisti, T., & Perez-Esparrells, C. (2010). Comparing efficiency in a cross-country perspective: the case of Italian and Spanish state universities. *Higher Education*, *59*(1), 85-103.

Agasisti, T., & Pohl, C. (2012). Comparing German and Italian public universities: Convergence or divergence in the higher education landscape? *Managerial and Decision Economics*, *33*(2), 71-85.

Agasisti, T., Barra, C., & Zotti, R. (2017). *Research, knowledge transfer and innovation: the effect of Italian universities’ efficiency on the local economic development 2006-2012*.

Aghion, P., Blundell, R., Griffith, R., Howitt, P., & Prantl, S. (2009). The effects of entry on incumbent innovation and productivity. *The Review of Economics and Statistics*, *91*(1), 20-32.

Aigner, D., Lovell, C. K., & Schmidt, P. (1977). Formulation and estimation of stochastic frontier production function models. *Journal of econometrics*, *6*(1), 21-37.

Arellano, M., & Bover, O. (1995). Another look at the instrumental variable estimation of error-components models. *Journal of econometrics*, *68*(1), 29-51.

Aubyn, M. S., Garcia, F., & Pais, J. (2009). *Study on the efficiency and effectiveness of public spending on tertiary education* (No. 390). Directorate General Economic and Financial Affairs (DG ECFIN), European Commission.

Barra, C., & Zotti, R. (2017). Investigating the Human Capital Development–growth Nexus: Does the Efficiency of Universities Matter? *International Regional Science Review*, *40*(6), 638-678.

Barro, R. J., & Lee, J. W. (2013). A new data set of educational attainment in the world, 1950–2010. *Journal of development economics*, *104*, 184-198.

Battese, G. E., & Corra, G. S. (1977). Estimation of a production frontier model: with application to the pastoral zone of Eastern Australia. *Australian Journal of Agricultural and Resource Economics*, *21*(3), 169-179.

Baumol, W., Panzar, J., & Willig, R. (1982). Contestable markets and the theory of market structure. *Nueva York, Harcourt Brace Javanovich, Inc*.

Becker, G. (1964). human capital, NY. *GS Becker*.

Bekkers, R., & Freitas, I. M. B. (2008). Analysing knowledge transfer channels between universities and industry: To what degree do sectors also matter? *Research policy*, *37*(10), 1837-1853.

Belenzon, S., & Schankerman, M. (2013). Spreading the word: Geography, policy, and knowledge spillovers. *Review of Economics and Statistics*, *95*(3), 884-903.

Bluestone, B. (1993). *An Economic Impact Analysis*. UMASS/Boston

Bogetoft, P., & Otto, L. (2010). *Benchmarking with Dea, Sfa, and R* (Vol. 157). Springer Science & Business Media.

Bonaccorsi, A., Daraio, C., Lepori, B., & Slipersaeter, S. (2007). Indicators on individual higher education institutions: addressing data problems and comparability issues. *Research Evaluation*, *16*(2), 66-78.

Burgette, L. F., & Reiter, J. P. (2010). Multiple imputation for missing data via sequential regression trees. *American journal of epidemiology*, *172*(9), 1070-1076.

Capello, R. (2011). Location, Regional Growth and Local Development Theories. *Aestimum*, *58*(1), 1-25.

Chirikos, T. N., & Sear, A. M. (2000). Measuring hospital efficiency: a comparison of two approaches. *Health Services Research*, *34*(6), 1389.

Demidova, O. (2015). Spatial effects for the eastern and western regions of Russia: a comparative analysis. *International Journal of Economic Policy in Emerging Economies*, *8*(2), 153-168.

Demidova, O., & Ivanov, D. (2016). Models of Economic Growth with Heterogenous Spatial Effects: The Case of Russian Regions. *HSE Economic Journal*, *20*(1), 52-75.

Egorov, A., Leshukov, O., & Gromov, A. (2017). The Role of Universities in Economic Development of Russian Regions.

Elliott, D. S., Levin, S. L., & Meisel, J. B. (1988). Measuring the economic impact of institutions of higher education. *Research in Higher Education*, *28*(1), 17-33.

Etzkowitz, H. (1993). Enterprises from science: The origins of science-based regional economic development. *Minerva*, *31*(3), 326-360.

Farrell, M. J. (1957). The measurement of productive efficiency. Journal of the Royal Statistical Society. Series A (General), 120(3), 253-290.

Ferrier, G. D., & Lovell, C. K. (1990). Measuring cost efficiency in banking: Econometric and linear programming evidence. *Journal of econometrics*, *46*(1-2), 229-245.

Froumin, I., & Leshukov, O. (2015). National-Regional Relationships in Federal Higher Education Systems: The case of Russian Federation. In *Higher education forum: a COE publication* (Vol. 12, pp. 77-94).

Froumin, I., Kouzminov, Y., & Semyonov, D. (2014). Institutional diversity in Russian higher education: revolutions and evolution. *European Journal of Higher Education*, *4*(3), 209-234.

Carnoy, M., Froumin, I., Leshukov, O. & Marginson., S. (2018) *Higher education in federal countries.* SAGE Publishing.

Gimpelson, V., Kapeliushnikov, R., (2015) The  Russian  Labour  Market  Model:  Trial  by  Recession. *Journal of the New Economic Association, 26*(2), 249-254.

Hanushek, E. A. (2016). Will more higher education improve economic growth? *Oxford Review of Economic Policy*, *32*(4), 538-552.

Hanushek, E. A., & Kimko, D. D. (2000). Schooling, labor-force quality, and the growth of nations. *American economic review*, *90*(5), 1184-1208.

Hanushek, E. A., & Woessmann, L. (2007). *The role of school improvement in economic development* (No. w12832). National Bureau of Economic Research.

Hanushek, E. A., & Woessmann, L. (2015). *The knowledge capital of nations: Education and the economics of growth*. MIT Press.

Huggins, R., & Johnston, A. (2009). The economic and innovation contribution of universities: a regional perspective. *Environment and Planning C: Government and Policy*, *27*(6), 1088-1106.

Johnson, M. S. (2008). Historical legacies of Soviet higher education and the transformation of higher education systems in post-Soviet Russia and Eurasia. In *The worldwide transformation of higher education* (pp. 159-176). Emerald Group Publishing Limited.

Johnes, J. (2006). Data envelopment analysis and its application to the measurement of efficiency in higher education. Economics of Education Review, 25(3), 273-288.

Kosor, M. M. (2013). Efficiency measurement in higher education: Concepts, methods and perspective. *Procedia-Social and Behavioral Sciences*, *106*, 1031-1038.

Kufenko, V. (2015). The resource curse and human capital. In *Economic Growth and Inequality* (pp. 42-61). Springer Gabler, Wiesbaden.

Laredo, P. (2007). Revisiting the third mission of universities: Toward a renewed categorization of university activities?. *Higher education policy*, *20*(4), 441-456.

LeSage, J. P., & Pace, R. K. (2014). The biggest myth in spatial econometrics. *Econometrics*, *2*(4), 217-249.

Leshukov, O., Platonova, D., & Semyonov, D. (2015). Does Competition Matter? The Efficiency of Regional Higher Education Systems and Competition: The Case of Russia.

Loh, W. Y. (2011). Classification and regression trees. Wiley Interdisciplinary Reviews: *Data Mining and Knowledge Discovery*, 1(1), 14-23.

Lucas Jr, R. E. (1988). On the mechanics of economic development. *Journal of monetary economics*, *22*(1), 3-42.

Mankiw, N. G., Romer, D., & Weil, D. N. (1992). A contribution to the empirics of economic growth. *The quarterly journal of economics*, *107*(2), 407-437.

McMillan, M. L., & Chan, W. H. (2006). University efficiency: A comparison and consolidation of results from stochastic and non‐stochastic methods. *Education economics*, *14*(1), 1-30.

Meeusen, W., & van Den Broeck, J. (1977). Efficiency estimation from Cobb-Douglas production functions with composed error. *International economic review*, 435-444.

Mincer, J. A. (1974). Schooling and earnings. In *Schooling, experience, and earnings* (pp. 41-63). NBER.

Morsy, L., Khavenson, T., & Carnoy, M. (2018). How international tests fail to inform policy: The unsolved mystery of Australia’s steady decline in PISA scores. *International Journal of Educational Development*, *60*, 60-79.

Nazarko, J., & Saparauskas, J. (2014). Application of DEA method in efficiency evaluation of public higher education institutions. *Technological and Economic development of Economy*, *20*(1), 25-44.

Parteka, A., & Wolszczak-Derlacz, J. (2013). Dynamics of productivity in higher education: Cross-European evidence based on bootstrapped Malmquist indices. *Journal of Productivity Analysis*, *40*(1), 67-82.

Pinheiro, R., Benneworth, P., & Jones, G. A. (Eds.). (2012). *Universities and regional development: A critical assessment of tensions and contradictions*. Routledge.

Raghunathan, T. E., Solenberger, P. W., & Van Hoewyk, J. (2002). IVEware: Imputation and variance estimation software. Ann Arbor, *MI: Survey Methodology Program, Survey Research Center, Institute for Social Research, University of Michigan.*

Rebelo, S. (1991). Long-run policy analysis and long-run growth. *Journal of political Economy*, *99*(3), 500-521.

Romer, P. M. (1986). Increasing returns and long-run growth. *Journal of political economy*, *94*(5), 1002-1037.

Roodman, D. (2006). How to do xtabond2: An introduction to difference and system GMM in Stata.

Sala-i-Martin, X. (1994). Cross-sectional regressions and the empirics of economic growth. *European Economic Review*, *38*(3-4), 739-747.

Sardadvar, S., & Vakulenko, E. (2016). Interregional migration within Russia and its East-West divide: evidence from spatial panel regression. *Review of Urban & Regional Development Studies*, 28(2), 123-141.

Shephard, R.W., (1953) *Cost and production functions*. Princeton University Press, Princeton.

Simar, L., & Wilson, P. W. (1999). Of course we can bootstrap DEA scores! But does it mean anything? Logic trumps wishful thinking. *Journal of Productivity Analysis*, *11*(1), 93-97.

Simar, L., & Wilson, P. W. (2007). Estimation and inference in two-stage, semi-parametric models of production processes. *Journal of econometrics*, *136*(1), 31-64.

Solow, R. M. (1956). A contribution to the theory of economic growth. *The quarterly journal of economics*, *70*(1), 65-94.

Solow, R. M. (1957). Technical change and the aggregate production function. *The review of Economics and Statistics*, 312-320.

Valero, A., & Van Reenen, J. (2016). *The economic impact of universities: Evidence from across the globe* (No. w22501). National Bureau of Economic Research.

Vandenbussche, J., Aghion, P., & Meghir, C. (2006). Growth, distance to frontier and composition of human capital. *Journal of economic growth*, *11*(2), 97-127.

Varga, A. (1997). Regional economic effects of university research: a survey. *Unpublished manuscript, West Virginia University, Regional Research Institute, Morgantown, WV*.

Windmeijer, F. (2005). A finite sample correction for the variance of linear efficient two-step GMM estimators. *Journal of econometrics*, *126*(1), 25-51.

Wolszczak-Derlacz, J., & Parteka, A. (2011). Efficiency of European public higher education institutions: a two-stage multicountry approach. *Scientometrics*, *89*(3), 887-917.

Wooldridge, J. (2002). Econometric Analysis of Cross-Section and Panel Data. *MIT Press*, pp. 130, 279, 420-449.

**Annex. Figures and tables**

***Figure 1****.* Number of organization of Higher Education in Russian Federation.

*Source:* Russian Federation Federal State Statistics (<http://www.gks.ru/free_doc/new_site/population/obraz/vp-obr1.htm>)



*Notes:* Plot presents the averaged values of efficiency scores of regional HE systems and GRP growth rates over 2012–2015. Regional HE system is a set of universities located within the administrative borders of the region.

***Figure 2.***Scatter plot ofDEA efficiency scores and GRP growth rates (average values for the period 2012–2015).

*Source:* Authors’ calculations from «Annual Monitoring of Efficiency of Higher Education Institutions» and Russian Federation Federal State Statistics.

­­ 

*Notes:* Efficiency scores are distributed between 0 and 1. A value of 1 indicates that a university is efficient and lies on the best practice frontier.

***Figure 3****.* DEA efficiency distribution on institutional level.

*Source:* Authors’ calculations from «Annual Monitoring of Efficiency of Higher Education Institutions».



*Notes:* Plot presents the averaged efficiency scores of Regional Higher Education (HE) systems over 2012–2015. Regional HE system is a set of universities located within the administrative borders of the region. A value NA indicates the lack of the data for this region. The sample includes 78 regions.

***Figure 4.***DEA efficiency scores of Regional Higher Education (HE) systems.

*Source:* Authors’ calculations from «Annual Monitoring of Efficiency of Higher Education Institutions» data.

**Table 1.**

 **Descriptive statistics of the variables used for efficiency evaluation.**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variables | **2012**  | **2013**  | **2014** | **2015**  |
| (N=449) | (N=449) | (N=449) | (N=449) |
| Mean | Std dev | Mean | Std dev | Mean | Std dev | Mean | Std dev |
| **Inputs** |
| The income of educational organization from all sources per number of Faculty members, thousand rubles | 2832 | 1448 | 2892 | 1333 | 3097 | 1395 | 2577 | 1180 |
|  |  |  |  |  |  |  |  |  |
| The relative weight of academic staff with advanced degrees, % | 64 | 16 | 66 | 15 | 69 | 15 | 70 | 15 |
|  |  |  |  |  |  |  |  |  |
| The average entrance exam score of entering students | 65 | 9 | 68 | 10 | 65 | 10 | 64 | 8 |
| **Outputs** |
| The number of publications in science journals indexed in RSCI, Web of Science and Scopus, per capita of academic staff  | 78 | 49 | 84 | 51 | 128 | 77 | 184 | 119 |
|  |  |  |  |  |  |  |  |  |
| The total quantity of R&D per one employee of academic staff, thousand rubles | 236 | 327 | 238 | 333 | 214 | 281 | 176 | 228 |
|  |  |  |  |  |  |  |  |  |
| Students to academic staff ratio | 9 | 3 | 10 | 3 | 10 | 3 | 10 | 2 |
| **Environmental variables** |
| Share of masters’ students in total number of students, % | 4.4 | 8.2 | 5.0 | 8.2 | 5.3 | 7.5 | 7.3 | 7.9 |
|  |  |  |  |  |  |  |  |  |
| Share of full-time students in total number of students, % | 62.1 | 18.1 | 62.6 | 18.4 | 63.9 | 18.3 | 61.0 | 22.7 |
|  |  |  |  |  |  |  |  |  |
| Location of university in the capital city of the region (dummy), % | 91 | - | 91 | - | 91 | - | 91 | - |
|  |  |  |  |  |  |  |  |  |
| Market share of university – share of students in the university in total number of students in the region, % | 10.6 | 13.2 | 11.0 | 14.0 | 14.2 | 11.9 | 12.3 | 15.3 |
|  |  |  |  |  |  |  |  |  |
| Presence of medical faculty (dummy), % | 11 | - | 11 | - | 11 | - | 11 | - |

*Notes:* Information on incomes is adjusted to the level of December 2015 by using the annual national CPI. Data in the case of missing values an imputation procedure based on classification and regression trees algorithm was employed (Loh at al.,2011).

*Source:* Authors’ calculations from «Annual Monitoring of Efficiency of Higher Education Institutions» data.

**Table 2.**

**Descriptive statistics of the variables used for economic growth model.**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Variables** | **2012**  | **2013**  | **2014** | **2015**  |
| (N=77) | (N=77) | (N=77) | (N=77) |
| Mean | Std dev | Mean | Std dev | Mean | Std dev | Mean | Std dev |
| GRP growth rates (year to year), % | 3.41 | 5.77 | 2.07 | 5.68 | 1.95 | 4.74 | -0.20 | 2.52 |
|  |  |  |  |  |  |  |  |  |
| Efficiency measured by DEA | 0.71 | 0.08 | 0.75 | 0.07 | 0.76 | 0.08 | 0.76 | 0.07 |
|  |  |  |  |  |  |  |  |  |
| Population growth rates, % | - 0.09 | 0.01 | - 0.08 | 0.01 | - 0.01 | 0.01 | - 0.07 | 0.01 |
|  |  |  |  |  |  |  |  |  |
| Investment growth rates, % | 2.59 | 13.32 | - 3.01 | 14.27 | - 1.88 | 17.73 | - 9.89 | 13.72 |
|  |  |  |  |  |  |  |  |  |
| Gross regional product in previous period, bln rub. | 910.8 | 190.1 | 927.3 | 196.3 | 949.7 | 198.9 | 827.9 | 169.4 |
|  |  |  |  |  |  |  |  |  |
| Share of employed with higher education,% | 27.85 | 4.89 | 29.14 | 5.01 | 29.71 | 4.78 | 30.65 | 4.61 |
|  |  |  |  |  |  |  |  |  |
| Share of public sector in GRP,% | 17.87 | 7.46 | 18.99 | 7.68 | 18.53 | 7.39 | 17.12 | 7.19 |
|  |  |  |  |  |  |  |  |  |
| Share of commercial minerals extraction in GRP, % | 8.13 | 12.91 | 7.95 | 12.82 | 7.75 | 13.03 | 8.36 | 13.43 |
|  |  |  |  |  |  |  |  |  |
| Share of industries in GRP, % | 17.89 | 9.94 | 17.43 | 9.34 | 17.48 | 9.92 | 18.02 | 10.21 |

*Notes:* Information on incomes is adjusted to the level of December 2015 by using the annual national CPI.

*Source:* Authors’ calculations from Russian Federation Federal State Statistics Service.

**Table 3.**

**Descriptive statistics of DEA efficiency scores over 2012–2015.**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Statistics** | **2012**  | **2013**  | **2014** | **2015**  |
| (N=449) | (N=449) | (N=449) | (N=449) |
| **Institutional level** |
| Mean  | 0.665 | 0.722 | 0.739 | 0.741 |
| Median | 0.685 | 0.756 | 0.767 | 0.756 |
| Std deviation | 0.173 | 0.156 | 0.152 | 0.143 |
| Minimum | 0.185 | 0.259 | 0.275 | 0.262 |
| Maximum | 0.967 | 0.981 | 0.970 | 0.966 |
| **Statistics** | **2012**  | **2013**  | **2014** | **2015**  |
| (N=77) | (N=77) | (N=77) | (N=77) |
| **Regional level** |
| Mean  | 0.713 | 0.751 | 0.763 | 0.759 |
| Median | 0.725 | 0.758 | 0.771 | 0.762 |
| Std deviation | 0.077 | 0.070 | 0.078 | 0.070 |
| Minimum | 0.435 | 0.475 | 0.489 | 0.583 |
| Maximum | 0.897 | 0.900 | 0.912 | 0.882 |

*Notes:* Regional efficiency scores conform to estimates of regional HE systems. Regional HE system is a set of universities located within the administrative borders of the region.

*Source:* Authors’ calculations from «Annual Monitoring of Efficiency of Higher Education Institutions» data.

**Table 4.**

**Results of regional growth model estimation (standard errors are presented in the brackets)**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Variables** | **Model1** | **Model2** | **Model3** | **Model4** |
| Growth rate in previous period | 0.035\*(0.017) | 0.217\*\*\*(0.021) | 0.185\*\*\*(0.032) | 0.058\* (0.028) |
| Investment growth rate | 4.311\*\*(1.475) | 2.487\*\*\*(0.529) | 3.232\*\*\*(0.511) | 4.334\*\*\*(0.863) |
| Employed population growth rate | 1.282\*\*\*(0.173) | 0.278\*\*\*(0.043) | 0.356\*\*\*(0.082) | 0.276\*\*\*(0.062) |
| Gross regional product in previous period (log) | -26.031\*\*\*(1.857) | -4.272\*\*\*(0.601) | -6.885\*\*\*(0.911) |  -4.725\*\*\*(0.342) |
| Share of employed with HE | 0.372\*\* (0.126) | 0.202\* (0.078) | 0.158\*(0.087) | 0.320\*\*(0.111) |
| Share of commercial minerals extraction in GRP | 0.318\*\*(0.094) | 0.077 (0.080) | 0.105(0.321) | 0.043(0.053) |
| Share of industries in GRP  | 0.060(0.109) | 0.011(0.043) | 0.075\*(0.033) | 0.038 (0.041) |
| Share of public sector in GRP | -1.150\*\*\*(0.218) | -0.389\*\*\*(0.085) | -0.489\*\*\*(0.109) | -0.455\*\*\*(0.089) |
| Efficiency DEA |  | 8.283\*(4.493) | 9.062\*\*(3.332) | 12.272\*\*\*(1.989) |
| Efficiency (DEA)standard deviation |  |  | 23.883\*\*\*(5.824) | 15.249\*\*\*(4.381) |
| Efficiency spatial interaction |  |  |  | -0.041\*(0.019) |
| Growth spatial interaction  |  |  |  | 7.831\*\*\*(0.648) |
| Constant | 279.953\*\*\*(19.151) | 45.882\*\*(12.392) | 59.824\*\*\*(11.713) | 42.257\*\*(11.342) |
| Hansen-Sargan | 0.832 | 0.852 | 0.919 | 0.842 |
| AR(2) | 0.384 | 0. 389 | 0.215 | 0.639 |
| # of observations | 308 | 308 | 308 | 308 |

*Notes:* All equations are estimated through a two-step system generalized method moment estimator with Windmeijer (2005) corrected standard error (in brackets).

*Source:* Authors’ calculations from «Annual Monitoring of Efficiency of Higher Education Institutions» and Russian Federation Federal State Statistics.

**Annex 1. Data envelopment analysis calculations details**

For DEA efficiency evaluation we used double-bootstrap procedure proposed in (Simar & Wilson, 2007).

 The idea of the algorithm is as follows:

1. Compute for all universities from the sample by solving equation (7).
2. Estimate equation (8) by using the truncated regression model of on to obtain regression parameters and residuals variance .
3. For each repeat the following four steps (3.1–3.4) times to obtain sets of bootstrap estimates :
	1. For each draw from the truncated normal distribution with left-truncation at .
	2. For each compute .
	3. For each correct the outputs vector for the ration of efficiency score without environmental variables to efficiency score with environmental variables like the follows: .
	4. Using the corrected outputs vector for each estimate new efficiency scores by solving equation (7).
4. For each compute the bias-corrected efficiency scores using the first bootstrap estimates in and the efficiency scores without environmental variables as follows: where .
5. Use the method of maximum likelihood to estimate the truncated regression model of on to obtain regression parameters and residuals variance .
6. For each repeat the next three steps (6.1–6.3) times to yield sets of bootstrap estimates :
	1. For each draw from the truncated normal distribution with left-truncation at .
	2. For each compute .
	3. Use the maximum likelihood method to estimate the truncated regression of on obtaining estimates and .
7. Use the bootstrap values and the estimates , from the 5 step to construct estimated confidence intervals for each element of and for as described below.

Full description of the algorithm and underlying assumption can be found in original paper (Simar & Wilson, 2007).

The results of the second-stage regression are presented in the Table 1. The confidence intervals for most environmental variables do not contain zero, so this factors can be considered as statistically significant related to universities’ efficiency scores. However, the significance of effect of most determinants has decreased over time, and even has changed the sign for some parameters. Such results can be explained by the economic crises that occurred in Russia during the reporting years. Economic turbulences neutralize some dependencies that existed in 2012.

Table 2 contains intermediate results of the Simar and Wilson algorithm and final efficiency scores. All these results, including confidence intervals values, were aggregated to the regional level.

**Table 1.**

**Results of the second-stage truncated regression in double-bootstrap DEA procedure.**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variables | 2012(N=449) | 2013(N=449) | 2014(N=449) | 2015(N=449) |
|  | 95% Conf. In. |  | 95% Conf. In |  | 95% Conf. In |  | 95% Conf. In |
| Constant | -0.447 | [-1.70, 0.53] | -0.101 | [-0.92, 0.54] | -0.035 | [-0.84, 0.59] | 0.654 | [0.13, 1.08] |
|  |  |  |  |  |  |  |  |  |
| Share of masters’ students in total number of students | -0.083 | [-0.14, -0.04] | -0.014 | [-0.04, 0.01] | -0.012 | [-0.04, 0.01] | -0.015 | [-0.03, 0.00] |
|  |  |  |  |  |  |  |  |  |
| Share of full-time students in total number of students | 2.055 | [1.05, 3.13] | 1.3 | [0.69, 2.02] | 1.368 | [0.73, 2.07] | 0.788 | [0.28, 1.33] |
|  |  |  |  |  |  |  |  |  |
| Location of university in the capital city of the region (dummy) | 0.583 | [-0.03, 1.37] | 0.375 | [-0.03, 0.90] | 0.079 | [-0.26, 0.53] | -0.037 | [-0.34, 0.27] |
|  |  |  |  |  |  |  |  |  |
| Market share of university – share of students in the university in total number of students in the region | -6.162 | [-9.39, -3.76] | -2.408 | [-3.81, -1.16] | -1.134 | [-2.32, -0.13] | -1.1 | [-2.04, -0.30] |
|  |  |  |  |  |  |  |  |  |
| Presence of medical faculty (dummy) | 0.603 | [0.15, 1.05] | 0.778 | [0.47, 1.12] | 0.838 | [0.54, 1.18] | 0.654 | [0.39, 0.92] |

*Notes:* Confidence intervals are estimated by using bootstrap procedure (N = 2000).

*Source:* Authors’ calculations from «Annual Monitoring of Efficiency of Higher Education Institutions» data.

**Table 2.**

**Intermediate results of double-bootstrap DEA procedure and final efficiency scores on the regional level.**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |   | Lower bound | Upper bound | Mean over 2012-15 |
| id\_region | 2012 | 2013 | 2014 | 2015 | 2012 | 2013 | 2014 | 2015 | 2012 | 2013 | 2014 | 2015 | 2012 | 2013 | 2014 | 2015 |  |  |
| 1 | 1.43 | 1.32 | 1.24 | 1.25 | 1.49 | 1.37 | 1.28 | 1.29 | 1.45 | 1.33 | 1.25 | 1.26 | 1.53 | 1.41 | 1.31 | 1.31 | 1.31 | 1.36 |
| 2 | 1.47 | 1.26 | 1.32 | 1.28 | 1.51 | 1.30 | 1.37 | 1.32 | 1.48 | 1.27 | 1.33 | 1.29 | 1.54 | 1.33 | 1.41 | 1.35 | 1.33 | 1.38 |
| 3 | 2.18 | 2.00 | 1.97 | 1.47 | 2.31 | 2.11 | 2.04 | 1.54 | 2.23 | 2.03 | 1.99 | 1.49 | 2.42 | 2.19 | 2.10 | 1.60 | 1.90 | 2.00 |
| 4 | 1.26 | 1.27 | 1.28 | 1.21 | 1.33 | 1.32 | 1.32 | 1.25 | 1.27 | 1.29 | 1.29 | 1.21 | 1.39 | 1.36 | 1.35 | 1.28 | 1.25 | 1.30 |
| 5 | 1.07 | 1.11 | 1.09 | 1.07 | 1.14 | 1.16 | 1.17 | 1.14 | 1.08 | 1.12 | 1.10 | 1.09 | 1.19 | 1.21 | 1.24 | 1.20 | 1.08 | 1.15 |
| 6 | 1.19 | 1.15 | 1.07 | 1.20 | 1.24 | 1.20 | 1.11 | 1.23 | 1.20 | 1.16 | 1.07 | 1.20 | 1.28 | 1.24 | 1.13 | 1.26 | 1.15 | 1.20 |
| 7 | 1.29 | 1.25 | 1.24 | 1.29 | 1.34 | 1.29 | 1.30 | 1.34 | 1.30 | 1.26 | 1.26 | 1.31 | 1.39 | 1.33 | 1.35 | 1.39 | 1.27 | 1.32 |
| 8 | 1.23 | 1.24 | 1.25 | 1.45 | 1.34 | 1.33 | 1.32 | 1.52 | 1.25 | 1.26 | 1.26 | 1.46 | 1.42 | 1.40 | 1.38 | 1.58 | 1.29 | 1.38 |
| 9 | 1.26 | 1.11 | 1.19 | 1.23 | 1.29 | 1.14 | 1.22 | 1.27 | 1.26 | 1.12 | 1.20 | 1.24 | 1.30 | 1.17 | 1.24 | 1.29 | 1.20 | 1.23 |
| 10 | 1.31 | 1.21 | 1.23 | 1.26 | 1.38 | 1.29 | 1.29 | 1.32 | 1.32 | 1.23 | 1.25 | 1.27 | 1.42 | 1.35 | 1.35 | 1.36 | 1.25 | 1.32 |
| 11 | 1.57 | 1.43 | 1.42 | 1.35 | 1.68 | 1.53 | 1.51 | 1.42 | 1.59 | 1.46 | 1.44 | 1.36 | 1.77 | 1.61 | 1.58 | 1.48 | 1.44 | 1.54 |
| 12 | 1.68 | 1.50 | 1.51 | 1.50 | 1.80 | 1.62 | 1.61 | 1.58 | 1.72 | 1.54 | 1.53 | 1.52 | 1.89 | 1.71 | 1.68 | 1.64 | 1.55 | 1.65 |
| 14 | 1.38 | 1.34 | 1.41 | 1.63 | 1.41 | 1.39 | 1.45 | 1.70 | 1.38 | 1.36 | 1.42 | 1.65 | 1.44 | 1.42 | 1.48 | 1.76 | 1.44 | 1.49 |
| 15 | 1.41 | 1.31 | 1.30 | 1.31 | 1.50 | 1.40 | 1.37 | 1.37 | 1.44 | 1.33 | 1.31 | 1.32 | 1.56 | 1.47 | 1.42 | 1.42 | 1.33 | 1.41 |
| 16 | 1.15 | 1.23 | 1.32 | 1.41 | 1.21 | 1.26 | 1.36 | 1.46 | 1.15 | 1.23 | 1.32 | 1.42 | 1.27 | 1.28 | 1.40 | 1.49 | 1.28 | 1.32 |
| 17 | 1.23 | 1.20 | 1.21 | 1.17 | 1.32 | 1.28 | 1.28 | 1.23 | 1.26 | 1.23 | 1.23 | 1.19 | 1.38 | 1.33 | 1.34 | 1.27 | 1.21 | 1.28 |
| 18 | 1.59 | 1.42 | 1.35 | 1.43 | 1.66 | 1.48 | 1.41 | 1.50 | 1.60 | 1.43 | 1.36 | 1.45 | 1.72 | 1.53 | 1.47 | 1.55 | 1.45 | 1.51 |
| 19 | 1.44 | 1.41 | 1.45 | 1.49 | 1.52 | 1.48 | 1.50 | 1.54 | 1.46 | 1.43 | 1.47 | 1.50 | 1.60 | 1.53 | 1.54 | 1.58 | 1.45 | 1.51 |
| 20 | 1.52 | 1.38 | 1.19 | 1.12 | 1.57 | 1.42 | 1.21 | 1.14 | 1.53 | 1.39 | 1.19 | 1.12 | 1.61 | 1.45 | 1.23 | 1.16 | 1.30 | 1.34 |
| 21 | 1.37 | 1.44 | 1.34 | 1.46 | 1.45 | 1.51 | 1.40 | 1.51 | 1.39 | 1.46 | 1.36 | 1.48 | 1.52 | 1.56 | 1.46 | 1.54 | 1.40 | 1.47 |
| 22 | 1.50 | 1.39 | 1.25 | 1.17 | 1.67 | 1.42 | 1.28 | 1.19 | 1.47 | 1.39 | 1.25 | 1.16 | 1.82 | 1.44 | 1.30 | 1.20 | 1.33 | 1.39 |
| 23 | 1.46 | 1.33 | 1.33 | 1.31 | 1.52 | 1.39 | 1.39 | 1.36 | 1.48 | 1.34 | 1.35 | 1.32 | 1.56 | 1.43 | 1.43 | 1.39 | 1.36 | 1.41 |
| 24 | 1.47 | 1.22 | 1.18 | 1.14 | 1.51 | 1.26 | 1.22 | 1.18 | 1.48 | 1.23 | 1.19 | 1.15 | 1.55 | 1.28 | 1.25 | 1.22 | 1.25 | 1.29 |
| 25 | 1.24 | 1.16 | 1.20 | 1.35 | 1.28 | 1.20 | 1.24 | 1.38 | 1.25 | 1.17 | 1.20 | 1.35 | 1.32 | 1.23 | 1.28 | 1.41 | 1.24 | 1.28 |
| 26 | 1.24 | 1.17 | 1.13 | 1.14 | 1.29 | 1.22 | 1.18 | 1.18 | 1.25 | 1.19 | 1.15 | 1.15 | 1.33 | 1.26 | 1.22 | 1.21 | 1.17 | 1.22 |
| 27 | 1.34 | 1.19 | 1.21 | 1.38 | 1.45 | 1.27 | 1.28 | 1.45 | 1.37 | 1.21 | 1.22 | 1.39 | 1.53 | 1.33 | 1.33 | 1.51 | 1.28 | 1.36 |
| 28 | 1.22 | 1.19 | 1.24 | 1.33 | 1.24 | 1.21 | 1.27 | 1.36 | 1.22 | 1.19 | 1.24 | 1.33 | 1.26 | 1.23 | 1.29 | 1.38 | 1.25 | 1.27 |
| 29 | 1.33 | 1.36 | 1.28 | 1.19 | 1.42 | 1.46 | 1.34 | 1.23 | 1.35 | 1.40 | 1.30 | 1.19 | 1.50 | 1.55 | 1.38 | 1.26 | 1.29 | 1.36 |
| 30 | 1.30 | 1.22 | 1.11 | 1.16 | 1.36 | 1.27 | 1.16 | 1.22 | 1.31 | 1.22 | 1.11 | 1.16 | 1.40 | 1.31 | 1.19 | 1.27 | 1.20 | 1.25 |
| 31 | 1.38 | 1.41 | 1.57 | 1.65 | 1.45 | 1.46 | 1.63 | 1.69 | 1.39 | 1.42 | 1.58 | 1.66 | 1.51 | 1.50 | 1.67 | 1.72 | 1.50 | 1.56 |
| 32 | 1.48 | 1.26 | 1.26 | 1.26 | 1.57 | 1.35 | 1.34 | 1.33 | 1.49 | 1.28 | 1.28 | 1.28 | 1.64 | 1.43 | 1.40 | 1.39 | 1.32 | 1.40 |
| 33 | 1.61 | 1.56 | 1.40 | 1.31 | 1.67 | 1.63 | 1.45 | 1.36 | 1.62 | 1.59 | 1.41 | 1.33 | 1.71 | 1.69 | 1.49 | 1.39 | 1.47 | 1.53 |
| 34 | 1.35 | 1.33 | 1.34 | 1.31 | 1.43 | 1.40 | 1.43 | 1.37 | 1.36 | 1.34 | 1.36 | 1.32 | 1.49 | 1.45 | 1.50 | 1.42 | 1.33 | 1.41 |
| 35 | 1.59 | 1.56 | 1.50 | 1.49 | 1.66 | 1.63 | 1.57 | 1.54 | 1.61 | 1.58 | 1.52 | 1.49 | 1.72 | 1.69 | 1.62 | 1.58 | 1.54 | 1.60 |
| 36 | 1.37 | 1.26 | 1.24 | 1.22 | 1.45 | 1.33 | 1.31 | 1.28 | 1.39 | 1.28 | 1.26 | 1.23 | 1.51 | 1.39 | 1.38 | 1.33 | 1.27 | 1.34 |
| 37 | 1.37 | 1.26 | 1.22 | 1.24 | 1.44 | 1.31 | 1.28 | 1.30 | 1.39 | 1.27 | 1.24 | 1.26 | 1.49 | 1.35 | 1.33 | 1.35 | 1.27 | 1.33 |
| 38 | 1.44 | 1.34 | 1.30 | 1.29 | 1.50 | 1.39 | 1.36 | 1.34 | 1.45 | 1.35 | 1.32 | 1.29 | 1.54 | 1.42 | 1.39 | 1.38 | 1.34 | 1.39 |
| 39 | 1.28 | 1.30 | 1.19 | 1.26 | 1.34 | 1.36 | 1.25 | 1.31 | 1.29 | 1.32 | 1.20 | 1.26 | 1.39 | 1.42 | 1.31 | 1.35 | 1.26 | 1.32 |
| 40 | 1.29 | 1.19 | 1.06 | 1.10 | 1.35 | 1.24 | 1.13 | 1.17 | 1.30 | 1.19 | 1.07 | 1.10 | 1.40 | 1.28 | 1.20 | 1.23 | 1.16 | 1.22 |
| 41 | 1.52 | 1.40 | 1.30 | 1.33 | 1.60 | 1.47 | 1.36 | 1.38 | 1.53 | 1.41 | 1.31 | 1.34 | 1.65 | 1.54 | 1.41 | 1.41 | 1.39 | 1.45 |
| 42 | 1.55 | 1.49 | 1.40 | 1.32 | 1.64 | 1.59 | 1.49 | 1.39 | 1.56 | 1.50 | 1.42 | 1.34 | 1.72 | 1.68 | 1.57 | 1.45 | 1.44 | 1.53 |
| 43 | 1.31 | 1.28 | 1.20 | 1.28 | 1.35 | 1.34 | 1.26 | 1.33 | 1.32 | 1.29 | 1.21 | 1.28 | 1.38 | 1.39 | 1.30 | 1.37 | 1.27 | 1.32 |
| 44 | 1.43 | 1.43 | 1.40 | 1.36 | 1.51 | 1.48 | 1.46 | 1.43 | 1.44 | 1.44 | 1.40 | 1.37 | 1.58 | 1.52 | 1.51 | 1.49 | 1.40 | 1.47 |
| 45 | 1.44 | 1.45 | 1.68 | 1.66 | 1.49 | 1.50 | 1.75 | 1.71 | 1.45 | 1.46 | 1.69 | 1.67 | 1.54 | 1.53 | 1.80 | 1.76 | 1.56 | 1.61 |
| 46 | 1.34 | 1.30 | 1.25 | 1.33 | 1.41 | 1.36 | 1.32 | 1.38 | 1.36 | 1.32 | 1.27 | 1.34 | 1.46 | 1.40 | 1.37 | 1.43 | 1.30 | 1.37 |
| 47 | 1.35 | 1.23 | 1.32 | 1.32 | 1.41 | 1.27 | 1.36 | 1.37 | 1.36 | 1.24 | 1.33 | 1.33 | 1.45 | 1.30 | 1.39 | 1.40 | 1.31 | 1.35 |
| 48 | 1.27 | 1.17 | 1.16 | 1.22 | 1.38 | 1.26 | 1.23 | 1.26 | 1.30 | 1.19 | 1.18 | 1.22 | 1.46 | 1.34 | 1.30 | 1.29 | 1.21 | 1.28 |
| 50 | 1.00 | 1.22 | 1.29 | 1.25 | 1.11 | 1.25 | 1.32 | 1.29 | 1.00 | 1.22 | 1.30 | 1.26 | 1.22 | 1.26 | 1.35 | 1.31 | 1.19 | 1.24 |
| 51 | 1.79 | 1.62 | 1.64 | 1.64 | 1.89 | 1.71 | 1.74 | 1.71 | 1.81 | 1.65 | 1.66 | 1.65 | 1.98 | 1.78 | 1.82 | 1.77 | 1.67 | 1.76 |
| 52 | 1.15 | 1.24 | 1.24 | 1.20 | 1.21 | 1.31 | 1.29 | 1.25 | 1.14 | 1.24 | 1.24 | 1.21 | 1.27 | 1.37 | 1.33 | 1.29 | 1.21 | 1.26 |
| 53 | 1.43 | 1.31 | 1.20 | 1.13 | 1.49 | 1.34 | 1.24 | 1.17 | 1.44 | 1.32 | 1.21 | 1.13 | 1.55 | 1.38 | 1.27 | 1.21 | 1.26 | 1.31 |
| 54 | 1.72 | 1.46 | 1.55 | 1.42 | 1.83 | 1.54 | 1.60 | 1.47 | 1.75 | 1.48 | 1.57 | 1.42 | 1.92 | 1.61 | 1.63 | 1.51 | 1.54 | 1.61 |
| 56 | 1.52 | 1.42 | 1.44 | 1.44 | 1.62 | 1.52 | 1.53 | 1.51 | 1.54 | 1.44 | 1.46 | 1.46 | 1.70 | 1.61 | 1.62 | 1.58 | 1.45 | 1.55 |
| 57 | 1.45 | 1.42 | 1.30 | 1.25 | 1.51 | 1.48 | 1.35 | 1.31 | 1.45 | 1.44 | 1.32 | 1.26 | 1.56 | 1.54 | 1.39 | 1.35 | 1.36 | 1.42 |
| 58 | 1.45 | 1.32 | 1.22 | 1.23 | 1.57 | 1.41 | 1.29 | 1.29 | 1.47 | 1.34 | 1.23 | 1.24 | 1.67 | 1.48 | 1.34 | 1.34 | 1.30 | 1.39 |
| 59 | 1.28 | 1.31 | 1.36 | 1.35 | 1.31 | 1.35 | 1.42 | 1.39 | 1.28 | 1.32 | 1.38 | 1.36 | 1.33 | 1.39 | 1.46 | 1.42 | 1.33 | 1.37 |
| 60 | 1.56 | 1.43 | 1.58 | 1.37 | 1.63 | 1.49 | 1.64 | 1.40 | 1.56 | 1.44 | 1.59 | 1.37 | 1.69 | 1.53 | 1.68 | 1.42 | 1.48 | 1.54 |
| 61 | 1.41 | 1.27 | 1.37 | 1.28 | 1.49 | 1.34 | 1.45 | 1.34 | 1.43 | 1.28 | 1.39 | 1.29 | 1.55 | 1.40 | 1.52 | 1.39 | 1.33 | 1.40 |
| 62 | 1.31 | 1.18 | 1.14 | 1.14 | 1.39 | 1.24 | 1.21 | 1.20 | 1.33 | 1.19 | 1.15 | 1.15 | 1.46 | 1.28 | 1.28 | 1.25 | 1.19 | 1.26 |
| 63 | 1.34 | 1.27 | 1.16 | 1.25 | 1.43 | 1.34 | 1.23 | 1.32 | 1.36 | 1.29 | 1.17 | 1.27 | 1.49 | 1.40 | 1.29 | 1.38 | 1.26 | 1.33 |
| 64 | 1.42 | 1.38 | 1.40 | 1.39 | 1.48 | 1.44 | 1.46 | 1.46 | 1.43 | 1.39 | 1.41 | 1.40 | 1.53 | 1.49 | 1.51 | 1.52 | 1.40 | 1.46 |
| 65 | 1.27 | 1.25 | 1.09 | 1.22 | 1.34 | 1.32 | 1.14 | 1.27 | 1.27 | 1.25 | 1.10 | 1.23 | 1.40 | 1.38 | 1.18 | 1.32 | 1.21 | 1.27 |
| 66 | 1.65 | 1.42 | 1.43 | 1.34 | 1.74 | 1.50 | 1.48 | 1.41 | 1.68 | 1.45 | 1.44 | 1.36 | 1.82 | 1.57 | 1.53 | 1.46 | 1.46 | 1.54 |
| 67 | 1.33 | 1.41 | 1.32 | 1.44 | 1.44 | 1.48 | 1.38 | 1.47 | 1.36 | 1.43 | 1.34 | 1.44 | 1.51 | 1.54 | 1.43 | 1.49 | 1.38 | 1.44 |
| 68 | 1.58 | 1.43 | 1.28 | 1.20 | 1.67 | 1.51 | 1.36 | 1.27 | 1.59 | 1.45 | 1.30 | 1.21 | 1.74 | 1.56 | 1.42 | 1.33 | 1.37 | 1.45 |
| 69 | 1.18 | 1.12 | 1.15 | 1.20 | 1.25 | 1.19 | 1.21 | 1.26 | 1.19 | 1.14 | 1.16 | 1.21 | 1.31 | 1.25 | 1.26 | 1.32 | 1.16 | 1.23 |
| 70 | 1.53 | 1.37 | 1.33 | 1.44 | 1.61 | 1.42 | 1.38 | 1.51 | 1.55 | 1.38 | 1.34 | 1.45 | 1.68 | 1.46 | 1.42 | 1.57 | 1.42 | 1.48 |
| 71 | 1.30 | 1.27 | 1.28 | 1.21 | 1.39 | 1.36 | 1.37 | 1.28 | 1.31 | 1.28 | 1.29 | 1.22 | 1.46 | 1.44 | 1.44 | 1.35 | 1.27 | 1.35 |
| 72 | 1.37 | 1.20 | 1.23 | 1.30 | 1.43 | 1.24 | 1.27 | 1.35 | 1.39 | 1.21 | 1.25 | 1.31 | 1.48 | 1.26 | 1.31 | 1.40 | 1.27 | 1.32 |
| 73 | 1.30 | 1.23 | 1.20 | 1.16 | 1.37 | 1.29 | 1.26 | 1.24 | 1.32 | 1.24 | 1.21 | 1.19 | 1.43 | 1.34 | 1.30 | 1.31 | 1.22 | 1.29 |
| 74 | 1.37 | 1.28 | 1.29 | 1.30 | 1.43 | 1.34 | 1.34 | 1.34 | 1.39 | 1.29 | 1.30 | 1.30 | 1.48 | 1.38 | 1.39 | 1.37 | 1.31 | 1.36 |
| 75 | 1.23 | 1.18 | 1.13 | 1.15 | 1.30 | 1.26 | 1.21 | 1.23 | 1.25 | 1.20 | 1.15 | 1.17 | 1.35 | 1.33 | 1.27 | 1.30 | 1.17 | 1.25 |
| 76 | 1.43 | 1.26 | 1.22 | 1.24 | 1.49 | 1.33 | 1.26 | 1.28 | 1.44 | 1.29 | 1.23 | 1.25 | 1.54 | 1.38 | 1.30 | 1.31 | 1.29 | 1.34 |
| 78 | 1.43 | 1.35 | 1.33 | 1.36 | 1.51 | 1.42 | 1.39 | 1.42 | 1.46 | 1.37 | 1.35 | 1.38 | 1.56 | 1.47 | 1.43 | 1.47 | 1.37 | 1.43 |
| 79 | 1.11 | 1.06 | 1.05 | 1.12 | 1.19 | 1.12 | 1.10 | 1.17 | 1.12 | 1.08 | 1.06 | 1.13 | 1.26 | 1.17 | 1.14 | 1.21 | 1.08 | 1.14 |
| 80 | 1.34 | 1.25 | 1.16 | 1.18 | 1.40 | 1.31 | 1.20 | 1.21 | 1.35 | 1.26 | 1.17 | 1.18 | 1.46 | 1.35 | 1.24 | 1.23 | 1.23 | 1.28 |
| 81 | 1.35 | 1.25 | 1.34 | 1.21 | 1.43 | 1.34 | 1.43 | 1.28 | 1.38 | 1.28 | 1.36 | 1.22 | 1.50 | 1.40 | 1.50 | 1.35 | 1.29 | 1.37 |

*Notes:* Efficiency scores and its’ confidence intervals from the institutional level were aggregated to the regional level. These aggregated values are weighted averages of universities efficiency scores and its’ confidence intervals by the total number of students of local universities in certain regions (for more detailed information see the main body of research).

*Source:* Authors’ calculations from «Annual Monitoring of Efficiency of Higher Education Institutions» data.

**Annex 2. Robustness check with Stochastic frontier analysis**

 Stochastic frontier analysis (SFA) was introduced by Aigner et al. (1977), Meeusen and Van den Broeck (1977), Battese and Corra (1977). This technique is parametric alternative to non-parametric data envelopment analysis. We employ this technique as a robustness check of the results obtained through data envelopment analysis. Using both DEA and SFA techniques we can get reliable information about efficiency ranking (McMillan and Chan, 2006)

In order to implement SFA technique in multi-output case we use input distance function of the following form (Shephard distance function (Shephard, 1953)):

(2) , where:

Farrell input [Farrell, 1957];

technology set;

input vector;

output vector.

We used the variable of income of educational organization from all sources per number of Faculty members as normalizing input – dependent variable in stochastic frontier regression. All inputs and outputs are the same as in case of data envelopment analysis (see Table 1 at the main body of research). In order to estimate stochastic frontier model Cobb-Douglas type production function in logarithmic form was assumed. The results of the SFA regression estimation are presented in the Table 3.

**Table 3.**

**Results of the efficiency estimation on institutional level (standard errors are presented in the brackets)**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dependent variable – the income of educational organization from all sources per number of Faculty members | 2012(N=449) | 2013(N=449) | 2014(N=449) | 2015(N=449) |
| The relative weight of academic staff with advanced degrees,% | -6.452\*\*\*(0.836) | -9.034\*\*\*(0.763) | -11.976\*\*\*(0.761) | -8.102\*\*\*(0.094) |
| The average entrance exam score of entering students | -27.72\*\*\*(1.028) | -25.95\*\*\*(0.847) | -28.78\*\*\*(0.899) | -25.66\*\*\*(1.062) |
| The number of publications in science journals indexed in RSCI, Web of Science and Scopus, per capita of academic staff | 0.039\*\*\*(0.011) | 0.048\*\*\*(0.011) | 0.041\*\*\*(0.010) | -0.016\*\*(0.006) |
| The total quantity of R&D per one employee of academic staff, thousand rubles | 0.022\*\*(0.008) | 0.019\*\*(0.006) | 0.010(0.006) | 0.007(0.006) |
| Students to academic staff ratio | -0.045(0.038) | -0.127\*\*\*(0.030) | -0.122\*\*\*(0.319) | -0.136\*\*\*(0.034) |
| Share of masters’ students in total number of students, % | 0.035\*\*\*(0.010) | 0.034\*\*\*(0.009) | 0.044\*\*\*(0.008) | 0.052\*\*\*(0.010) |
| Share of full-time students in total number of students, % | 0.365\*\*\*(0.092) | 0.371\*\*\*(0.074) | 0.389\*\*\*(0.066) | 0.418\*\*\*(0.068) |
| Location of university in the capital city of the region (dummy), % | 0.049(0.034) | 0.043.(0.026) | 0.009(0.028) | 0.022(0.031) |
| Market share of university – share of students in the university in total number of students in the region, % | 0.222\*(0.097) | 0.073(0.075) | 0.064(0.069) | 0.076(0.070) |
| Presence of medical faculty (dummy), % | 0.309\*\*\*(0.043) | 0.386\*\*\*(0.036) | 0.348\*\*\*(0.041) | 0.268\*\*\*(0.047) |
| Lambda | 4.414 | 4.952 | 3.965 | 3.412 |
| Log likelihood | 224.02 | 315.73 | 339.35 | 326.24 |

*Source:* Authors’ calculations from «Annual Monitoring of Efficiency of Higher Education Institutions» data.

Kernel density plot for the efficiency scores obtained using stochastic frontier analysis is presented on the Fig.1. Distributions of SFA efficiency scores have negative asymmetry and higher kurtosis than for DEA efficiency scores (see Figure 1 at the main body of research).



*Notes:* Efficiency scores are distributed between 0 and 1. A value of 1 indicates that a university is efficient and lies on the best practice frontier.

***Figure 1****.* SFA efficiency distribution on institutional level.

*Source:* Authors’ calculations from «Annual Monitoring of Efficiency of Higher Education Institutions».

The mean SFA efficiency scores have increased by in 4%: from 0.84 in 2012 to 0.87 in 2015. This dynamic of growth correspond to tendency with DEA efficiency estimates. The values of SFA efficiency scores at the institutional level are on average 15 per cent lower than DEA efficiency scores. The SFA efficiency estimates discriminate universities in the sample worse than DEA scores: the standard deviation of SFA scores are on average 40% lower than standard deviation for DEA estimates. The descriptive statistics of SFA efficiency scores at the institutional level are presented in the Table 4.

SFA efficiency scores were also aggregated to the regional level using weighted average formula with total number of students in the university as a weight. The above mentioned conclusions for institutional efficiency scores are relevant too for regional efficiency values, because the last ones are based on the first. The descriptive statistics for SFA efficiency scores at the regional level are shown in Table 4.

**Table 4.**

**Descriptive statistics of SFA efficiency scores over 2012–2015.**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Statistics** | **2012**  | **2013**  | **2014** | **2015**  |
| (N=449) | (N=449) | (N=449) | (N=449) |
| **Institutional level** |
| Mean  | 0.839 | 0.859 | 0.869 | 0.871 |
| Median | 0.864 | 0.889 | 0.893 | 0.893 |
| Std deviation | 0.109 | 0.097 | 0.086 | 0.084 |
| Minimum | 0.226 | 0.513 | 0.552 | 0.490 |
| Maximum | 0.982 | 0.989 | 0.983 | 0.984 |
| **Statistics** | **2012**  | **2013**  | **2014** | **2015**  |
| (N=79) | (N=79) | (N=79) | (N=79) |
| **Regional level** |
| Mean  | 0.841 | 0.862 | 0.875 | 0.877 |
| Median | 0.855 | 0.883 | 0.885 | 0.884 |
| Std deviation | 0.071 | 0.067 | 0.056 | 0.056 |
| Minimum | 0.591 | 0.641 | 0.669 | 0.649 |
| Maximum | 0.932 | 0.965 | 0.961 | 0.952 |

*Notes:* Regional efficiency scores conform to estimates of regional HE systems. Regional HE system is a set of universities located within the administrative borders of the region.

*Source:* Authors’ calculations from «Annual Monitoring of Efficiency of Higher Education Institutions» data.

 The distribution of Regional Higher Education systems based on SFA efficiency scores is represented by Figure 2. This distribution corresponds to the analogous regional distribution of DEA efficiency scores (see Figure 4 at the main body of research). The most efficient Regional HE systems are concentrated in the central part of Russia and less at the western regions. While the most efficient Regional HE systems according to DEA scores are located more at the western part of Russia and less at the central regions.

****

*Notes:* Plot presents the averaged efficiency scores of Regional Higher Education (HE) systems over 2012–2015. Regional HE system is a set of universities located within the administrative borders of the region. A value NA indicates the lack of the data for this region. The sample includes 78 regions.

***Figure 4.***SFA efficiency scores of Regional Higher Education (HE) systems.

*Source:* Authors’ calculations from «Annual Monitoring of Efficiency of Higher Education Institutions» data.

The Pearson correlation coefficient between DEA and SFA regional efficiency scores amounts on average 0.15 that corresponds to a low strength of relationship. This correlation is statistically significant (p<0.05). The result about low strength of relationship between scores estimated by DEA and SFA was achieved in more early literature (for example, Chirikos and Sear (2000), Ferrier and Lovell (1990)). There are numerous causes for the variance of DEA and SFA results: different functional form of production function and true/false assumption on it, presence of statistical noise, different distributions of inefficiency term and true/false assumption on it, correlation of inefficiency scores with explanatory variables, omission of relevant/inclusion of irrelevant variables etc. The analysis of listed causes goes beyond the scope of the present paper. But since the DEA efficiency scores discriminate universities in the sample better than SFA estimates, the economic growth model estimation involves mainly DEA efficiency values and use SFA scores to check the robustness of the results. The estimation results of economic growth with SFA efficiency scores are presented in Table 5.

**Table 5.**

**Results of regional growth model estimation with SFA efficiency scores (standard errors are presented in the brackets)**

|  |  |
| --- | --- |
| Variables  | Estimated parameter |
| Growth rate in previous period | 0.038\*(0.019) |
| Investment growth rate | 4.312\*\*\*(0.729) |
| Employed population growth rate | 0.342\*\*\*0.044 |
| Gross regional product in previous period | -4.879\*\*\*(0.825) |
| Share of employed with HE | 0.338\*(0.142) |
| Share of commercial minerals extraction in GRP | 0.043(0.038) |
| Share of industries in GRP | 0.055\*(0.025) |
| Share of public sector in GRP | -0.527\*\*\*(0.125) |
| Efficiency SFA | 18.245\*\*(6.111) |
| SFA efficiency standard deviation | 51.234\*(26.124) |
| Efficiency spatial interaction | -0.045\*\*\*(0.005) |
| Growth spatial interaction  | 7.865\*\*\*(0.905) |
| Constant | 24.814\*(13.424) |
| Hansen-Sargan | 0.589 |
| AR(2) | 0.842 |
| # of observations | 308 |
| Significance levels: \*\*\* p-value<0.001; \*\* p-value<0.01; \* p-value<0.05; . p-value<0.1 |

*Notes:* Equation is estimated through a two-step system generalized method moment estimator with Windmeijer (2005) corrected standard error (in brackets).

*Source:* Authors’ calculations from «Annual Monitoring of Efficiency of Higher Education Institutions» and Russian Federation Federal State Statistics.

 All relationships that were revealed based on the model with DEA scores (see Table 6 at the main body of research) are preserved in the model with SFA efficiency scores as one of the explanatory variables. This robustness check confirms the conclusions about positive influence of regional higher education system efficiency on economic growth rates, negative spatial effects of efficiency and positive relationship between standard deviation of universities’ efficiency within the region and economic growth rates.

 The scatter plot of SFA efficiency and GRP growth rates averaged for the period under review is presented on the Figure 3. As in the case of DEA, scatter plot shows positive relationship between regional rates of economic growth and regional HE systems efficiency scores. However, the two-dimensional regression line has a smaller slope than in case of DEA.

 In overall, we can conclude that results of the paper are robust to the change of efficiency measure.



*Notes:* Plot presents the averaged values of efficiency scores of regional HE systems and GRP growth rates over 2012–2015. Regional HE system is a set of universities located within the administrative borders of the region.

***Figure 3.***Scatter plot ofSFA efficiency scores and GRP growth rates (average values for the period 2012–2015).

*Source:* Authors’ calculations from «Annual Monitoring of Efficiency of Higher Education Institutions» and Federal State Statistics Service.

**Annex 3. Handling missing data and outliers**

*MICE with CART algorithm for missing data imputation*

The basic idea of the multiple imputations by chained equations (MICE) (Raghunathan et al, 2002) is as follows. Suppose we have a sample of variables that contains missing values. These variables are considered separately starting from . In order to predict missing values in this variable we built a regression for using as predictors:

The valueof predicted by regression (3) is used as a proxy for missing value. This procedure should be repeated for each variable in the sample. So in other words, we have to consider equations of the form (1) and use them for predicting missing values.

 However, sometimes relationships between variables in the sample are interactive and non-linear, thus simple linear regression may give predictions with rather high errors. Moreover, sometimes variables may have difficult distribution that are not captured by linear models (Burgette and Reiter, 2010). That is why it is reasonable to implement alternative models for prediction of missing values, for instance, Classification and Regression Trees algorithm (CART).

 In order to implement MICE based on CART algorithm we have to change equation (3) by the regression tree model that is formulated as follows. Suppose we have outcome variable and matrix containing explanatory variables. Regression tree suggests that the space of predictors should be divided into regions in such way that minimize the sum of squared errors and the fitted value in each . More detailed description of CART algorithm can be found in (Loh, 2011).

*Capping procedure for outliers*

 The capping procedure suggests the following algorithm. Suppose we have variable in our dataset. At the first step we identify outliers according to the formula (4).

 where

;

;

comulative distribution function.

 When outliers are identified their values are substituted using the following rule (5):

 So this procedure replaces the outliers by the tail distribution quantiles, for example by quantiles of the levels 0.975 and 0.025.

1. The terms “third mission” or “knowledge transfers” indicate the diverse channels through which HEIs interact with community and contribute to his economic and social development (Agasisti et al., 2017 ). [↑](#footnote-ref-1)
2. In that framework a Decision Making Unit is a University. [↑](#footnote-ref-2)
3. [www.miccedu.ru](http://www.miccedu.ru) [↑](#footnote-ref-3)
4. These calculations are based on the data from Russian Federation Federal State Statistics ([www.gks.ru](http://www.gks.ru)). [↑](#footnote-ref-4)
5. <http://www.gks.ru/wps/wcm/connect/rosstat_main/rosstat/en/main/> [↑](#footnote-ref-5)