NATIONAL RESEARCH UNIVERSITY
HIGHER SCHOOL OF ECONOMICS

Tommaso Agasisti, Aleksei Egorov,
Pavel Serebrennikov

HOW DO THE CHARACTERISTICS OF THE ENVIRONMENT INFLUENCE UNIVERSITY EFFICIENCY? EVIDENCE FROM A CONDITIONAL EFFICIENCY APPROACH

BASIC RESEARCH PROGRAM
WORKING PAPERS

SERIES: ECONOMICS
WP BRP 238/EC/2020

How do the characteristics of the environment influence university efficiency? Evidence from a conditional efficiency approach

Tommaso Agasisti[a] Aleksei Egorov[b*] Pavel Serebrennikov[b]

[a] Department of Management, Economics & Industrial Engineering,
Politecnico di Milano School of Management Milan, 20156, Milan, Italy
[b] Laboratory for University Development of Institute of Education,
National Research University Higher School of Economics, 101000, Moscow, Russia
* Corresponding author

Abstract

This paper explores the black box behind efficiency measurements in higher education and define the determinants of university efficiency. Particularly, it investigates how the efficiency of universities is affected by the characteristics of the territory in which they operate. We propose an analysis that combines two perspectives: 1) the resource dependence theory, suggesting that the location of university can determine the amount of resources available to it; 2) institutional isomorphism, according to which the characteristics of other higher education institutions located in the same area may shape the university production function and the efficiency of its operations.

In order to test this framework we use the data on Russian universities and non-parametric conditional order-m efficiency estimator with two categories of exogenous variables. The first group includes the social, economic and cultural characteristics of the region where the university is located. The second set includes the characteristics of other higher education institutions located in the same region. Our findings highlight that the managerial efficiency of universities is strongly associated with the characteristics of the environment in which they operate.

Keywords
Universities, conditional efficiency, order-m, DEA, exogenous variables.

JEL codes
C44, I23

[*]aegorov@hse.ru
16 Potapovskiy Pereulok, Building 10, 101000, Moscow, Russia

1 Introduction

Universities in most developed and developing countries accumulate significant amount of financial resources, a substantial part of which comes from public sources [Pruvot et al., 2015]. A significant increase in scale of higher education systems observed during last decades [Cantwell et al., 2018] expanded the demands from different groups of university stakeholders. These demands are mostly related to the questions regarding how to objectively measure the performance of higher education institutions, what the impact of universities is on the development of different spheres of society, and how to justify public investments in higher education. With most higher education systems experiencing budget constraints [Agasisti, Abalmasova, et al., 2019] the only way of maximizing university performance is to maximize the efficient utilization of resources. The efficiency level of universities can be also considered an important pre-condition for their effectiveness.

Efficiency concept suggests that a university is considered as a production unit that transforms the set of available resources into a set of outputs, primarily teaching and research. In this study by "inefficiency" we mean how much the university could potentially increase its output with the same level of resources. However, the final performance of a higher education institution is determined not only by amount of available resources and the particular technology that is adopted by the organization in order to convert input into output (including the quality of this production technology and operations, the quality of management practices and other "internal" factors). External or "exogenous" factors that cannot be influenced by organization's management at least in the short-run also play a significant role. In the context of the higher education sector, these external factors include different public policies and interventions such as additional funding, merger policies, changes in university autonomy, as well as the different characteristics of the geographical area where the university operates.

The literature suggests that accounting for exogenous variables may significantly influence the distribution of decision making units (DMUs) according to their efficiency scores [Yu, 1998]. Most papers focused on explaining the variation of university efficiency scores use student-related variables [Bradley et al., 2010] and the characteristics of public policies in higher education sector [Zinchenko & Egorov, 2019]. Much less attention is paid to different characteristics of the environment in which universities operate. The question about how the regional environment influences university efficiency is especially important when different policy implications are derived based on efficiency analysis. The ignorance of regional factors in this case may lead to the situation when universities located in a particular area will be in a more advantageous or disadvantageous position depending on the characteristics of the geographical area. To the best of our knowledge, there are no studies applying efficiency models that incorporate the exogenous factors associated with a geographic area in context of higher education. This paper is aimed at filling this gap in the literature.

The innovative contribution of this paper is twofold. First, we propose an analytical framework that analyzes how the environment in which university operates may influence its production process and, consequently, its efficiency and performance. Particularly, this framework combines the resource dependency theory [Pfeffer & Salancik, 2003], suggesting that that the location of university can determine the amount of resources available to it, and institutional isomorphism [DiMaggio & Powell, 1983; Meyer & Rowan, 1977] according to which the characteristics of other higher education institutions located in the same area may shape the university production function and the efficiency of its operations. Based on this conceptual framework we describe a mechanism through which the regional environment can influence university efficiency. Second, in order to test this framework empirically, we employ modern non-parametric methodology - order-m conditional efficiency estimator [Daraio & Simar, 2005], which is applied in analysing the efficiency of HE institutions for the first time.

Data on Russian universities is used in this study. Russia is a federal country with a significant level of regional differentiation in terms of the social, economic and cultural development of regions and of regional higher education systems [Agasisti, Egorov, et al., 2020]. Therefore, higher education institutions included in the sample are characterized by substantial differentiation due to the environment in which they operate. To anticipate the main results, we find a statistically significant association between efficiency scores and variables reflecting the social, economic and cultural development of region, and the characteristics of regional higher education systems.

The paper has 7 sections of which this is the first. The second section provides a brief overview of the Russian higher education sector and its regional dimension. The third section provides the literature review discussing efficiency measurement issues in the context of higher education. The fourth section describes the theoretical framework that is used in order to study the influence of external environment on university efficiency levels. The fifth section presents the details of the methodological strategy and describes the dataset used for the analysis. The sixth section discuss the empirical results. The final section contains a discussion of the results, concluding remarks and policy implications.

2 The Russian higher education sector and its regional dimension

The modern Russian higher education system consists of 1264 institutions, among which 555 are branches, i.e. structural parts of universities usually located in different geographical areas. These branches are small and represent just around 8% of the total student body. There are both public and private higher education institutions. Like branches, private universities are usually small and represent around 4% of total student population. The core of the Russian higher education system consists of 494 public universities with 2.28 million students (88% of the total student population).

The system is a successor of the Soviet one which was a quasi-corporate system of higher education [Froumin et al., 2014]. The design, structure and activities of the Soviet higher education system were subordinated to the needs of national economy. The central government tightly controlled number of higher education institutions in different regions, the total number of graduates, the distribution of graduates across different fields of study and a lot of other parameters. The culmination of this centralization and control was the compulsory distribution of graduates to the workplaces [Egorov et al., 2019].

After the collapse of Soviet union in 1991, the Russian higher education system experienced the emergence of market mechanisms, diversification, an increase in university autonomy and many structural reforms. However, it is still one of the most centralized higher education systems in the World [Carnoy et al., 2018]. Only 44 out of 494 public universities (around 2% of total student population) are governed at the regional level. All other public higher education institutions are governed by Ministry of Science and Higher Education or other federal agencies. Moreover, regional governments have a limited set of support channels for universities located on their territories.

Despite the centralization of governance, it is still can be argued that the location of a university in a particular region can influence significantly how it operates. This can be due to the huge differentiation of Russian regions in terms of their social and economic development. Particularly, as shown in [Agasisti, Egorov, et al., 2020], in 2017 the gross regional product growth rate in different regions varied between -6% and 12%; share of employed people with higher education varied between 8.7% and 46.2%; the share of students among population aged 17-25 varied between 13% and 44%.

Therefore, despite the Russian higher education system preserving a lot of characteristics of the Soviet one in terms of centralization, it still represents a good example for investigating how regional disparities in terms of social, economic and cultural development can influence the operations of universities.

3 Literature review

Efficiency measurements in education in general and in higher education sector in particular grounded in the education economics literature have focused on educational production functions that relate the resources used by education organization and the outputs produced [Hanushek, 1987; Monk, 1989]. The studies of higher education institution efficiency evolved quite recently primarily because of the difficulties associated with the specification of universities' production function. The main difficulties are the non-profit nature of this type of organization, the absence of input and output prices [J. Johnes, 2006] and the multiplicity of consumed inputs and produced outputs. The first attempts to measure universities' efficiency were based on the regression analysis [J. Johnes & Taylor, 1990], subsequent studies were divided into two streams – efficiency estimations based on non-parametric techniques such as DEA ([Charnes et al., 1978]) and parametric approaches based on stochastic frontier analysis ([Aigner et al., 1977; Meeusen & van Den Broeck, 1977]). Interested readers can find an excellent review of the vast literature on the efficiency analysis of higher education institutions in [De Witte & López-Torres, 2017].

Recent papers studying higher education efficiency emphasize the role of exogenous factors that are out of management control at least in the short-run, but can significantly influence the university production process. The first studies with this focus were based on two-stage methodology that combined efficiency estimation techniques (usually DEA or SFA) with a tobit regression model, i.e. at the second stage different exogenous factors of interest are regressed on the efficiency scores estimated at the first stage. An example is [Kounetas et al., 2011] where the efficiency of Greek university departments is evaluated using DEA and at the second stage the authors examine the degree to which environmental factors impact departmental efficiency. The main disadvantages of this type of methodologies are possible serial correlation [Xue, Harker, et al., 1999] and the possible endogeneity at the second stage regression [Simar, Lovell, et al., 1994].

Most studies focusing on how external factors affect higher education efficiency utilize another two-stage methodology proposed by [Simar & Wilson, 2007]. The main idea of this approach is a double-bootstrap procedure where, first, the bootstrap is applied to the efficiency scores estimation and, second, a bootstrap is also applied to the second stage truncated regression. This technique overcomes the main drawbacks of standard two-stage procedures outlined above. Some examples of studies that utilize this methodology to analyze the influence of exogenous factors on university efficiency are as follows. [Agasisti, Barra, et al., 2019] include in the efficiency analysis such external factors as a dummy variable for having a medical school in the university, the number of years as a technology transfer office have been opened in the university, the percentage of dropouts by the end of the first year, the size of the lump-sum subsidy that central government transfers to the university, the size of the students body. [Agasisti, Egorov, et al., 2020] in a similar context used a list of exogenous variables that includes the share of Master's students in the university, share of full-time students, a dummy variable representing a university being in the region capital, the presence of a medical school and the university's share of the regional higher education market. [Zinchenko & Egorov, 2019] consider exogenous variables that are associated with different public policies and programs implemented in the higher education sector.

Another methodology that can incorporate exogenous variables in efficiency analysis is a one-stage approach suggesting that exogenous factors are considered as free disposal inputs and outputs, but which are not active in the optimization process [Färe et al., 1989]. The main drawbacks of this technique are that, first, we have to assume free disposability and, second, we have to know a priori what the particular role of environmental factors is in the production process [Daraio & Simar, 2007]. This approach was not applied in the context of higher education efficiency measurement.

All these approaches have been subject to considerable criticism. [Daraio & Simar, 2007] argue that even corrected by the bootstrap two-step techniques have disadvantages. The first one is that this technique relies on a separability condition between the input-output space and the space of environmental variables. Secondly, the second stage regression always relies on parametric assumptions, and modelling the shape of the relationship between efficiency and exogenous variables is difficult in this setting.

The newest technique that can fix the issues outlined above is a conditional efficiency estimator proposed by [Daraio & Simar, 2005] and which we use in this paper. There are two studies that apply this techniques in the sphere of education in general. [De Witte & Kortelainen, 2013] use this methodology in order to investigate how students' background characteristics influence their performance. [Haelermans & De Witte, 2012] studied school performance with a focus on exogenous variables that reflects educational innovations. This estimator has not been implemented in the context of university efficiency measurements.

4 Theoretical framework: resource dependence theory and institutional isomorphism

The theoretical framework to investigate how external factors influence universities' efficiency and performance combines resource dependence theory and institutional isomorphism. This section reviews these two theoretical approaches to the analysis of organizational behavior. The next section proposes the particular variables that should be included in the model for explaining the variation of universities' efficiency scores based on this review.

4.1 Resource dependence theory

Resource dependence theory (RDT) was proposed by [Pfeffer & Salancik, 2003] and is widely used in the discussions about organizational behavior and organizational dynamics. RDT covers three primary themes: the environmental effects on organizations; organizational efforts to manage environment constraints; and how environmental constraints affect internal organizational dynamics [Powell & Rey, 2015]. Although many studies employing RDT are focused on for-profit organizations, there are also studies that apply RDT in the nonprofit sector and, particularly, in the higher education sector (see, for example, [Askin, 2007]). The main proposition of this theory is that organizations always depend on particular resources, and some of these resources can be out of organization's management control. In other words, the environment provides critical resources needed by the organization [Niehüser, 2008], and organizations' activities and decisions can be influenced by different external forces that control these critical resources.

In the context of our research question, we assume that the regional environment in general is one of these external forces that influence the activities of universities. Particularly, we assume that 3 regional characteristics – economic, social and cultural - can significantly determine the amount of resources available for universities located in the region. The Russian higher education system is highly centralized [Carnoy et al., 2018] and public funding is distributed by the federal government [Agasisti, Abalmasova, et al., 2019]; and regional governments have limited incentives and channels through which they can cooperate with higher education institutions [Egorov et al., 2019]. Russian regions are characterized by tremendous differentiation in terms of different socio-economic indicators [Agasisti, Egorov, et al., 2020] and these regional differences can create fundamentally different environments in which universities operate.

We assume that the characteristics of a region can significantly influence the availability of the two main types of resources needed by the universities – financial and human. Therefore, we define two particular channels through which the regional environment in which a university operates may influence its resources endowment. First, regions with more favorable economic, social and cultural conditions are more attractive for living, therefore, school graduates will be more likely to apply to the universities in these regions and, consequently, it will be easier for the universities to attract school graduates of higher quality. The quality of incoming students is usually considered in education economics literature [Rothschild & White, 1995] as an important input in terms of the university production function. Thus, location in a region with favorable economic, social and cultural conditions allows universities to attract high-quality entrants at a lower cost[1]. Second, universities located in regions with a relatively high level of economic development are more able to attract financial resources. Universities face greater demand for their services and greater willingness to pay. This corresponds to both demand for higher education paid by tuition fees[2] and locals firm demand for R&D implemented by universities. Universities located in economically developed regions may have an access to additional funds from the regional budget[3].

It is important to mention here that there are alternative theories explaining the same phenomena as RDT. [Niehüser, 2008] highlights two dimensions in which RDT is related to other organizational theories. First, resource dependence can be considered as a part of a more general theory, particularly, social exchange and power theories [Emerson, 1976]. Secondly, RDT has some interceptions with the same order theories such as the resource-based view [Helfat & Peteraf, 2003], institutional theory [Scott, 1987], etc.

4.2 Institutional isomorphism

Institutional isomorphism was proposed in [Meyer & Rowan, 1977] and suggests that organizations facing the same environmental conditions tend to become more and more similar to each other. [DiMaggio & Powell, 1983] highlight three different types of institutional isomorphism. The first type is a coercive isomorphism that is considered a result of the formal and informal pressures from other organizations and the cultural expectations of the society in which the organization exists. This type also includes different government regulations. The second type is represented by mimetic isomorphism suggesting that organizations can imitate the behavior of other organizations of the same type due to uncertainty. The third and final type is normative isomorphism which is associated with professionalization: a group of people who share the same profession may try to define principles and methods of implementing their work and working conditions.

There are numerous studies that apply institutional isomorphism to study institutional change in higher education. [Croucher & Woelert, 2016] empirically tests this concept using the data on Australian universities and find significant convergence in terms of formal organizational structures and student and staff numbers in the majority of academic fields. [Marginson & Considine, 2000] argue that mimetic isomorphism was a wide-spread strategy for less prestigious Australian universities in order to adopt to public policies. [Fay & Zavattaro, 2016] highlight the importance of isomorphic forces that influence branding and marketing strategies of the US higher education institutions.

Most studies of institutional isomorphism in higher education focus on the isomorphic convergence among universities that face the same governance regime. Due to high level of centralization of the Russian higher education system, most universities have the same conditions in terms of how the system is managed on the national level. But we assume that two channels within institutional isomorphism can explain how higher education institutions located in the same region can influence each other's activities. The first channel of pressure on university from other higher education institutions located in the same area is related to coercive isomorphism

[1] This channel is especially important in the Russian context, because many school graduates consider higher education as an opportunity to move to another region or city with a higher level of socio-economic development (around 50% of all school graduates who want to continue their education at the university).

[2] Russian universities offer both publicly-funded and tuition-fee based education (depending on student's achievements.)

[3] Russian universities are primary funded from the federal budget and private sources. Only the regions with a high level of economic development may provide additional subsidies from their budgets (see [Agasisti, Abalmasova, et al., 2019] for details).

and represents different pressures that can occur due to competition. According to the conceptual framework of university competition proposed by [Agasisti & Johnes, 2009], "the incentive in a growing competitive environment, for each university, is to improve its own performance relative to all others". The second channel corresponds to mimetic isomorphism according to which organizations that face the same environment tend to imitate each other's behavior. Particularly, universities located in one region face the same market structure for higher education and research and the same characteristics of social and economic environment. Thus, according to institutional isomorphism, these universities have the same uncertainty and overcome it by imitating each other.

5 Methodology and data

5.1 Efficiency estimation taking into account exogenous factors

We formulate an extension of the conventional non-parametric efficiency estimator (DEA) which is the order-m conditional efficiency estimator. It is based on the probabilistic formulation of the efficiency measurement problem proposed by [Cazals et al., 2002] and differs from standard DEA estimator in two main perspectives. First, it measures the efficiency with respect to the partial frontier that is constructed using only part of the available data points, i.e. it eliminates the drawback of DEA related to its high sensitivity to outliers. Secondly, this approach accounts for exogenous variables, which is crucial for our research question.

5.1.1 Probabilistic nature of the efficiency estimation problem

Order-m efficiency methodology is based on the representation of the production process that has a probabilistic nature. Suppose we have p inputs ($x \in \mathbb{R}_{+}^{p}$), q outputs ($y \in \mathbb{R}_{+}^{q}$) and production technology set Ψ is:

$$\Psi = \{(x, y) \in \mathbb{R}_{+}^{p+q} \mid x \; can \; produce \; y\} \tag{1}$$

As in the case of the order-m efficiency estimator, the production technology set is unobserved because of its theoretical nature and can be estimated only empirically using a random sample: $\chi_n = \{(x, y) \mid i = 1, ..., n\}$. Within the framework of probabilistic representation the production technology set is represented as a joint probability function that reflects the probability of being dominated at the level (x, y):

$$H_{XY}(x, y) = P(X \leqslant x, Y \geqslant y) \tag{2}$$

For an output-oriented setting this joint probability function can be represented as:

$$H_{XY}(x, y) = P(\, Y \geqslant y \,|X \leqslant x \,)\, P(X \leqslant x \,) = S_{Y|X}(y \mid x)\, F_X(x) \tag{3}$$

where $S_{Y|X}(y|x)$ is the conditional survival function of Y given that $X \leqslant x$ and $F_X(x)$ is the distribution function of X. Note that it is assumed that $S_Y(y) > 0$ and $F_X(x) > 0$ - in other words, they exist. Then we can introduce the definition of output efficiency measure $\lambda(x, y)$:

$$\lambda(x, y) = \sup\{\lambda \mid S_{Y|X}(\lambda y \mid x) > 0\} = \sup\{\lambda \mid H_{XY}(x, \lambda y) > 0\} \tag{4}$$

In order to obtain efficiency scores in practice we have to estimate the empirical analogues of $H_{XY}(x, y)$ and $S_{Y|X}(y|x)$. These two terms can be estimated using:

$$\hat{H}_{XY,n}(x, y) = \frac{1}{n} \sum_{i=1}^{n} I(x_i \leqslant x, y_i \geqslant y) \tag{5}$$

and

$$\hat{S}_{Y|X,n}(y \mid x) = \frac{\hat{H}_{XY,n}(x, y)}{\hat{H}_{XY,n}(x, 0)} \tag{6}$$

where $I()$ is the indicator function.

5.1.2 Partial frontier

Conventional non-parametric techniques for efficiency measurements like DEA of FDH can be sensitive to outliers. In order to overcome this drawback [Cazals et al., 2002] used the expected value of maximum of $m \geqslant 1$ random variables drawn from conditional distribution function $F_{X|Y}(x \mid y)$:

$$\phi_m(x) = E[max(Y^1, ..., Y^m) \mid X \leqslant x] = \int_0^\infty (1 - F_{Y|X}(y \mid x)^m) dy \tag{7}$$

The main idea of this approach is to estimate a partial frontier based on m random units which have input level less than x. Therefore, we consider $\phi_m(x)$ as a benchmark for an organization consuming input x.

The empirical analogue of $\phi_m(x)$ is defined as:

$$\hat{\phi}_{m,n}(x) = \hat{E}[max(Y^1, ..., Y^m) \mid X \leqslant x] = \int_0^\infty (1 - \hat{F}_{Y|X}(y \mid x)^m) dy \tag{8}$$

The first step to obtain efficiency scores with respect to the partial frontier is to define the order-m attainable set $\Psi_m(x)$ for a given input x where Y_i, $i = 1, ..., m$, is an i.i.d. random variable drawn from conditional distribution function $F_{Y|X}(y|x)$:

$$\Psi_m(x) = \{(u, v) \in \mathbb{R}_+^{p+q} \mid u \leqslant x,\ v \geqslant Y_i \quad i = 1, ..., m\} \tag{9}$$

Secondly, we obtain the following radial efficiency measure:

$$\widetilde{\lambda}_m(x, y) = \sup\{\lambda \mid (x, \lambda y) \in \Psi_m(x)\} \tag{10}$$

It is important to note that the efficiency measure described above is a random variable because of the stochastic nature of Y_i. We can rewrite it in a following way to define the order-m efficiency measure:

$$\lambda_m(x, y) = E(\widetilde{\lambda}_m(x, y) \mid X \leqslant x) = \int_0^\infty (1 - F_{Y|X}(uy \mid x)^m) du \tag{11}$$

In practice $F_{Y|X}(y|x)$ and $E(\widetilde{\lambda}_m(x, y) \mid X \leqslant x)$ are replaced by their empirical analogues:

$$\hat{\lambda}_m(x, y) = \hat{E}(\widetilde{\lambda}_m(x, y) \mid X \leqslant x) = \int_0^\infty (1 - \hat{F}_{Y|X}(uy \mid x)^m) du \tag{12}$$

5.1.3 Conditional order-m efficiency estimator

[Daraio & Simar, 2005] extended the partial frontier approach to incorporate exogenous variables in this analysis. The main advantage of this method is that it estimates the production frontier and avoids the separability condition between input-output space and Z-space ($Z \in \mathbb{R}_+^r$) when exogenous factors, in fact, do not influence either the attainable set Ψ or the production frontier.

Therefore, order-m conditional efficiency is based on the probabilistic formulation mentioned above but with additional condition such that $Z = z$:

$$H_{XY|Z}(x, y \mid z) = P(X \leqslant x, Y \geqslant y \mid Z = z) \tag{13}$$

This equation can be decomposed for an output conditional distribution:

$$H_{XY|Z}(x, y \mid z) = S_{Y|XZ}(y \mid x, z)\, F_{X|Z}(x \mid z) \tag{14}$$

After that, we can again define efficiency measure $\lambda(x, y \mid z)$:

$$\lambda(x, y \mid z) = \sup\{\lambda \mid S_{Y|XZ}(\lambda y \mid x, z) > 0\} = \sup\{\lambda \mid H_{XY|Z}(\lambda x, y \mid z) > 0\} \tag{15}$$

Similarly to the previous case without exogenous variables, $H_{XY|Z}(x, y \mid z)$ and $S_{Y|Z}(y \mid z)$ should be approximated by their empirical analogues. Finally, we can consider in the same way the partial frontier instead of the full frontier using the order-m efficiency estimator to define the conditional order-m effciency measure:

$$\lambda_m(x, y \mid z) = E(\widetilde{\lambda}_m^z(x, y) \mid X \leqslant x, Z = z) = \int_0^\infty (1 - F_{Y|XZ}(uy \mid x, z)^m)du \tag{16}$$

where $\widetilde{\lambda}_m^z(x, y) = \sup\{\lambda \mid (x, \lambda y) \in \Psi_m^z(x)\}$ and $\Psi_m^z(x) = \{(u, v) \in \mathbb{R}_+^{p+q} \mid u \leqslant x,\ v \geqslant Y_i \ \ i = 1, ..., m\}$ but with $Y_i,\ i = 1, ..., m$ produced by the distribution function $S_{Y|XZ}(y \mid x,\ z)$. In practice the estimator $\hat{S}_{Y|XZ}(y \mid x, z)$ and $\hat{E}(\widetilde{\lambda}_m^z(x, y) \mid X \leqslant x, Z = z)$ are used.

5.2 Data and descriptive analysis

The data used for the analysis comes from the Monitoring of Performance of Higher Education Institutions conducted by Russian Ministry of Science and Higher Education. Data on different regional socio-economic characteristics of regions used as exogenous variables for the efficiency analysis were collected from the Federal State Statistic Service (Rosstat) database. All data refers to 2017.

Our sample consists of public Russian higher education institutions without their branches. Among these institutions the following constraints are imposed: we consider universities in which the total number of academic staff is more than 100 and the number of students exceeds 500. We also restricted our sample by excluding universities with zero income obtained from R&D projects. The fourth restriction is to remove universities with zero citations of articles, books, monographs, etc. published in journals indexed by Web of Science, Scopus or the Russian Science Citation Index (RSCI). By imposing these limitations we exclude higher education institutions with very specific types of production function in order to homogenize our sample. The final sample is 371 universities.

It is important to make some notes concerning input and output variables used for the efficiency analysis. Following [Cohn et al., 1989], we consider universities as multi-product organizations that process inputs to produce outputs. Following the literature on efficiency measurements in higher education, we suggest three inputs for the efficiency model. The first one is the financial resources of the university measured by income from all sources [Agasisti & Johnes, 2009; Agasisti & Pérez-Esparrells, 2010]. The second input is the total number of academic staff, which measures the human capital resources available [Agasisti & Johnes, 2009; Agasisti & Pérez-Esparrells, 2010; Agasisti & Pohl, 2012; Wolszczak-Derlacz & Parteka, 2011]. The third input is the average unified state exam score which is accepted by all Russian public universities as an entrance exam. This variable reflects the quality of entrants. The literature on education economics suggests that the quality of students should be an important resource available for educational organizations [Hoxby, 2009].

Three output variables approximate three different missions of universities: research, teaching and university-industry collaboration. The first output is the total number of citations (over the last 5 years) of articles published in journals indexed in Web of Science, Scopus and RSCI. This variable reflects university scientific performance. The second output is the total income obtained from research and development projects conducted by the university which approximates the university's collaboration with industrial partners. The third output is the total number of graduates of higher education institution employed one year after graduation, which is an indicator of university teaching performance[4].

[4]This is an approximated variable calculated as the number of university's graduates (computed as the ratio

We define the set of exogenous variables included in the efficiency analysis in coherence with our theoretical framework. According to the resource-dependence theory we picked indicators that reflect the social, economic and cultural development of the region. These indicators are (1) the share of employed population with higher education, (2) gross regional product, (3) life expectancy at birth and (4) the total number of theater and museum visitors per 1000 of population. Based on institutional isomorphism - (1) the weighted (by the number of university students) average unified entrance exam scores of universities located in the region, (2) the Herfindahl—Hirschman index (HH index) which is calculated as the sum of the squares of the ratio of students of the university to the total number of students in the region, (3) average citations over the last 5 years of articles published by journals indexed in RSCI, Web of Science or Scopus of all universities in the region (weighted by the total number of university staff), and (4) the total number of leading universities in the region[5]. The full list of variables used for efficiency analysis is represented by Table 1. Descriptive statistics for these variables are presented in Table 2.

of the university's students to all students in the region × the number of graduates in the region) × the share of employed graduates during the year after graduation.

[5]Leading universities are institutions that have special status given by the Government. Most of these universities participate in Russian excellence initiative - project 5-100, i.e. these institutions receive additional funding to improve their positions in international rankings.

Table 1: Characteristics of universities

Variable name	**Description**	**Unit of measurement**
Inputs		
IncomeUniv	The income of university from all sources	Thousand rubles
Faculty	Total number of full-time faculty	Persons
Exam	The average entrance exam score of entering students	Scores
Outputs		
Citations	Total number of citations of articles published by journals indexed by RSCI, WoS and Scopus	Units
R&D	The total quantity of R&D	Thousand rubles
EmployedGrad	Total number of universities' graduates employed during the year after graduation	Persons
Environmental variables: resource-dependence theory		
Employment	Share of employees with higher education	%
GRP	Gross regional product	Million rubles
Life	Life expectancy at birth	Years
Culture	Total number of theatre and museum visitors per 1000 population	Persons
Environmental variables: institutional isomorphism theory		
WMeanExam	Weighted average entrance exam score across universities located in the region	Scores
HHindex	Herfindahl—Hirschman index for the ratio of students of the university to the total number of all students in the region	Points
WMeanCitations	Weighted average number of citations of articles published by journals indexed in RSCI, WoS and Scopus across universities located in the region	Units
TotalLeaders	Total number of leading universities in the region	Units

Source: Annual Monitoring of Efficiency of Higher Education Institutions, Federal State Statistics Service (Rosstat)

Table 2: Descriptive statistics of the variables used for efficiency evaluation

Variable name	**2017** (N = 371)				
	Mean	Std dev	Median	Min	Max
Inputs					
IncomeUniv	1 758 584	2 470 338	973 026	164 214	25 418 823
Faculty	577	639	412	101	8 768
Exam	67.8	10.1	66.6	48.3	95.8
Outputs					
Citations	10 106	20 109	4 262	136	191 793
R&D	230 615	547 657	56 119	450	6 069 631
EmployedGrad	1 314	1 004	1 040	89	6 965
Environmental variables: resource-dependence theory*					
Employment	32.83	5.04	31.62	25.58	49.49
GRP	913 741	1 931 926	419 203	44 572	15 724 910
Life	72.05	2.28	71.78	66.10	81.59
Culture	875	804	666	145	5769
Environmental variables: institutional isomorphism theory*					
WMeanExam	62.26	14.97	64.79	0	80.22
HHindex	0.435	0.3	0.332	0	1
WMeanCitations	45 723	125 992	13 036	0	975 480
TotalLeaders	0.5	1.69	0	0	14

* Descriptive statistics of environmental variables refer to regions
Source: Annual Monitoring of Efficiency of Higher Education Institutions, Federal State Statistics Service (Rosstat)

Universities included in the sample have different sizes (total income from all sources ranged between 164 million to 25 billion roubles; total number of full-time faculty from 101 to 8,768). They also represent different segments of the higher education system in terms of its quality. Averaged unified state exam score is ranged between 48.3% and 95.8%. Therefore, despite excluding from our sample very specific institutions with specific productions functions, our sample is still represents different types of universities and includes more than a half of all Russian public institutions.

The significant differentiation of Russian universities can be also viewed in terms of their productivity. The total number of citations of publications indexed by Web of Science , Scopus and RSCI ranges between 136 and 191 793, and the total value of R&D projects ranged from 0.45 to 6,070 million rubles. Although we restricted our sample to only universities with more than 500 students, the number of employed graduates is between 89 and 6,965.

As for environmental variables associated with Russian regional context, there is a high degree of variation observed among the factors of resource-dependence theory and institutional-isomorphism, which coincides with the considerable differentiation of Russian regions [Agasisti, Egorov, et al., 2020].

6 Results

Before estimating the effect of different environmental factors on university' efficiency, we obtain standard efficiency estimates and efficiency estimates corrected for exogenous factors. In order to obtain standard non-parametric efficiency scores we use order-m methodology that gives efficiency scores corrected for the presence of outliers in the sample. Efficiency scores corrected for environmental variables are obtained through combining the conditional efficiency and partial frontier approaches which are described in detail in Section 5.1. In other words, our aim is to compare the distribution of the unconditional efficiency scores and the distribution of efficiency scores which account for exogenous factors.

In order to implement the order-m efficiency estimator we use $m = 100$ meaning that each DMU is benchmarked against 100 random observations with greater output level. This particular choice eliminates the potential effects of outliers. The number of DMUs used for building the efficiency frontier is large enough to provide the robustness of our estimations. The procedure of partial frontier construction was repeated 2,000 times in order to eliminate the randomness of the set of DMUs selected for frontier construction.

All efficiency estimates are based on output-oriented models, meaning that we assume that amount of available resources is fixed for the university, and managers are focused on maximizing outputs rather than minimizing costs holding a particular level of output. This assumption seems to be reasonable for the higher education sector, since we consider only public institutions for which amount of available resources is exogenously determined (by the federal Ministry of Science and Higher education, regional government and so on). In order to obtain empirical estimates of the distributions required to calculate corrected efficiency scores we use smoothing technique. Assuming that q exogenous variables are taken and product kernel is used ($K(u) = k(u_1)...k(u_q)$), optimal bandwidth for smoothing is based on the following formula:

$$h_j = C_v(k, q)\, \hat{\sigma_j}\, n^{-1/(2v+q)} \tag{17}$$

where n is the number of observations; $\hat{\sigma_j}$ denotes empirical standard deviation of variable $j, j = 1, ..., q$; $C_v(k, q)$ is a rule-of-thumb constant depending on type the of kernel $k(u_j)$, the order of kernel v and the number of external factors q. For the second-order Epanechnikov kernel and 8 environmental variables, $C_v(k, q)$ equals 2.0037.

Descriptive statistics for the efficiency scores (conventional and corrected for exogenous factors) are presented in Table 3. Accounting for exogenous factors significantly reduces the variation of efficiency scores - the standard deviation is less half of corrected efficiency scores (0.09 vs 0.23), indicating that part of the inefficiency distribution is actually related to the environment in which universities operate.

Table 3: Descriptive statistics of conditional and unconditional order-m efficiency scores

(output-oriented case, m = 100)

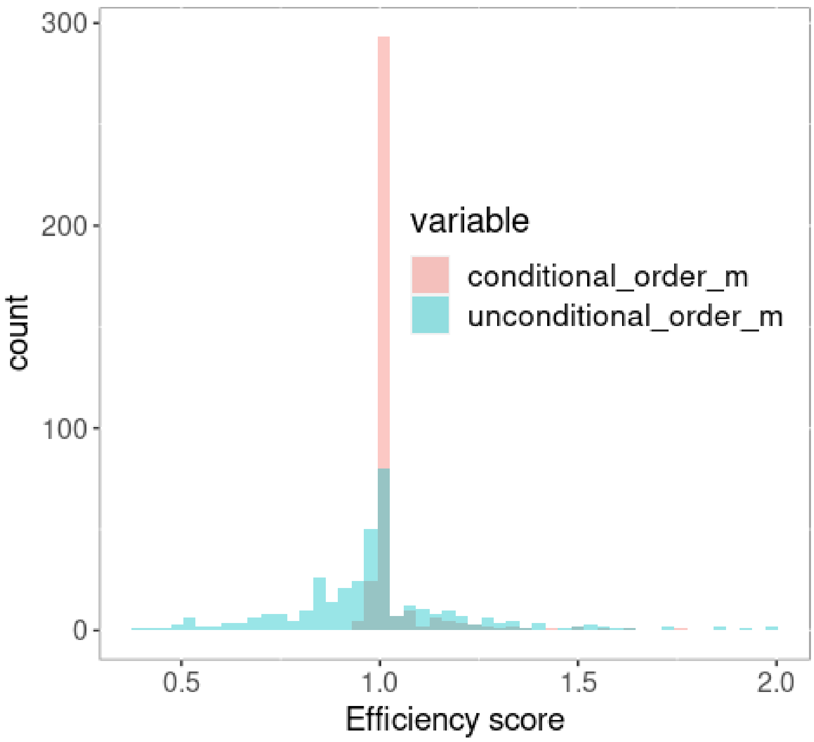
Type of order-m	**Mean**	**Std dev**	**Median**	**Min**	**Max**
Unconditional	0.9887	0.2315	0.9897	0.4125	2.0156
Conditional	1.0246	0.0902	1.0000	0.9494	1.7522

Source: authors' calculations

We also compare standard and corrected efficiency scores by plotting their densities. The histograms of standard and corrected efficiency scores are presented in Figure 1. This figure suggests that the variation of efficiency scores is significantly lower when we account for exogenous factors. This drop in variation after taking into consideration different variables reflecting regional environment, and the different characteristics of the regional higher education system, allows us to hypothesize that these external variables partially explain the variation of universities' efficiency levels.

In order to identify the impact of particular exogenous variables on the efficiency level of universities we use partial regression plots. If the smoothed regression line is increasing, it means that this particular exogenous factor has a favourable impact on the efficiency level. Otherwise, this exogenous factor negatively affects an organization's performance. Figures 2

Figure 1: Ordinary and conditional order-m efficiency distributions



represents conventional partial regression plots based on local-linear regressions. The solid lines are non-parametric regression lines, while the dashed lines represent bootstrapped confidence intervals.

It is also important to test the statistical significance of environmental factor effects. For this purpose we utilize test proposed by [Racine, 1997]. The null and alternative hypotheses H_0 and H_1 are as follows:

$$H_0 : \beta = 0 \tag{18}$$
$$H_1 : \beta \neq 0 \tag{19}$$

where β is the coefficient of exogenous variables in partial regression.

Table 4: Nonparametric significance test

Variable	**p-value**	**Effect**
Employment	0.2983	No effect
GRP	0.0000***	Positive
Life	0.4236	No effect
Culture	0.1328	No effect
WMeanExam	0.0000***	Positive
HHindex	0.0000***	Negative
WMeanCitations	0.0000***	Positive
TotalLeaders	0.0000***	Positive
Significance levels: *** p-value < 0.001; ** p-value < 0.01		

Source: authors' calculations

The results of non-parametric testing that reveals statistical significance of the effects are

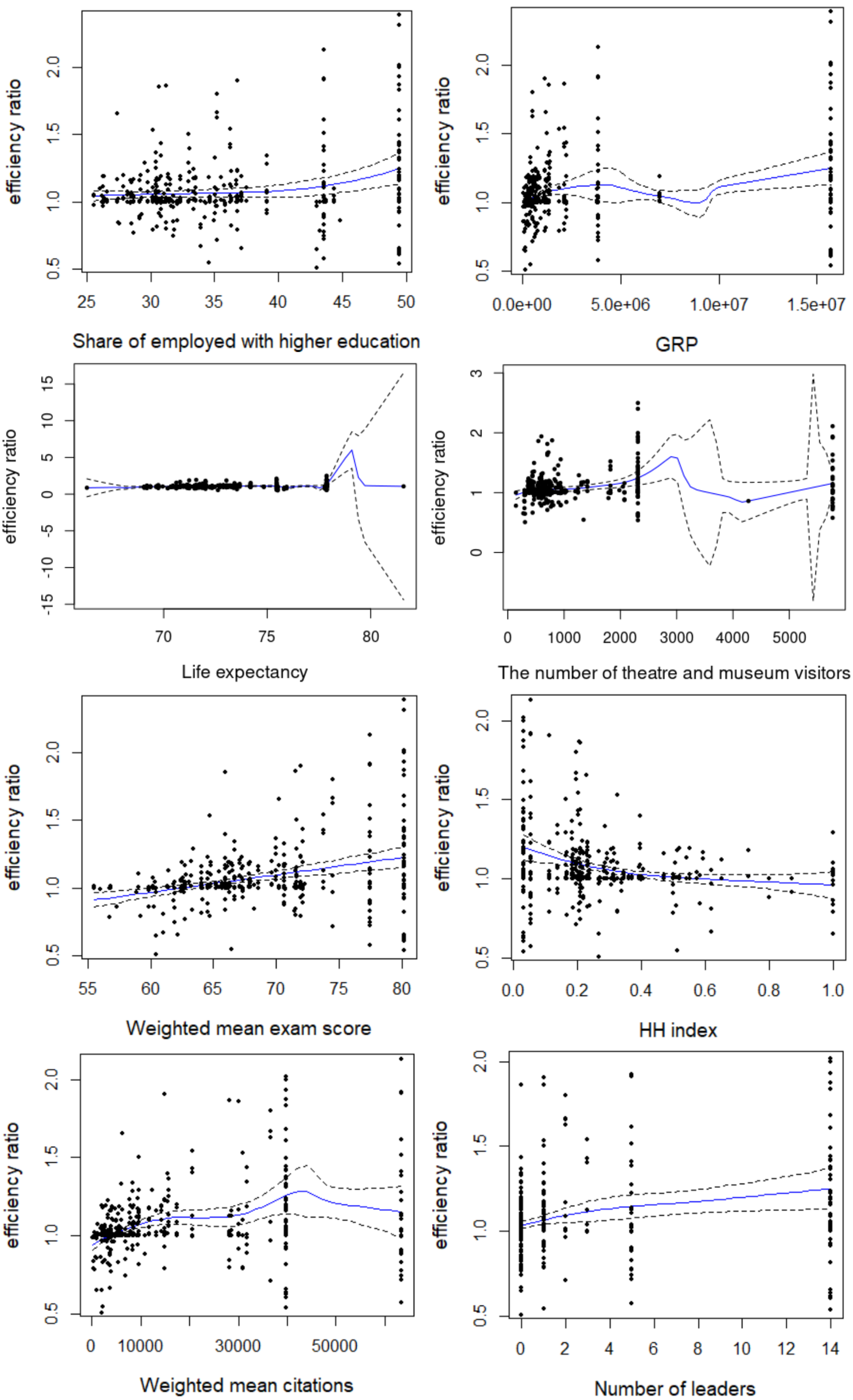


Figure 2: Effect of different exogenous factors on order-m frontier using local linear regression

presented in Table 4. The findings indicate that the indicators highlighted, based on resource-dependence theory, are not statistically significant for exploring university efficiency. The only exception is gross regional product. The relationship between university efficiency and gross regional product has non-linear nature, and the latter has a weak positive impact on the former. Life expectancy is not statistically significant. The number of theatre and museum visitors per 1000 of population is also not significant. In general, we can conclude that among the regional indicators only economic ones are significant predictors of university efficiency and there is no impact of social and cultural factors. For universities located in more economically developed regions, it is easier to attract additional human and financial resources as explained in the theoretical framework, and it is easier to build collaborations with different external partners.

As for variables reflecting the characteristics of regional higher education systems that we derived based on institutional isomorphism, all of them are statistically significantly related to university efficiency. Figure 2 shows the positive relationship between university efficiency and entrance exam score averaged across all universities located in the region. The HH index has negative effect on efficiency level where the low level of the HH index corresponds to a strong competitive environment in the region. A weaker positive effect is observed for the average number of citations and for number of leading higher education institutions located in the region. This finding highlights the importance of a competitive environment for university efficiency. The location of a university in a regional higher education system with strong peer institutions creates incentives for increasing its productivity.

To summarize, we observe that the efficiency level of higher education institutions is determined not only by the resources and outputs as predicted by the standard educational production function theory, but it is also dependent on different external variables reflecting both the specific features of region and the structure of the regional higher education system.

7 Discussion and concluding remarks

This study represents the first attempt to analyze how different characteristics of the territory where a university is located influence its efficiency by using modern non-parametric order-m conditional efficiency methodology. The results described in the previous section have several important implications.

First, due to high level of centralization in the Russian higher education sector, many policies and programs pay little attention to the regional environments in which particular universities operate [Egorov et al., 2019]. In other words, higher education institutions are often treated by the federal Government as a homogeneous population facing the same opportunities and constraints. The results of our analysis suggest, that given high regional disparities and the significant influence of these disparities on universities' operations, public policies in higher education should be context-specific and take into account regional factors which are out of control of university managers.

The second policy recommendation is to develop different mechanisms that can intensify the role of regional governments in the developments of higher education. As mentioned in Section 2, regional authorities in Russia have a limited set of channels through which they can support universities located on their territories [Carnoy et al., 2018]. The development of links with regions may create opportunities for the universities to influence the regional environment and make it more predictable, which is especially important given the high sensitivity (as suggested by our results) of university operations to the external context.

The third and final policy implication is that the level of competition in the regional higher education markets is a predictor of university performance. The effect of the concentration of universities in the region or the effect of the number of leading universities in the region may influence efficiency through different mediating factors, but we assume that in the Russian context competition can be considered the most important one. During the last decade Russian higher education sector has experienced several reforms which reduced the number of universities and, consequently, the level of competition in some regions. These reforms were mostly related to university mergers of different types (the establishment of large federal universities by merging several institutions, the mergers of small and low-performing universities, etc) . Our results suggest that these policies should be implemented carefully, without radical changes in structure of regional higher education markets.

Finally, this paper has some limitations, the main one being that our research design cannot study the particular mechanisms through which regional environment may influence university operations in detail. In order to overcome this limitation more narrowly focused studies are needed - case studies of particular regions, interviews with managers, etc. Another limitation of this study is that we do not discuss specifically the issue of education and labor mobility. The rate of university graduate migration may influence different regional characteristics which consequently, according to our conceptual framework, may influence university efficiency. We assume that these effects should not be sufficiently strong to influence our main results, however, this problem requires additional investigations.

Acknowledgements.
This paper is an output of a research project implemented as part of the Basic Research Program at the National Research University Higher School of Economics (HSE).

References

[1] Agasisti, T., Abalmasova, E., Shibanova, E., & Egorov, A. (2019). The causal impact of performance-based funding on university performance: Quasi-experimental evidence from a policy in russian higher education. *Higher School of Economics Research Paper No. WP BRP*, *221*.

[2] Agasisti, T., Barra, C., & Zotti, R. (2019). Research, knowledge transfer, and innovation: The effect of italian universities' efficiency on local economic development 2006-2012. *Journal of Regional Science*, *59*(5), 819–849.

[3] Agasisti, T., Egorov, A., Zinchenko, D., & Leshukov, O. (2020). Efficiency of regional higher education systems and regional economic short-run growth: Empirical evidence from russia. *Industry and Innovation*, 1–28.

[4] 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.

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

[6] 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.

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

[8] Askin, J. A. (2007). Community college mission: Re(s)ources make a difference. *Community College Journal of Research and Practice*, *31*(12), 977–997.

[9] Bradley, S., Johnes, J., & Little, A. (2010). Measurement and determinants of efficiency and productivity in the further education sector in England. *Bulletin of Economic Research*, *62*(1), 1–30.

[10] Cantwell, B., Marginson, S., & Smolentseva, A. (2018). *High participation systems of higher education*. Oxford University Press.

[11] Carnoy, M., Froumin, I., Leshukov, O., & Marginson, S. (2018). *Higher education in federal countries: A comparative study*. SAGE Publishing India.

[12] Cazals, C., Florens, J.-P., & Simar, L. (2002). Nonparametric frontier estimation: A robust approach. *Journal of Econometrics*, *106*(1), 1–25.

[13] Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, *2*(6), 429–444.

[14] Cohn, E., Rhine, S. L., & Santos, M. C. (1989). Institutions of higher education as multi-product firms: Economies of scale and scope. *The Review of Economics and Statistics*, 284–290.

[15] Croucher, G., & Woelert, P. (2016). Institutional isomorphism and the creation of the unified national system of higher education in Australia: An empirical analysis. *Higher Education*, *71*(4), 439–453.

[16] Daraio, C., & Simar, L. (2007). *Advanced robust and nonparametric methods in efficiency analysis: Methodology and applications*. Springer Science & Business Media.

[17] Daraio, C., & Simar, L. (2005). Introducing environmental variables in nonparametric frontier models: A probabilistic approach. *Journal of Productivity Analysis*, *24*(1), 93–121.

[18] De Witte, K., & Kortelainen, M. (2013). What explains the performance of students in a heterogeneous environment? Conditional efficiency estimation with continuous and discrete environmental variables. *Applied Economics*, *45*(17), 2401–2412.

[19] De Witte, K., & López-Torres, L. (2017). Efficiency in education: A review of literature and a way forward. *Journal of the Operational Research Society*, *68*(4), 339–363.

[20] DiMaggio, P. J., & Powell, W. W. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *American Sociological Review*, 147–160.

[21] Egorov, A., Leshukov, O., & Froumin, I. (2019). Regional flagship university model in Russia: Searching for the third mission incentives. *Tertiary Education and Management*, 1–14.

[22] Emerson, R. M. (1976). Social exchange theory. *Annual Review of Sociology*, *2*(1), 335–362.

[23] Färe, R., Grosskopf, S., Lovell, C. K., & Pasurka, C. (1989). Multilateral productivity comparisons when some outputs are undesirable: A nonparametric approach. *The Review of Economics and Statistics*, 90–98.

[24] Fay, D. L., & Zavattaro, S. M. (2016). Branding and isomorphism: The case of higher education. *Public Administration Review*, *76*(5), 805–815.

[25] 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.

[26] Haelermans, C., & De Witte, K. (2012). The role of innovations in secondary school performance–evidence from a conditional efficiency model. *European Journal of Operational Research*, *223*(2), 541–549.

[27] Hanushek, E. A. (1987). Educational production functions, In *Economics of Education*. Elsevier.

[28] Helfat, C. E., & Peteraf, M. A. (2003). The dynamic resource-based view: Capability lifecycles. *Strategic Management Journal*, *24*(10), 997–1010.

[29] Hoxby, C. M. (2009). The changing selectivity of American colleges. *Journal of Economic Perspectives*, *23*(4), 95–118.

[30] 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.

[31] Johnes, J., & Taylor, J. (1990). *Performance indicators in higher education: UK universities*. Open University Press; the Society for Research into Higher Education.

[32] Kounetas, K., Anastasiou, A., Mitropoulos, P., & Mitropoulos, I. (2011). Departmental efficiency differences within a Greek university: An application of a DEA and Tobit analysis. *International Transactions in Operational Research*, *18*(5), 545–559.

[33] Marginson, S., & Considine, M. (2000). *The enterprise university: Power, governance and reinvention in Australia*. Cambridge University Press.

[34] Meeusen, W., & van Den Broeck, J. (1977). Efficiency estimation from Cobb-Douglas production functions with composed error. *International Economic Review*, 435–444.

[35] Meyer, J. W., & Rowan, B. (1977). Institutionalized organizations: Formal structure as myth and ceremony. *American Journal of Sociology*, *83*(2), 340–363.

[36] Monk, D. H. (1989). The education production function: Its evolving role in policy analysis. *Educational Evaluation and Policy Analysis*, *11*(1), 31–45.

[37] Niehüser, W. (2008). Resource dependence theory: How will does it explain behavior of organization. *Management Revue*, *9*, 9–32.

[38] Pfeffer, J., & Salancik, G. R. (2003). *The external control of organizations: A resource dependence perspective*. Stanford University Press.

[39] Powell, K. K., & Rey, M. P. (2015). Exploring a resource dependency perspective as an organizational strategy for building resource capacity: Implications for public higher education universities. *Management in Education*, *29*(3), 94–99.

[40] Pruvot, E. B., Claeys-Kulik, A.-L., & Estermann, T. (2015). Strategies for efficient funding of universities in Europe, 153–168.

[41] Racine, J. (1997). Consistent signicance testing for nonparametric regression. *Journal of Business and Economic Statistics*, *15*(3), 369–379.

[42] Rothschild, M., & White, L. J. (1995). The analytics of the pricing of higher education and other services in which the customers are inputs. *Journal of Political Economy*, *103*(3), 573–586.
[43] Scott, W. R. (1987). The adolescence of institutional theory. *Administrative Science Quarterly*, 493–511.
[44] Simar, L., Lovell, C. K., & van den Eeckaut, P. (1994). Stochastic frontiers incorporating exogenous influences on efficiency. *Discussion Papers*, *9403*.
[45] 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.
[46] Wolszczak-Derlacz, J., & Parteka, A. (2011). Efficiency of european public higher education institutions: A two-stage multicountry approach. *Scientometrics*, *89*(3), 887.
[47] Xue, M., Harker, P. T. Et al. (1999). Overcoming the inherent dependency of DEA efficiency scores: A bootstrap approach. *Unpublished Working Paper, Wharton Financial Institutions Center, University of Pennsylvania.*
[48] Yu, C. (1998). The effects of exogenous variables in efficiency measurement—a Monte Carlo study. *European Journal of Operational Research*, *105*(3), 569–580.
[49] Zinchenko, D., & Egorov, A. (2019). Efficiency modeling of Russian universities. *HSE Economic Journal*, *23*(1), 143–172.

Contact details and disclaimer:

Aleksei Egorov
Laboratory for University Development of Institute of Education, National Research University Higher School of Economics, 101000, Moscow, Russia
E-mail: aegorov@hse.ru