SERIES: ECONOMICS

Tommaso Agasisti[1], *Ekaterina Shibanova*[2], *Daria Platonova*[3], *Mikhail Lisyutkin*[4]

*Corresponding author: eshibanova@hse.ru

THE RUSSIAN EXCELLENCE INITIATIVE FOR HIGHER EDUCATION: AN ECONOMETRIC EVALUATION OF SHORT-TERM RESULTS[5]

This research studies the short-term effects of the Russian Excellence Initiative Project 5to100 on participating universities. To trace the effect, we develop a quasi-experimental econometric methodology. A control group of universities comparable to the Project 5to100 universities at the starting point of the program's implementation was singled out using propensity score matching. Data envelopment analysis was conducted, and the Malmquist productivity index was calculated to trace how and why the efficiency of the "participants" and "non-participants" of the Project 5to100 has changed due to the project. We also investigate the direct impact of the policy on the research productivity of universities, using the average treatment effect, and difference-in-difference approaches. The final step consists of an explanatory analysis of the factors apart from the policy potentially affecting efficiency scores. We find statistically significant positive effects of the policy both on the productivity and on the efficiency of the participating universities.

JEL Classification. I23, I28

Keywords. efficiency in higher education; excellence initiative; management of universities; data envelopment analysis; Tobit regression; Malmquist index

[1] Politecnico di Milano School of Management, Associate professor, tommaso.agasisti@polimi.it
[2] National Research University Higher School of Economics, Laboratory for University Development, Intern researcher; E-mail: eshibanova@hse.ru
[3] National Research University Higher School of Economics, Laboratory for University Development, Laboratory head; E-mail: dplatonova@hse.ru
[4] National Research University Higher School of Economics, Institute of Education, Research fellow; E-mail: mlisyutkin@hse.ru

[5] **Acknowledgements**

We are very thankful for fruitful discussion and valuable comments on the previous version of this paper at the following conferences and workshops: IPA Higher Education Symposium for Early-Career Scholars (Moscow, 2017), 5th Workshop on Efficiency in Education (Budapest 2017; Huddersfield 2018), CHER Conference (Moscow, 2018), Conference Higher Education in Modern Ecosystems: Efficiency, Society and Policies (Augsburg 2018). All eventual errors are our own responsibility. This article 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).

1. Motivation and research questions

Global competition in higher education (HE) has had a great influence on the priorities of national governments in recent years (Chirikov 2016). The global ranking of universities has become a powerful tool, influencing the perception of success and excellence in higher education at national and institutional levels (Hazelkorn 2011, 2014, Altbach et al. 2017). Many countries have launched programs to develop a group of so-called ‘world-class’ universities (Altbach & Salmi, 2011). Such policies are known as Excellence Initiatives (ExIn), and are aimed at pushing particular higher education institutions (HEIs) to compete successfully in international education and research markets. Since 2000, more than 40 excellence-driven initiatives have been launched in more than 20 countries. More than US$60 billion has been invested in these initiatives (Salmi 2015).

Inspired mostly by the success of the Chinese ExIn, Russia initiated Project 5to100 in 2012. The basic aim of the project is that at least five Russian universities will be among the world’s top 100 universities according to key world university rankings by 2020. Fifteen universities were selected on a competitive basis in 2013 and have been receiving additional funding each year.

ExIns in higher education are mostly aimed at boosting the production of the key output, that is the international research intensity of universities (Salmi 2012). A number of academic papers have evaluated the changes in publication productivity for Chinese universities (Zhang et al. 2013, Zong and Zhang 2017, Yang and You 2017, Yufang 2017), German universities (Möller et al 2016, Klarl et al. 2016), Russian universities (Turko et al. 2016, Poldin et al. 2017), and Korean universities (Seong et al. 2008, Shin 2009). The design of recent studies is quasi-experimental. Zong and Zhang (2017), Yufang (2017) and Poldin et al. (2017) compare results for participants and non-participants of ExIn. Zong and Zhang (2017) and Klarl et al. (2016) use difference-in-difference (DID) models. Yufang (2017) develops Propensity Score Matching (PSM) model (specifically nearest neighbor matching) to evaluate the effects of Project-211.

ExIns are not only funding programs to achieve higher output in terms of research productivity; they also aim to bring organizational transformations to institutions and institutional environments. These include changes in the internal activities and efficiency of universities. To the best of our knowledge only two studies assess the effects of ExIn through the evaluation of university efficiency. Gawellek and Sunder (2016) examine effects of the German ExIn on the efficiency of universities using data envelopment analysis (DEA), the Malmquist index and its decomposition. Yaisawarng and Ng (2014) also use DEA and compute the annual efficiency scores to test if Chinese Project-211 universities perform better than non-211 universities, and compute the Malmquist index to examine whether productivity changes and technological advancement took place over a three-year period.

The main objective of this study is to estimate the changes in the activities of Project 5to100 universities. There are two key research questions:

- What are the effects of the participation in ExIn on university performance?
- Have there been any changes in the efficiency of the participating universities since the implementation of the project?

A striking inadequacy among studies evaluating the effects of ExIns on university activities is the lack of evaluation of performance and efficiency score changes between participant and non-participant groups solving the attribution challenge using PSM. This study addresses this gap by deploying an innovative quasi-experimental design to study the effects of ExIns. We employ an empirical analysis based on five steps:

Step 1: Single out HEIs comparable to excellence-driven universities at the start of the program which are not part of ExIn, using PSM to create a control group.

Step 2: Check whether after the launch of Project 5to100 there is a significant difference between the control and treatment groups in their key performance indicator – publication activity, by means of quasi-experimental techniques (DID and PSM).

Step 3: Estimate the efficiency of universities participating in the excellence initiative and their control group over a 5-year period (2012–2016) using DEA.

Step 4: Estimate the dynamics of the efficiency of universities measured using the Malmquist productivity index, and its decomposition into (i) efficiency change due to internal change and (ii) technical change due to the overall shift of the efficiency frontier.

Step 5: Identify the explanatory factors at an organizational level, which could potentially affect university efficiency in addition and beyond ExIn.

We observe positive and statistically significant effects of the policy on the efficiency and productivity of universities. The policy's impact on research performance on the participating universities is also positive and significant. We also trace the positive effects of research and development activity, the homogeneity of the academic preparation of students and autonomous status on the efficiency of universities.

The paper is organized as follows. The first section discusses ExIn as a global phenomenon and provides information about Project 5to100. The second section provides a literature review about the effects of ExIn on university performance, the mechanisms/sources of performance changes and the policy background. The third section discusses the performance and efficiency of Russian HEIs on the basis of previous research. The fourth section describes the methodology of the five-step analysis: PSM, DID, DEA, the Malmquist productivity index and Tobit regression. The key results of the research are presented in the fifth section.

2. Excellence Initiatives: Aims and Design

2.1. Excellence Initiative as a Worldwide Movement

Aiming global competitiveness, excellence and a world-class level of higher education, national policies use different tools such as mergers, the establishment of new universities and upgrading existing ones (Shattock 2017). Excellence initiatives pursue ambitious goals in changing national universities. Governments push universities to become more visible agents of international education and research, and to dramatically improve their productivity measured by the objectives set by governments (Froumin, Lisyutkin, 2015, Salmi 2016).

The particular aims of the initiatives vary greatly from country to country. A comparative analysis of ExIns launched by different countries allowed us to single out some practical rationales for their implementation. First, some governments consider universities to have a direct economic impact producing a significant part of the national GDP, by selling educational services. Secondly, as universities attract foreign students and better professors, the rationale for an ExIn is to develop future leaders of innovative economies and make a greater contribution to the development of human capital. Thirdly, governments focus increasingly on research development as a part of public policy. ExIns promote a more or less universal model of the research university (Mohrman, Ma and Baker 2008). The German ExIn supports this, largely with regards to its support of advanced research at universities. The French ExIn was the part of the "Investments for the Future" program to create multidisciplinary education and research centers complying with global standards (Gesson, 2015). The Chinese projects "211" and "985" have been discussed extensively. Project-211 supported around a hundred leading Chinese universities in training elite specialists for the implementation of national economic and social development programs (Lixu, 2004). Project-985 was launched in China in cooperation with local governments to propel a group of leading universities to a world-class level (Litao, Jinjing, 2010).

2.2. Russian Excellence Initiative: The Design of Project 5to100

In 2012 Russia joined the race for global competitiveness in higher education by launching Project 5to100. The basic idea of the project was that at least five Russian universities would enter the world's top 100 universities (according to the international rankings – ARWU, THE, QS) by 2020. This initiative gives governmental support (both financial and managerial) to help stimulate academic excellence and the global competitiveness of particular HEIs to the level of the best universities in the world. The implementation of Project 5to100 and the assessment of its results is one of the most important political issues in HE, as a significant amount of money is being invested in a relatively small share of the public HEIs.

In 2013, the first year of the project's deployment, 15 universities were selected on a competitive basis and have been receiving additional funding. The group of the "5to100 universities" was enlarged to 21 universities at the second stage of the project in 2015. Universities aspiring to become part of Project 5to100 prepared detailed road maps for their development by 2020 according to criteria established by the government. The road maps anticipated the achievement of specific performance indicators set by the government. An International Council was established by the government to make the road map assessment more objective and to evaluate the road maps according to international standards. In general, the approaches and design of the Projects is very similar to the Chinese Project-985 (Chirikov 2018 in press).

During the five years in which the project was being implemented, the 5to100 universities have received more than 50 billion rubles (about US$850 million) from the federal budget. In relative numbers, the annual project subsidy is only 2% of the federal budget going to higher education. The subsidy within universities differs according to their achievements. In 2017 the disparity in the subsidy for high-achievers was twice that for low-achievers (within 15 universities of the first wave).

The characteristics of 5to100 universities vary dramatically in scale. In terms of budget, they differ eight-fold. Project funding does not play a crucial role in all universities. For some it makes up about 28% of overall funding, but for others only 5%. From this perspective, it is also likely that the effects generated by this additional money might be heterogeneous across institutions, because of their different sizes and structures (for example, the different mix of disciplines and departments). Almost all the 5to100 universities operate in STEM fields.

Despite the differences, most of the universities involved in the project have been successful in achieving their principal goals. The dynamics are positive in terms of their positions in international university subject rankings. In 2017, 15 universities included in Project 5to100 entered the subject rankings of the QS World University Rankings. These subject rankings include "physics and astronomy", "electronic engineering", "mathematics", "information technology", "chemistry", and "engineering, aerospace and industrial engineering". Five universities are already among the top 100 universities of the world, according to QS subject rankings already (Table 1).

[Table 1] here

The Ministry of Education and Science (MoES) has implemented several other projects of large-scale federal support since 2000 (Platonova and Semyonov 2018). The Russian Universities Program (2008) supported the renovation of research equipment at the best national universities with a large proportion of the advanced equipment purchased from abroad. The extensive network of 10 Federal Universities and 29 National Research Universities appeared in 2006–2012, with an active internationalization agenda and project management approach – strategic development programs, KPI, etc. The majority of universities that then participated in Project 5to100 also participated in these initiatives. Since the launch of Project 5to100 in 2012, the issue of the global competitiveness of Russian universities has been at the top of the agenda. Besides, in 2009 the oldest and largest universities – Moscow State University and St. Petersburg State University – received special status and special funding for their development.

3. Conceptual framework: How Do Excellence Initiatives Change University Activities and Their Performance?

3.1. Mechanisms of Change

The majority of studies determine several basic sources of the changes in HEI performance due to ExIn. The selective investment and the concentration of funds aims to improve research and boost research outcomes. The top-down national policy promotes scientific research and identifies priorities and very specific outputs. The investments change the inputs: building ‘hardware’ and infrastructure, training high-level postgraduates and faculty members and attracting the most talented students (Huang 2015). The promotion of internationalization is not limited to attracting leading researchers from abroad, but also changing the processes and standards, such as research oriented to international agendas and publications in internationally referred journals. The changes in activities can appear due to more feasible competitive mechanisms (Seong 2008) and the effects of public announcements (Klarl et al. 2016).

These changes of processes and standards are key transformations of universities in ExIns. To fulfill the government’s targets and expectations universities employ different strategies. Chirikov (2018 in press) conducted research based mostly on interviews with government officials and the administration and faculty in the Russian excellence-driven universities in Moscow. The interviews cover strategy transformations influenced by various definitions of global competition in HE established by Project 5to100. The author identifies four definitions of global competition determining different strategies: to be at the top of global rankings, to develop the right institutions, to be capable of radical change, and to engage in evolutionary transformation. The study revealed that there are four sets of mechanisms for organizational transformation to respond to the different definitions of global competition: paralleling, power play, imitation, and gaming (see Figure 1). In the short term all of the strategies can be successful in reaching the targets and producing higher outputs, however in the longer term the efficiency of activities can also worsen in some cases.

[Figure 1] here

Although the aims and general results of ExIns are much debated (Salmi 2012, 2016) the estimations of how universities change their performance and activities due to participation in ExIns are limited (see the case studies in Altbach et al. 2016). The key reason is the duration of programs and their delayed effect (the time challenge). Real university modernization takes many years, at least eight to ten (Salmi 2012), and most initiatives were implemented only recently. Moreover, the design of a study evaluating the effects of the policy should take into account the objectives of the policy, changing contexts, and other factors affecting university changes (the attribution challenge) (Yufang 2017). Empirical evidence mainly addresses the changes in publication activity, but there is some literature on the estimation of efficiency transformations. These provide an argument for conducting an evaluation of the short-term results of the policy in Russia.

3.2. Potential effects on publication activity

ExIns are mostly aimed at upgrading the research capacity of universities (Salmi 2009, 2016), a number of papers evaluate changes in university publication productivity. The study of Chinese universities participating in Project-985 by Zhang et al. (2013) found that the rate of growth of publications for universities as a whole increased more quickly after the implementation of the ExIn. A quasi-experimental study based on DID conducted by Zong and Zhang (2017) to evaluate the effectiveness of Project-985, using 15-year panel data (1998–2013) also showed that the Project-985 had a positive effect on publication output of participating universities. Yufang (2017) evaluated the effects of another Chinese excellence initiative in higher education, Project-211. Using PSM, Yujang found that the project had a positive effect on the research of participant universities, increasing the number of the academic papers published and research projects implemented by the universities. Yufang also tries to solve the time challenge covering the period before, during and after the Project (6 years). He finds that Project-211 had positive effects on university research and teaching performance but negative effects on university service performance (income from technology transfer). Using a survey of 30 faculty members from 30 public universities in China, Yang and You (2017) showed Projects 985 and 211 had a positive effect on international publications, but also identified that they had no significant impact on domestic publications and technology transfer.

Shin (2009) evaluated the effects of South Korea’s Brain Korea 21 (BK 21) project by examining the frequency of article publications in SCI journals 1995–2005. It was revealed that the growth of research publications by Korean research universities intensified significantly following the implementation of the Korean excellence initiative in 1999.

The results of the first German ExIn by Klarl et al. (2016), based on quantitative and qualitative measures (1998–2012) suggest that the improvements were not due to ExIn, but to the announcement of the policy which "triggered diverging performance paths within the German higher education system, thus positively contributed to augmented research performance of the promoted universities". Bibliometric analysis conducted by Möller et al. (2016) to assess the effects of the German ExIn showed that the program succeeded in concentrating excellent research and intensifying collaboration between HEIs and the non-university research sector.

Some research projects assessing the effects of ExIns are mostly based on a quasi-experimental design. For example, Zong and Zhang (2017), Yufang (2017) and Poldin et al. (2017) compare the results for participants and non-participants of excellence initiatives by means of econometric methods.

There a few studies on the effects of Project 5to100. Analysis by Turko et al. (2016) identified that Project 5to100 had a positive impact on publication productivity and enhanced the growth of global competitiveness expressed by promotion in the rankings. Applying mixed method models, Poldin et al. (2017) confirmed that the relationship between the number of publications in general and publications in high-impact journals and participation in ExIn was positive and significant for two the years after the program began. However, as the authors mention, for more sustainable results a longer period should be taken into consideration.

3.3. Potential effects on efficiency

ExIns are not only funding programs to achieve higher output in terms of research productivity, but they are also aimed to bring organizational transformations of institutions and institutional environment. Gawellek and Sunder (2016) examine the effects of the German ExIn on the efficiency of universities using DEA, the Malmquist index and its decomposition components, regressed on a set of dummy variables, indicating whether a certain university applied for the program and won the subsidy in 2006 or 2011. The results mainly suggest that applying for the program was expensive for universities, as they had to risk a large amount of resources. They lost considerably in efficiency and productivity, but those who finally received the grant, successfully recovered in productivity and efficiency in the second period of the analysis (2006–2011), especially if compared to their losses due to the application in 2001–2006. The authors state that the participants managed to show efficiency and productivity growth due to additional funding within the program, but the main concern is whether these universities will continue their movement toward excellence, as the funding is limited in time.

Yaisawarng and Ng (2014), in the context of Chinese HE, use DEA and computed the annual efficiency scores to test if the Project-211 universities perform better than non-211 universities; and compute the Malmquist index to examine whether productivity changes and technological advancement took place over a three-year period. They found a positive effect of participation in Project-211, however the best non-Project 211 universities, despite limited resources, also show productive dynamics.

4. What do we know about performance of Russian universities?

The Russian HE system is one of the largest in the world. In 2017 more than 4.4 million students were studying at 502 public HEIs with 480 satellites and 266 private HEIs with 171 satellites. These universities operate with relatively high financial constraints. Only about 1.61% of public spending goes on HE (Roskazna 2017). Higher education accumulates US$13,3 billion, 44% of which comes from non-public resources and 53% from the federal budget (MoES, 2016). The ongoing trend is a concentration of public funding in leading universities (Abankina et al. 2017). Since 2012 the MoES policy to develop new public management mechanisms in HE, including performance-based accountability and funding. Thus, the MoES launched the monitoring of HEI performance. This tool provides 8 indicators for university activities to determine under-performing universities.

Despite this policy agenda, research about Russian HEI performance and efficiency is still very limited. A number of studies show a high level of system stratification in terms of available resources (Lisyutkin 2017, Abankina et al. 2018) and types of activities (Platonova and Semyonov 2018, Smolentseva et al. 2018). Although there is intense international discussion about university efficiency (e.g. Wolszczak-Derlacz, 2014; Johnes and Johnes, 2015; Agasisti and Johnes, 2015; Lee and Worthington, 2016; Sagarra, Mar-Molinero, Agasisti, 2017), there are few Russian studies that provide empirical evidence of efficiency.

Some papers (e.g. Abankina et al. 2013, 2016) develop a typology of Russian HEIs (for 2010) and verify the results by assessing university efficiency scores. They identify two groups of universities with the highest efficiency scores (‘market leaders’ and ‘universities of good standing’). Zinkovsky et al. (2016a, 2016b) present an analysis of efficiency changes of the universities which merged with vocational schools and colleges. The results show positive trends for organizational agglomeration. Gromov (2017) studies what kind of returns of scale Russian universities have and evaluates the factors influencing the efficiency of 120 universities, 2012–2014. The study shows that 55% of HEIs operate with decreasing returns to scale, 36% operate with increasing returns to scale and 9% of institutions face constant returns to scale. The number of HEIs with increasing returns to scale increased during the period. Productivity growth was revealed to be the result of a frontier shift, not individual efficiency growth.

5. Methodological and Data

This study develops a five-step analysis to assess the effects of ExIn and efficiency changes of participating universities. We use data from the monitoring of performance of Russian HEIs, conducted by the MoES (for a detailed description of the Monitoring see Sokolov and Tsivinskaya, 2018a, 2018b). The data covers a 5-year period, from the 2012/13 to 2016/17 academic years. The descriptive statistics for each stage of the analysis are presented in Table 2.

Initially, there were 543 higher education institutions in total. Taking into account the official requirements for potential participants in the ExIn and data availability, we implement the following limitations to our data sample:

- there must be publicly funded students within the university’s educational programs;
- the minimal overall number of students enrolled in the educational programs is 4,000;
- the minimal Unified State Exam (standardized national entry exam) grade must be equal to or greater than 64.

We also excluded all universities which had been reorganized during the period (2012–2016), to avoid any bias caused by the structural differences of the units. We also excluded universities participating in the excellence initiative at its second stage. After implementing the limitations, we had a sample of 125 universities, including the 15 first participants of the ExIn program. The total sample used for PSM is 125 universities. After performing the PSM we have a total of 30 universities, 15 participants and 15 in the control group. Only these 30 universities are analyzed at the DiD, DEA, Malmquist index and post-estimation stages.

[Table 2] here

At the first step, we implement a PSM procedure with a twofold objective. PSM is used to assess whether universities participating in the ExIn performed at higher levels than the control group, namely those institutions that are similar to them for several observable characteristics. The propensity scores also guide the selection of the control group for checking the dynamics of university efficiency scores, so that efficiency of ExIn institutions over time can be observed in parallel with the controls.

At the second step, we evaluate the ExIn's direct impact on the research activity of the participating universities, comparing them with the control group by applying two econometric techniques: the average treatment effect based on PSM and DID.

At the third stage we employ both a traditional and a bootstrap DEA technique to measure efficiency scores over a five-year period, including the "zero" year (2012/13 academic year), i.e. the year of selection for the program without any managerial and financial changes, later followed by a period of implementation of the ExIn, i.e. 2013/14–2016/17 academic years. The idea here is to understand whether efficiency scores of 5to100 universities are higher than those of the control group, in a static comparison based on year-by-year data. We apply a bootstrap procedure (Daraio & Simar 2007) to ensure our model's fit and a robustness check procedure as well.

The fourth step explicitly takes into account how the productivity index of universities varied over time, and we decompose this dynamic into different causes (pure efficiency gains/losses and a modification of the efficiency frontier) by means of the Malmquist Index procedure.

Lastly, the fifth step includes an analysis of the efficiency determinants which are not in full control of university managers in the short term, for example, the legal status of the organization. The aim of this analysis is to take into consideration all the potential factors that affect efficiency beyond the implementation of the policy. Details about the steps are provided in the next subsections.

We note the importance of the Stable Unit Treatment Value Assumption (SUTVA) (Rubin, 1978) which imposes the absence of direct interaction of the unities of our analysis. University activities are open to the market, but we consider the treatment effect to be pure, because the scope of the treatment was strictly limited to the participating universities. Even if we assume that the policy was more widespread and occasionally affected the control group, we follow Imbens & Wooldridge (2009) and consider the indirect effects much smaller than the direct treatment; we assume that the SUTVA condition is not violated in our research.

Step 1. Propensity Score Matching: control group selection

One of the most popular methodologies to assess government programs is PSM (Rosenbaum and Rubin, 1985). This is defined as the probability of a treatment that is conditional on a set of observable variables:

$$propensity\ score_i = probability(D_i = 1|X_i)\ (1)$$

where D_i is the status of the treatment (0 if not present and 1 if present) and X_i is a set of observed covariates. Matching entities from the treatment group with entities from the control group is needed to construct the control group. As our study is an observable one, the true propensity score is unknown, but we can calculate it through a logistic model, 0 or 1 being the outcome of treatment assignment. The propensity score, which describes the probability of being in the treatment group for entities, must be calculated according to their characteristics. A correctly conducted randomization leads to a balanced distribution of indicators between the groups. PSM consists of a set of the following steps:

- the selection of variables potentially influencing the probability of entering the ExIn program and a statistical verification of this choice;
- calculations of propensity scores for each university within the sample;
- using variables to balance control between the matched pairs;
- an evaluation of the average effect on the treated based on the matched samples.

One or more matching universities are place in correspondence with each of the universities participating in the ExIn. Every time this procedure is realized, a trial is performed, after which we can measure, how the treatment influenced the participant and how this effect differed from the control university. This approach is frequently used in medical and social studies: an experimental group is matched with a control group to measure the efficiency of a certain therapy or policy. The basic methodology of our research design suggests that only one type of treatment exists, i.e. there is no subdivision of treatment according to amounts of financing within the program.

According to the government regulations of Project 5to100, all the participating universities are obliged to fulfill a certain set of objectives within the framework of the program. Only the activities related to the following objectives are additionally financed by the government within the framework of the project.

- Human resource development and internationalization (the formation of a management pool; attracting managers, young research and educational staff with work experience in leading foreign and Russian universities and research institutions; mobility programs for research and educational employees).
- The development and internationalization of the educational activities (the introduction of new educational programs conjointly with leading foreign and Russian universities and research organizations; the development of post-graduate and doctoral programs).
- The development of research activities (the implementation of fundamental and applied research projects in partnership with leading research and industry organizations).
- The internationalization of the student body and attracting talented students (students support programs; attracting students from leading foreign universities to study in Russian HEIs).

We combine the indicators reflecting the changes expected by the government (Table 3). Most of indicators that we use are valid for the ExIn, although we do not use all the indicators listed in the official documents as not all of them are quantitative and therefore observable. We believe them to be the best as they represent the two key fields of activities – internationalization and research productivity – and reflect the overall economic context in which the universities perform. The outcome in the PSM model also reflects the framework of the Russian ExIn and is expressed by the number of publications indexed in Web of Science and Scopus. We chose a propensity score nearest neighbor matching 1:1 procedure to be the best in a stage-by-stage selection process. Matching each participating university with another specific one to prove that there are no structural difference effects in the program outcome and performance is crucial. We build a PSM relying on a range of parameters:

- methodological (used by the government to select the participant);
- environmental (such as regional development indicators);
- structural-related variables (such as the share of STEM and medical students).

[Table 3] here

Additional assumptions must also be taken into account. The Conditional Independence Assumption (CIA) implies that selection is solely based on observable characteristics and that all variables influencing the treatment assignment and potential outcomes simultaneously are observed by the researcher. To satisfy the CIA, the values of the characteristics before treatment should be used so that we can be sure all the universities had an equal chance of being chosen. Common Support ensures that entities with the same characteristic value have a positive probability of being both participants and non-participants. The Common Support Problem means that there can be entities without a matching pair. If the number of such entities is low, they can be dropped. To anticipate, we do not find evidence of problems in the validity of CIA. We conduct the PSM using data collected before the actual implementation of Project 5to100, i.e. we use the data for the 2012/13 academic year, following (López-Torres, Prior, & Santín, 2016)

Step 2. Impact on research productivity

The first quantitative analysis of the ExIn uses two different techniques. PSM can be applied to calculate the 'average treatment effect on the treated' (ATT). For our topic of interest, the selection of universities as participants of Program 5to100 is the 'treatment', and selected universities are 'treated':

$$\tau_{ATT} = E(\tau|D = 1) = E[Y(1)|D = 1] - E[Y(0)|D = 1] \quad (2)$$

where τ is a treatment indicator, and $[Y(1)|D = 1]$ and $Y(0)|D = 1]$ are potential outcomes for entities of the treated group if they receive treatment and do not receive treatment respectively. The latter is not observed and the methodology for the selection of a substitute is needed.

DID is a popular method for accessing the impact of reforms in health care (Dimick & Ryan 2014) and education (Pedraja-Chaparro 2016; Lopez-Torres et al. 2016). The following empirical model is being estimated with a simple OLS:

$$Y_i = \alpha + \lambda * D_i + \delta * D_i * year + \gamma * year + u_i \ (3)$$

where Y_i is the outcome variable used to estimate the ExIn effect and D_i is the treatment status. For the research activity indicators, we choose the number of citations and publications indexed in Scopus and Web of Science databases (per 100 faculty members).

Step 3. Measuring university efficiency

Measuring university performance using DEA is common practice. Worthington (2001), Johnes (2004) and Lopez-Torres & De Witte (2015) illustrate an overview of using frontier efficiency applications in education, by means of useful surveys of the literature.

DEA, a non-parametric linear programming method, provides a measurement of the efficiency and productivity scores of a Decision Making Unit (DMU) – universities in our case. DEA is based on a programmed envelopment of observed multiple input-output vectors (Boussofiane, Dyson & Thanassoulis 1991) without additional issues of data distribution. DEA is suitable for efficiency estimation to consider a multi-input and multi-output production function in the absence of all the market prices of the components (Ray 2004). The efficiency of each DMU is measured through the changing proportion of inputs or outputs. A DEA model can be input- or output-oriented, depending on whether a minimization or a maximization problem is being solved, it can also be modified depending on a constant or a variable return to scale. An output-oriented DEA model is produced to test if a DMU is capable of increasing its outputs with the same inputs. In our research, an output-oriented DEA will be used for the calculation of efficiency scores.

The most debated point in measuring the efficiency of universities deals with different ways to define the variables used as inputs and outputs, as this provides a certain freedom in determining the design of an efficiency measuring model (De Witte & López-Torres 2017). Within the Russian context the diversity of possible input and output combinations is also present. Abankina et al., (2015) use two separate models to estimate the teaching and academic potential of Russian universities and do not take into consideration financial resources. Gromov (2017) presents a model with finance-based inputs and both teaching and research indicators as outputs.

In practice, the choice of indicators for the assessment should not be limited to those available, nor should it use 'traditional' indicators. Instead, the evaluation procedure should be informed by the discretionary choices set by policy-makers. The present study is limited to assessing the main effects of the policy over a five-year period and it assumes that traditional efficiency measurements approximate the main goals set by policy-makers for improving the Russian HE system.

Achievement assessment at the national level is mostly related to international competitiveness indicators. These indicators include world university rankings and additional performance criteria set by the governments reflecting the universities global position also (e.g. an increase in the quality of incoming students, new (mainly research) facilities, more international partnerships (Hazelkorn 2007; Froumin & Salmi 2013).

The specific model used in this paper is shown in Table 4. We expect each variable to reflect a certain area of university resources (capital, size or labor) and performance (teaching, research or rankings).

[Table 4] here

The universities face a technology which can be realized through a combination of inputs (x) and outputs (y), or a production possibility set:

$$PPS = \{(x, y): (x; y) \in R_n^+; y \in R_m^+\} \ (4)$$

The combination of inputs and outputs is feasible only for the efficiency frontier of the production possibility set. For the output-oriented model, the technical efficiency is

$$Max\ \varphi_k + \varepsilon \sum_{r=1}^{s} s_r + \varepsilon \sum_{i=1}^{m} s_i \ (5)$$

under the condition $\varphi_k y_{rk} - \sum_{j=1}^{n} \lambda_j y_{rj} + s_r = 0, r = 1, \dots s,$ (6)

$$x_{ik} - \sum_{r=1}^{n} \lambda_j x_{ij} - s_i = 0, i = 1, \dots m, \ (7)$$

$$\sum_{j=1}^{n} \lambda_j = 1, \ (8)$$

$$\lambda_j, s_r, s_i \geq 0\ \forall j = 1, \dots, n; r = 1, \dots, s; i = 1, \dots, m \ (9)$$

Where *s* are outputs, *m* are inputs, y_{rk} is the volume of output of type *r*, belonging to university *k*, $x_{ik}$is the volume of input of type *i*, belonging to university *k*, s_i and s_r are the slack in outputs and inputs respectively. The efficiency rate of university k is defined as $\tau_y = \frac{1}{\varphi_k}$; university k is efficient, if the efficiency rate $\varphi^* = 1$ and there is no slack in the volumes of inputs and outputs. If $\varphi^* = 1$, then the university under evaluation is a frontier point. i.e., there is no other universities that are operating more efficiently than this particular one.

This analysis provides efficiency scores in terms of input used to produce outputs as efficiently as possible. The findings of this stage will be discussed in the results section, but the statistics we derived from the DEA analysis are used to finalize our research in the Malmquist index for productivity measurement. A variable return to scale assumption is implemented for both DEA efficiency estimations and the Malmquist productivity index, as previous research indicates Russian universities operate under variable returns to scale (Gromov, 2017).

Step 4. The Malmquist productivity index

We measure the Malmquist productivity index, showing the change in university productivity over time. The methodology is based on DEA and can be modified for input-orientation or output-orientation. This index is measured as a ratio of two distance functions representing the efficiency performance in two different (equation 5 for period *t* and 6 for period *t+1*) time periods (Lee, Leem, 2011):

$$\text{Malmquist Index}_t = \frac{E^t(x^{t+1}, y^{t+1})}{E^t(x^t, y^t)} \quad (10)$$

$$\text{Malmquist Index}_{t+1} = \frac{E^{t+1}(x^{t+1}, y^{t+1})}{E^{t+1}(x^t, y^t)} \quad (11)$$

$$\text{MPI}_i = \left(\frac{E_i^t(x^{t+1}, y^{t+1})}{E_i^t(x^t, y^t)}\left(\frac{E_i^{t+1}(x^{t+1}, y^{t+1})}{E_i^{t+1}(x^t, y^t)}\right)\right)^{1/2} \quad (12)$$

Where *x* and *y* are the levels of output produced by university *i* and level of inputs used respectively in period *t* or *t+1*. $E_i^{t+1}(x^t, y^t)$ is the production frontier that could be achieved by the combination of inputs used and outputs produced in period *t* if operating under technology in period *t+1*. $E_i^t(x^{t+1}, y^{t+1})$ is the maximum output that could be produced in period *t* given the outputs and the inputs of period *t+1*. $E_i^t(x^t, y^t)$ is the actual production combination of inputs and outputs in period *t* under technology of the same period, as $E_i^{t+1}(x^{t+1}, y^{t+1})$ is for period *t+1*.

By extracting the geometric mean of the two indices we obtain the Malmquist productivity index for each university, where the first ratio is the technical efficiency change of university *i*, i.e. the closer the university approaches the frontier, the more efficient it is. The second ratio represents the technological change or the shift of the technology frontier and reflects the changes of the technology within the sample as a whole (i.e. something that affects the productivity possibly for all the universities):

$$MPI_i = \left(\frac{E_i^{t+1}(x^{t+1},y^{t+1})}{E_i^t(x^t,y^t)}\right) \times \left(\frac{E_i^t(x^{t+1},y^{t+1})}{E_i^{t+1}(x^{t+1},y^{t+1})} \times \frac{E_i^t(x^t,y^t)}{E_i^{t+1}(x^t,y^t)}\right)^{1/2} \quad (13)$$

Measuring the index, we can show whether the total productivity, the technical efficiency (or the pure efficiency change) and the technological change (or the technical change) of a certain university is increasing, decreasing or stagnating over time as the Malmquist index will be greater than, less than or equal to unity respectively.

Step 5. Explanatory analysis of DEA efficiency

Tobit regressions are often used as a second-stage technique to explain the DEA efficiency scores (McDonald 2009). Due to the panel structure of our data (30 universities and 5 years of analysis in total) we estimate the following random-effects panel regression model and apply a bootstrap procedure (Mooney & Duval 1993; Simar & Wilson 2007):

$$y_{it} = \beta X_{it} + \upsilon_i + \epsilon_{it} \quad (14)$$

where panels *i=1,…n, t = 1, …*n_i and υ_i are the random effects. As the maximum possible efficiency score is 1, we must apply a censoring constraint for the dependent variable:

$$y_{it} \leq y_{it}^* \quad (15)$$

the y_{it}^{*} being a latent unobservable term that equals 1.

The model specification relies on the external conditions of HE in the context of the Russian education system. However, it is based on operational indicators which reflect university activities apart from the ExIn roadmaps. Thus, we use the "autonomous institution" dummy variable: the status of an autonomous institution can be gained by a non-profit organization, and it gives higher level of freedom in managing financial resources. The gap between entry exam scores of tuition fee paying students and state-funded students measures the internal differentiation in terms of the academic preparation of students. "Share of full-time students" reflects the proximity to the model of a research university (in contrast to a teaching university). "Market share" is calculated as the share of a university's students in the total student body of the region (excluded from the final model due to bad fitness). "Proxy for Moscow/Saint Petersburg/Tomsk region" captures the fact of positive migration flows of 17-20 years olds to these cities. However, "market share" and "proxy variable for Moscow/Saint Petersburg/Tomsk region" did not fit the model and were excluded.

6. Results

6.1. Control group selection

To make the evaluation unbiased, we conduct the PSM to ensure that the treated and the control groups of universities are similar when considering their observable characteristics: we do not observe any difference between the two groups due to a successfully conducted PSM. Tables 5 and 6 show the results for the first and last years of the analysis, respectively.

Twelve (five non-participants and seven participants) of the 30 final items belong to the Moscow/Saint Petersburg region, most of the universities specializing in technical or engineering sciences. No structural differences were observed after the PSM. This will be later additionally proved via the lack of a statistically significant difference in efficiency scores among the two groups at the zero point of the analysis. In other words, the participants and the non-participants were not significantly different from each other in terms of their relative efficiency before the actual changes in the strategic and operational development. Thus, we can claim that while observing a difference in both the technical efficiency and the direct impact of ExIn on the research productivity of the universities after ExIn had started, we can be assured that this difference is caused by ExIn. The fulfillment of the common support condition was ultimate for the control group selection procedure. The obtained region of common support is [0.053, 0.998]

[Table 5] here

[Table 6] here

6.2. ExIn impact on publication activity: propensity score matching and difference-in-difference

As anticipated in the methodology section, to find the ExIn effect on research productivity two different methodologies were applied: the PSM-based average treatment effect on the treated and the DID approach. The results are presented in Table 7.

Average treatment effect on the treated

ATT is 130.18 ±27.37 publicationsindexed in Web of Science or Scopus per 100 faculty members. In terms of citations per 100 faculty members the observed effect was the following: 987.9 ±400.22 in favor of the participant group. Thus, on average universities participating in the ExIn had better research output than the non-participants 2013–16.

Difference-in-differences approach

The DID revealed similar results: the effects on both publications and citations are positive due to the time-treatment (i.e. ExIn) effect. The effect of the ExIn in terms of the number of publications by faculty members of participating universities is 90.5 for publications indexed in Scopus or Web of Science (per 100 faculty members), on average a faculty member of a university participating in the ExIn produces 0.9 more publications than her/his counterpart from the control group. As for citations, the effect is also positive: 833.00 more citations were indexed in Scopus or Web of Science (per 100 faculty members), or 8.33 more citations per faculty member. A comparative presentation of both methods' results is presented in Table 7. The results confirm the positive effects of the policy on the university productivity.

[Table 7] here

6.3. Efficiency analysis of Russian universities after the policy

The efficiency indices (Table 8) do not differ significantly among the two groups of universities over the whole period. In the 2012/13 academic year, participants and non-participants did not differ significantly in relative efficiency. We consider 2012/13 the zero year, in other words, time *t*=0 in which the policy had not yet been implemented. The following dynamics can be explained through the ExIn effect. On average, neither the participants nor the control group reached the efficiency frontier. At the ExIn's starting point a quick positive effect of the ExIn is observed in the participants. This can be partly explained by the fact that the zero year was a year of preparation for application as well as the year of the competition itself. A cumulative effect could be present. The second possible explanation for this quick first-year result could be related to both the intensive and extensive managerial shifts inside university structures.

The most important highlight of the efficiency analysis is related to the fact that we conduct a short-term measurement of a long-term program. However, even given this condition, we are able to observe a statistically significant difference in relative efficiency by the end of the observed period. ExIn universities are more efficient than their counterparts in the 2016/17 academic year and perform better than themselves in the same modeling conditions in the zero year (by 7.6 p.p.) and in the first year of the ExIn (by 5.8 p.p.) .

[Table 8] here

A closer look at the DEA efficiency scores in dynamics permits a better understanding of the difference between the two groups. There are some crucial highlights that can explain the drop (Fig. 2) in the growth of participant efficiency. First, all possible expected effects of the ExIn have delayed effects and they require time to settle. The ExIn can be regarded as a specific reputational signal in the HE system: participation in the project can improve a university's status but it can increase the competition among the universities, including non-participants.

The DEA-obtained efficiency scores are proved to be robust through a Pearson's r and significant and high correlations with the bootstrapped efficiency scores. A robustness check with a step-by-step exclusion of each of the model's parameters and a re-estimation is presented in the Appendix and confirms the findings. Finally, the Kolmogorov-Smirnov test proves the similarity of efficiency score distributions for the treated and control groups both in the zero year and the final year of our analysis. The difference in the means is tested using both parametric (t-test) and non-parametric (Wilcoxon signed-rank test) techniques (Table 9).

[Table 9] here

6.4. Malmquist Index dynamics: how productivity changed after the policy

The Malmquist index can be decomposed into efficiency change, measuring efficient modifications within the university, and technology change, measuring the change referring to the shift relative to the efficiency frontier. This means that the total productivity is explained through a combination of

- efficient managerial changes inside the universities (the efficiency change), i.e. to what extent universities are able to adopt changes in technology inside the small ExIn-control group economy;
- and the frontier shift or the technological change, affecting the system of universities.

During the period, the means of the productivity indices (including total productivity, efficiency change and technical efficiency), do not vary significantly between the two groups except for the last year (Table 10). Both participants and non-participants, on average, reach the frontier in terms of the total productivity factor at the beginning of the observed period, and improve their performance significantly over time.

[Table 10] here

The participants are those who tend to push the technological frontier and expand it, i.e. technological progress is observed over the period, while the non-participants tend to produce more outputs with the resources they possess. However, the total change in productivity through the observed period proves that the participants managed to increase their total productivity by more than 400%, which is mostly explained by the almost 370% increase in the frontier shift component and an 11.4% increase in efficiency change. The dynamic overlap of the DEA efficiency analysis and the Malmquist index shows that in the first to second year period (2013/14–2014/15 academic years) of the ExIn, the participants managed to considerably shift the frontier (by 116.4%) and to approach it in a more efficient way, increasing their technical efficiency by 190%. Just as the DEA analysis revealed, this progress was followed by a relative decline which recovered in the last year of the observed period (Table 10).

The Malmquist index and its decomposition suggest that the ExIn produces a positive effect not only on the participants, via an increase in the total productivity factor, a gain in pure efficiency and a considerable improvement in technological change. The non-participants, most of whom had applied for the ExIn, also improved their performance considerably. We claim that the ExIn had a positive spill-over effect: the participants introduced a new technology immediately after the program started, and the non-participants gradually started to adopt it. However, even though the trend in total productivity factor is almost parallel in the last years of our analysis, the growth rate of the frontier shift component for participants is greater than for the control group. Moreover, the treated universities, unlike the control group, experienced an increase (albeit limited) in pure efficiency.

6.5. Explanatory analysis of DEA efficiency scores

In the last stage of the analysis, the factors affecting university efficiency are investigated through a bootstrapped Tobit panel data regression. Alternative factors that can explain the change in efficiency scores between the two groups were included. Bootstrapped efficiency scores are regressed using factors which are not in full control of university management and the factors which reflect the way universities operate beyond the ExIn roadmaps. The results are presented in Table 11.

After a set of iterations for best model fit, we observe several significant determinants for university efficiency. The effect of Autonomous institution equals a 0.096 point increase ($p<0.01$). The financial structure of universities is reflected through the share of research and development income in the total financial resources. This parameter reflects the universities' engagement in research activity. The effect is moderate but positive and equals 0.0055 ($p<0.001$). The student body structure is another indicator we use to explain the efficiency dynamics. A high share of part-time students reflects profit-gaining in education as part-time education is an unselective option and is in demand (low opportunity costs and low tuition fees). The effect of the share of full-time students is significant ($p<0.01$) but moderate, 0.0012. Finally, the average entrance state exam score effect is examined. This indicator is divided into two separate determinants Russian HEIs have a dual-track system: the highest achievers compete for state-funded places and students with lower scores compete for paid places. If the student body structure is more homogeneous in terms of entry exam scores, the efficiency increases. The effect of the parameter is -0.0098 ($p<0.001$).

[Table 11] here

The second-step regression analysis reveals additional mechanisms that can determine positive managerial changes within the participants and can have an impact on the efficiency of the control group. Autonomous status provides greater freedom in the distribution of financial resources within university activities leading to an increase in efficiency. The more the university management can spread the resources – in adapting to new challenges and external factors – the more efficient the university is. Even with the design of the ExIn to fulfill two main aims: internationalization and increased research quality, the more the university is focused on research and development activity, the more relatively efficient it is in terms of the ExIn design. Research-oriented universities are more likely to enter the rankings, to gain a reputation and to remain efficient.

7. Concluding remarks

This study contributes to the discussion of the effects of the ExIn on university performance and efficiency. Our research is the first attempt to measure the efficiency of universities participating in the Russian ExIn. This paper develops a methodology for a complex causal assessment of ExIn effects. The issue of causality is present in the field of educational studies, especially if reform effects are evaluated (Schlotter et al., 2009). We reduce the causality ambiguity by forming a control group of universities, which are as similar to universities already participating in the ExIn as possible. One of the main external proofs of PSM quality is the fact that most of the control-group universities applied for the program and could compete with those who were finally accepted. After performing this step, we continue the evaluation via effectiveness (ATT, DID) and efficiency (DEA, the Malmquist index) analysis, later complemented by a second-stage double-bootstrap panel data explanatory analysis of efficiency determinants. Due to the high heterogeneity of the Russian HE system, the participants differ a lot from each other, i.e. in terms of the share of the ExIn subsidy within the total budget. This is one of the arguments to support measuring efficiency in terms of resource allocation and output maximization, not in terms of spending.

The combination of methods is has practical value for a quantitative evaluation of the reform effects. The main concern of our analysis is based on the discrepancy between the factual efficiency and the productivity of universities within the control group with their underrepresentation in world rankings. For instance, even being successful overall in terms of allocating their resources, they still cannot compete with the participants in terms of ranking inclusion. This might mean that not only efficiency matters, but what could be even more important while competing with the world best-performers, is reputation, status and internal managerial structure.

We agree that the ExIn design motivates universities to implement internal organizational changes which can lead to an increase in efficiency. In fact, our analysis revealed that the ExIn positively affects not only the participating universities, but the control group as well. The efficiency increase among the participants was followed by a positive trend in the control group's efficiency. The participants changed their production technology, which motivated the control group to adopt the new technology and even compete with the participants.

The factors which determine the efficiency increase apart from the ExIn itself, include a higher degree of management freedom in terms of financial resource allocation and a higher level of research and development orientation. This positively influences the efficiency within our modeling. The questionable inclusion of a rankings indicator into our efficiency modeling was checked for robustness: the difference in efficiency scores for modeling with and without raking is not statistically significant.

We examine the mechanisms for obtaining efficiency which can be used in decision making in the further development of the ExIn. On the one hand, participation in ExIn has a positive impact on total productivity and technical efficiency. Moreover, the participant universities perform better than non-participants. On the other hand, the difference between the groups is not sustainable in terms of this relative efficiency modeling. There are universities that have a high efficiency score and very positive changes even without additional project funding. This finding can be used for project design review: participants are currently evaluated on a yearly basis, but in fact, the kinds of changes that are awaited, demand a longer period of time (Poldin et al. 2017).

Finally, we observe quick growth in participant efficiency right after the ExIn started. This progress can be explained through the mobilization of recourses and was later followed by a decrease in efficiency growth and can also indicate a reputational effect of the ExIn, which can be examined in detail in future research. However, the participants manage to significantly outdo the control group by the end of the observed period. This progress in efficiency is complemented by progress in research effectiveness, measured by the number and the quality of publications. The latter is measured in two different ways: the average treatment effect on the treated, and the DID approach. The results confirm that that Project 5to100 had a statistically significant, positive effect on the efficiency and productivity of participating universities.

References

1. Abankina I. V., Aleskerov F. T., Belousova V. Yu., Zinkovsky K. V., Petrushenko V. V. (2013) Ocenka rezul`tativnosti universitetov s pomoshh`yu obolochechnogo analiza danny`x [Evaluation of Universities' Performance with Data Envelopment Analysis]. Voprosy obrazovaniya, (2), 15-48.

2. Abankina, I., Aleskerov, F., Belousova, V., Gokhberg, L., Kiselgof, S., Petrushchenko, V., ... & Zinkovsky, K. (2016). From equality to diversity: Classifying Russian universities in a performance oriented system. Technological Forecasting and Social Change, 103, 228-239.

3. Abbott, M., & Doucouliagos, C. (2003). The efficiency of Australian universities: a data envelopment analysis. Economics of Education review, 22(1), 89-97.

4. Ahn, T., & Seiford, L. M. (1993). Sensitivity of DEA to models and variable sets in a hypothesis test setting: The efficiency of university operations. Creative and innovative approaches to the science of management, 6, 191-208.

5. Ahn, T., Arnold, V., Charnes, A. B. R. A. H. A. M., & Cooper, W. W. (1989). DEA and ratio efficiency analyses for public institutions of higher learning in Texas. Research in Governmental and Nonprofit Accounting, 5(2), 165-185.

6. Altbach, P. G., & Salmi, J. (Eds.). (2011). The road to academic excellence: The making of world-class research universities. The World Bank.

7. Avkiran, N. K. (2001). Investigating technical and scale efficiencies of Australian universities through data envelopment analysis. Socio-Economic Planning Sciences, 35(1), 57-80.

8. Breu, T. M., & Raab, R. L. (1994). Efficiency and perceived quality of the nation's "top 25" National Universities and National Liberal Arts Colleges: An application of data envelopment analysis to higher education. Socio-Economic Planning Sciences, 28(1), 33-45

9. Breu, T. M., & Raab, R. L. (1994). Efficiency and perceived quality of the nation's "top 25" National Universities and National Liberal Arts Colleges: An application of data envelopment analysis to higher education. Socio-Economic Planning Sciences, 28(1), 33-45.

10. Carnoy, M., Loyalka, P., Dobryakova, M., Dossani, R., Kuhns, K., & Wang, R. (2013). University expansion in a changing global economy: Triumph of the BRICs?. Stanford University Press.

11. Carrington, R., Coelli, T. I. M., & Rao, D. P. (2005). The performance of Australian universities: Conceptual issues and preliminary results. Economic Papers: A journal of applied economics and policy, 24(2), 145-163.
12. Chirikov, I. (2018, in press) The Sputnik Syndrome: How Russian Universities Make 259 Sense of Global Competition in Higher Education In Oleksiyenko, V., Zha, Q., Chirikov, I., & Li, J. (Eds) *International Status Anxiety and Higher Education: The Soviet Legacy in China and Russia*. ISBN 978-988-14241-6-7

13. Eckles, J. E. (2010). Evaluating the efficiency of top liberal arts colleges. Research in Higher Education, 51(3), 266-293.

14. Flegg, A. T., & Allen, D. O. (2007). Does expansion cause congestion? The case of the older British universities, 1994–2004. Education Economics, 15(1), 75-102.

15. Froumin, I., & Lisyutkin, M. (2015). Excellence-driven policies and initiatives in the context of Bologna process: Rationale, design, implementation and outcomes. In The European Higher Education Area (pp. 249-265). Springer, Cham.

16. Gawellek, B., & Sunder, M. (2016). The German excellence initiative and efficiency change among universities, 2001-2011(No. 142). Working Paper, Universität Leipzig, Wirtschaftswissenschaftliche Fakultät.

17. Gromov, A. (2017). The Efficiency of Russian Higher Education Institutions and its Determinants.

18. Hazelkorn, E. (2014). Rankings and the global reputation race. New directions for higher education, 2014(168), 13-26.

19. Hazelkorn, E. (2015). Rankings and the reshaping of higher education: The battle for world-class excellence. Springer.

20. Imbens, G. W., & Wooldridge, J. M. (2009). Recent developments in the econometrics of program evaluation. *Journal of economic literature*, *47*(1), 5-86.

21. Johnes, J., & Li, Y. U. (2008). Measuring the research performance of Chinese higher education institutions using data envelopment analysis. China economic review, 19(4), 679-696.Kehm, B. M. (2006). The German'Initiative for Excellence'and rankings. International Higher Education, 44, 20-22.

22. Kopytok V., Ratnikova T (2017). Vliyaniye politiki inflyatsyonnogo targetirovaniya na dinamiku pryamykh inostrannykh investitsii [The Impact of the Inflation Targeting Policy on the Foreign Direct Investments Dynamics]. Higher School of Economics Economic Journal, 21 (1), 32-65.

23. Lee K. R. et al. Malmquist Productivity Index using DEA frontier in Stata. URL: https://pdfs.semanticscholar.org/847f/0d96a2b8e5581ba45e32406ec16a8858839c.pdf

24. Leshukov, O., Platonova, D., & Semyonov D. (2016). The Efficiency of Regional Higher Education Systems and Competition in Russia. Ekonomika regiona, 12(2), 417-426.

25. Lixu, L. (2004). China's higher education reform 1998–2003: A summary. Asia Pacific Education Review, 5(1), 14.

26. López-Torres, L., Prior, D., & Santín, D. (2016). Quality improvement programs in public schools. An evaluation of their effectiveness and efficiency.

27. Marginson, S. (2006). Dynamics of National and Global Competition in Higher Education. Higher Education, 52(1), 1–39.

28. Möller, T., Schmidt, M., & Hornbostel, S. (2016). Assessing the effects of the German Excellence Initiative with bibliometric methods. scientometrics, 109(3), 2217-2239

29. Ng, Y. C., & Li, S. K. (2009). Efficiency and productivity growth in Chinese universities during the post-reform period. China economic review, 20(2), 183-192.

30. Poldin, O., Matveeva, N., Sterligov, I., & Yudkevich, M. (2017). Publication Activities of Russian Universities: The Effects of Project 5–100. Educational Studies Moscow No 2, 2017, 59.

31. Salmi, J. (2009). The challenge of establishing world class universities. The World Bank.

32. Salmi, J. (2016). Excellence strategies and the creation of world-class universities. In Matching Visibility and Performance (pp. 15-48). SensePublishers, Rotterdam.

33. Salmi, J., & Expert, G. T. E. (2012). Attracting talent in a global academic world: How emerging research universities can benefit from brain circulation. Academic Executive Brief, 2(1).

34. Schlotter, M., Schwerdt, G., & Woessmann, L. (2011). Econometric methods for causal evaluation of education policies and practices: a non- technical guide. *Education Economics*, *19*(2), 109-137.

35. Seong, S., Popper, S. W., Goldman, C. A., Evans, D. K. & Grammich, C. A. (2008). Brain Korea 21 phase II: a new evaluation model. Santa Monica: Rand Corporation, 64-76.

36. Shin, J. C. (2009). Building world-class research university: The Brain Korea 21 project. Higher Education, 58(5), 669.

37. Sokolov M., Tsivinskaya A. 2018a. Monitoring the effectiveness of educational organizations as a tool for evaluating Russian universities (NNS-S (2) series of preprints). St. Petersburg: CIANO EUSPb

38. Sokolov, M., Tsivinskaya, A. 2018b. Monitoring the effectiveness of educational organizations as a source of data on Russian higher education (HHC-S (1) series of preprints). St. Petersburg: CIANO EUSPb

39. Turko, T., Bakhturin, G., Bagan, V., Poloskov, S., & Gudym, D. (2016). Influence of the program "5–top 100" on the publication activity of Russian universities. Scientometrics, 109(2), 769-782.

40. Worthington, A. C. (2001). An empirical survey of frontier efficiency measurement techniques in education. Education economics, 9(3), 245-268.

41. Yang, X., & You, Y. (2017). How the World-Class University Project Affects Scientific Productivity? Evidence from a Survey of Faculty Members in China. Higher Education Policy, 1-23

42. Yaisawarng, S., & Ng, Y. C. (2014). The impact of higher education reform on research performance of Chinese universities. China Economic Review, 31, 94-105.

43. Yudkevich, M., Altbach, P. G., & Rumbley, L. E. (Eds.). (2016). The global academic rankings game: Changing institutional policy, practice, and academic life. Routledge.

44. Yufang, O. Evaluating effects of world-class university program using propensity score matching: Evidence from China's project 211 (2017) ISSI 2017 - 16th International Conference on Scientometrics and Informetrics, Conference Proceedings, pp. 400-411.

45. Zhang, H., Patton, D., & Kenney, M. (2013). Building global-class universities: Assessing the impact of the 985 Project. Research Policy, 42(3), 765-775.

46. Zinkovsky K. V., Derkachev P. V. (2016) Restrukturizaciya sistemy` vy`sshego obrazovaniya: ocenka rezul`tatov ob``edinenij vuzov [Higher education Restructuration: Evaluation of Universities' Mergers]. Universities' management: practice and analysis, (6), 135-145

47. Zinkovsky, K., Derkachev, P., & Vinarik, V. (2016). Exploring the restructuring of higher education system: an evidence based evaluation of higher education institutions passed the process of joining to other ones. XVII April International Academic Conference on Economic and Social Development

48. Zong, X., & Zhang, W. (2017). Establishing world-class universities in China: deploying a quasi-experimental design to evaluate the net effects of Project 985. Studies in Higher Education, 1-15.

Authors:

Tommaso Agasisti, Politecnico di Milano School of Management, Associate professor, tommaso.agasisti@polimi.it

Ekaterina Shibanova, National Research University Higher School of Economics, Laboratory for University Development, Intern researcher; E-mail: eshibanova@hse.ru

Daria Platonova, National Research University Higher School of Economics, Laboratory for University Development, Laboratory head; E-mail: dplatonova@hse.ru

Mikhail Lisyutkin National Research University Higher School of Economics, Institute of Education, Research fellow; E-mail: mlisyutkin@hse.ru

Figures and tables

Table 1. 5to100 universities' positions in rankings

University	QS Subject Ranking	The highest position in the ranking 2016	2017
National University of Science and Technology "MISiS"	Mineral and Mining Engineering	-	31
Moscow Institute of Physics and Technology	Physics and Astronomy	101	42
Novosibirsk State University	Physics and Astronomy	51	50
National Research University Higher School of Economics	Politics and International Studies, Sociology, Economics and Econometrics	101	51
National Research Nuclear University MEPhI	Physics and Astronomy	51	51

Source: QS ranking

Figure 1. Structural responses to the definitions of global competition (Chirikov 2018).

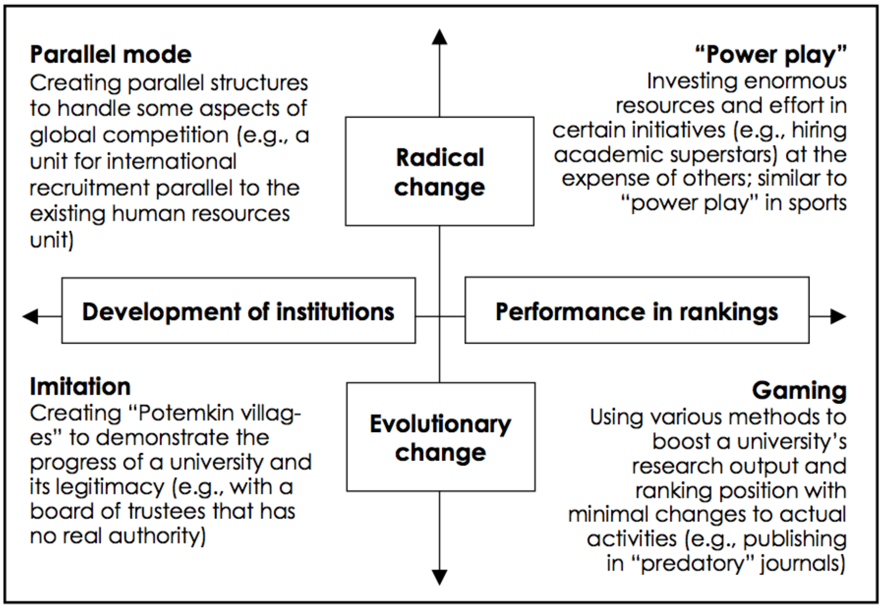


Table 2. Descriptive statistics for PSM-used variables in the end of the observed period

Group	Statistics	Entry exam average score	Share of PhD students	Number of citations indexed in RSCI, per 100 academic staff	R&D expenditures per a faculty member	Share of foreign students	Share of fixed assets less than 5 years old	Total number of students	Total number of teaching and research staff	Share of STEM students	Regional index of industrial production	Regional GDP per capita
Non-participants (15)	Average	68.65	6.21	771.85	451.84	2.6	31.22	13371.87	1088.2	54.24	103.85	765518.41
	St.Dev	5.69	6.62	497.29	448.83	1.99	22.19	5486.6	510.21	32.21	2.71	377204.43
	Min	60.4	1.74	116.09	67.51	0.28	0.12	4197	466	1.41	101.3	300186.2
	Max	82.83	29.85	287.83	2081.44	5.44	89.83	27035	2029	95.77	106.2	1500000
Participants (15)	Average	76.75	6.08	1365.99	541.1	3.51	43.51	16234.27	1479.53	62.31	102.78	677459.98
	St.Dev	8.12	2.83	1186.07	283.39	2.11	18.49	8617.51	736.08	24.37	2.26	311189.3
	Min	62.53	2.91	160.5	188.33	1.35	4.02	6095	541	16.36	97.8	363327.8
	Max	94.45	13.9	5229.35	970.84	8.39	65.95	32835	3011	100	107.7	1157373

Source: authors' calculations

Table 3. PSM model specification.

Indicator	Selection rationale
Research and development expenditures per one faculty member (research and teaching staff)	development of the research
Average entrance exam score[6] (unified state examination average score, full-time, state-funded students)	selectivity
Share of foreign students (excluding the students from Commonwealth of Independent stands for the aim of internationalization of Russian universities States counties; bachelor, specialist and master programs) in the student body, (per 100)	internationalization
Total number of students (bachelor, specialist and master programs; full-time students; state-funded and fee-paid)	reflects the size of the university as we want matching pairs to be almost equal
Share of PhD students	development of the research
Share of students of the education programs related to «STEM»	reflects the specialization of a certain university
Share of fixed assets less than 5 years old	reflects infrastructure of the university
Number of employees (research and teaching staff), incl. share in total	excludes a bias caused by differences in size
Number of citations indexed in Russian Science Citation Index (RSCI), per 100 academic staff reflects the academic and scientific performance within Russia	research performance
Regional domestic product index per capita; Regional index of industrial production	economic context of location

[6] So-called Universal State Exam is a high-stakes exam that combines both school-leaving exam and university entrance exam. It is an obligatory exam for all school graduates and universities enrollees (with some exceptions). The MoES often considers the average exam score of universities' enrollees as a proxy of university success in education filed and ability to attract talents.

Table 4. DEA model specification

Variable	**Selection rationale**
Inputs:	
Total income per student (without deduction of staff salary expenditures)	Proxy for basic capital description
Total number of academic staff (without external employment)	Proxy for basic labour description
Share of fixed assets less than 5 years old, %	Proxy for basic capital description
Outputs:	
<u>*Teaching*</u> Unified State Exam (entrance exam) score	Proxy for quality of application
Total number of Russian bachelor, specialist, masters students	Proxy for size
Share of PhD students, %	Development of the research potential indicator
Share of foreign students (excluding the students from Commonwealth of Independent; bachelor, specialist and master programs) in the student body, (per 100)	Proxy for internationalization
<u>*Research*</u> Number of publications indexed in Scopus or Web of Science, per capita of academic staff	Proxy for research performance on international quality level
Number of citations indexed in Scopus or Web of Science, per capita of academic staff	Proxy for research performance on international quality level
<u>*Ranking*</u> Inclusion into one of the worldwide university rankings (the Academic Ranking of World Universities, the Times World University Ranking or the QS ranking, 0-3)	Proxy for the main goal of the ExIn – world rankings inclusion

Table 5. DEA variables descriptive statistics after PSM, the first year of analysis

Group		Entry exam average score	Share of PhD students	Number of citations indexed in RSCI, per 100 academic staff	R&D expenditures per a faculty member	Share of foreign students	Share of fixed assets less than 5 years old	Total number of students	Total number of teaching and research staff	Share of STEM students	Regional index of industrial production	Regional GDP per capita
Non-participants	Average	74.49	6.55	104.08	844.92	1.79	58.24	14097.60	1051.20	49.93	99.20	960978.53
	St.Dev	7.29	3.57	70.78	600.27	1.65	23.36	5738.72	404.15	31.15	3.99	511602.42
	Minimim	64.60	3.47	15.77	72.39	0.00	0.00	5202.00	415.00	0.00	95.30	323669.00
	Maximum	86.14	14.16	287.83	2081.44	5.44	89.83	27035.00	2029.00	95.77	106.20	1500000.00
Participants	Average	72.73	5.66	150.07	1168.75	1.51	50.71	16653.07	1318.20	57.27	100.03	828499.53
	St.Dev	7.42	1.72	108.89	664.11	1.28	25.85	9720.37	704.36	32.32	2.98	415043.47
	Minimim	60.44	3.36	6.96	325.46	0.00	0.00	4977.00	443.00	4.86	95.30	422090.00
	Maximum	91.01	10.78	371.45	2665.64	4.94	83.40	36821.00	2886.00	96.81	105.20	1500000.00

Source: authors' calculations

Table 6. DEA variables descriptive statistics after PSM, 2012/2013-2016/2017 academic years

Year	Group	Statistics	Total income per capita of academic staff, 1000 rubles, 1000 rubles	Fixed assets less than 5 yearы old, %	R&D expenditures per capita of academic staff	Share of foreign staff, %	Share of foreign students, %	Number of publications (Scopus, WoS), per 100 of academic staff	Number of citaitons (Scopus, WoS), per 100 of academic staff	Number of rankings
2012/13	Participants	Average	4088.791	58.238	779.878	1.783	1.794	6.641	32.699	0
		St. Dev.	2354.154	23.355	554.06	3.259	1.649	5.876	34.538	0
	Non-paticipants	Average	4413.172	50.713	1078.779	1.411	1.513	30.143	197.429	0.467
		St. Dev.	1705.02	25.849	612.982	1.682	1.277	19.83	259.253	0.499
2013/14	Participants	Average	3227.718	38.043	309.779	1.064	2.921	12.856	129.159	0.067
		St. Dev.	1587.513	20.238	357.047	2.585	3.718	10.681	240.985	0.249
	Non-paticipants	Average	5031.197	57.936	356.391	0.892	1.88	66.473	708.521	0.667
		St. Dev.	1976.472	17.425	259.21	0.915	1.265	56.207	1017.656	0.471
2014/15	Participants	Average	3591.666	46.369	307.037	1.429	6.701	22.528	105.097	0.067
		St. Dev.	1577.928	21.898	257.673	3.499	4.855	12.583	103.773	0.249
	Non-paticipants	Average	4670.09	54.256	415.05	1.686	7.677	109.271	806.646	0.667
		St. Dev.	1550.076	21.146	289.622	1.606	4.791	75.818	1311.291	0.596
2015/16	Participants	Average	3366.09	43.212	434.307	0.879	7.569	26.581	134.459	0.467
		St. Dev.	1549.723	17.282	391.65	0.958	5.533	13.978	120.678	0.618
	Non-paticipants	Average	4585.449	52.384	534.666	2.451	9.413	156.761	1122.359	1.2
		St. Dev.	1829.029	21.849	320.915	1.653	5.464	101.449	1492.628	1.046
2016/17	Participants	Average	1665.925	31.216	451.845	1.249	7.433	44.301	219.689	0.4
		St. Dev.	876.826	22.193	448.833	1.657	4.988	30.662	174.773	0.8
	Non-paticipants	Average	1449.203	43.511	541.097	3.449	11.107	229.765	1941.811	1.667
		St. Dev.	553.54	18.489	283.391	3.214	6.021	146.935	2426.828	0.943

Source: authors' calculations

Table 7. Results of the impact-evaluation analysis: ATT and DiD approaches

Method	Research indicator					
ATT		Effect measure	St. error	T-statistics		
	Publications	130.180***	27.369	4.76		
	Citations	987.900**	400.223	2.47		
DiD		DiD intercept regression coefficient	Robust st. error	T-statistics	95% Confidence intervals	
	Publications	90.499***	16.584	5.460	57.723	123.275
	Citations	833.000***	236.166	3.530	366.255	1299.747

**** p<0.01, ** p<0.05, * p<0.1*

Source: authors' calculations

Note: The number of publications and the number of citations per 100 capita of academic staff are used as dependent outcome variables in both ATT and DiD.

Table 8. Descriptive statistics, DEA Efficiency Indexes by groups: 2012-2016.

Group	Academic year	2012/2013 the «zero» year	2013/2014 the 1st year	2014/2015	2015/2016	2016/2017
Non-participants	Average	0.695	0.616*	0.779	0.716	0.751*
	Maximum	1	1	1	1	1
	Minimum	0.115	0.099	0.252	0.198	0.119
	St. Dev.	0.241	0.324	0.271	0.316	0.286
Participants	Average	0.829	0.847*	0.852	0.877	0.905*
	Maximum	1	1	1	1	1
	Minimum	0.431	0.303	0.227	0.314	0.314
	St. Dev.	0.195	0.193	0.2	0.185	0.188

**** $p<0.01$, ** $p<0.05$, * $p<0.1$*

Source: authors' calculations

Table 9. Pearson correlations between original and bootstrapped DEA efficiency scores

Academic year		2012/2013 the «zero» year	2013/2014 the 1st year	2014/2015	2015/2016	2016/2017
Pearson's correlation		0.9852***	0.9901***	0.9002***	0.9883***	0.9902***
Kolmogorov-Smirnov test p-value	DEA distribution	0.373	0.373	0.543	0.891	0.543
	Bootstrapped DEA	0.373	0.152	0.543	0.678	0.184
T-test p-value (Pr T<t)	DEA	0.157	0.058	0.438	0.22	0.10
	Bootstrapped DEA	0.181	0.053	0.41	0.429	0.074
Wilcoxon test p-value	DEA	0.20	0.046	0.648	0.1074	0.0913
	Bootstrapped DEA	0.307	0.02	0.777	0.496	0.049

***** p<0.01, ** p<0.05, * p<0.1*

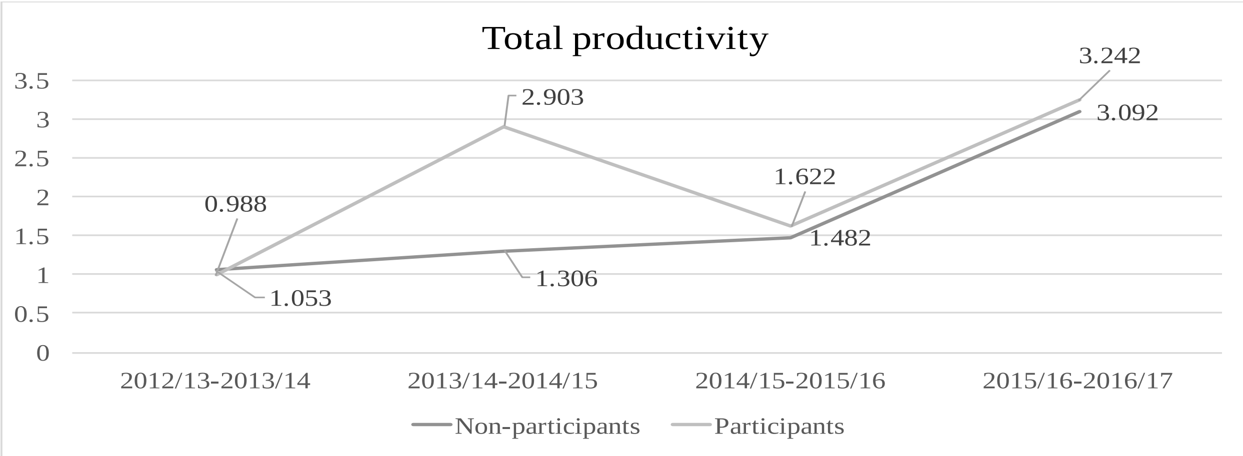
Source: authors' calculations

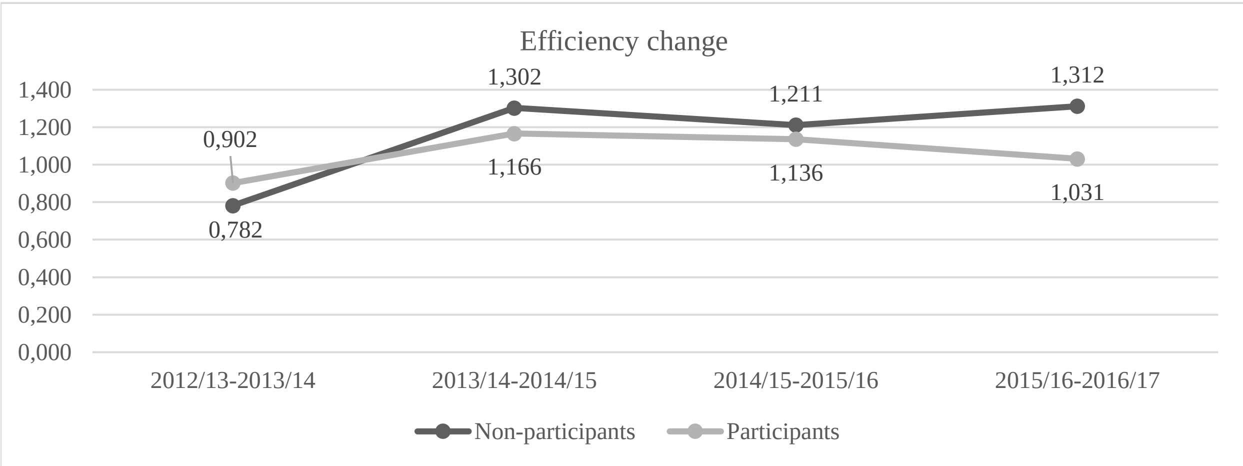
Table 10. Averages Malmquist Indexes by groups and its components: 2012-2016.

Academic years	Group	Total productivity	Efficiency change	Frontier shift
2012/2013-2013/2014	Non-participants	1.053	0.782	1.707
	Participants	0.988	0.902	1.130
2013/2014-2014/2015	Non-participants	1.306	1.302	1.030
	Participants	2.903	1.166	2.164
2014/2015-2015/2016	Non-participants	1.482	1.211	1.245
	Participants	1.622	1.136	1.370
2015/2016-2016/2017	Non-participants	3.092	1.312	2.470
	Participants	3.242	1.031	3.071**
Total change for 2013 - 2017	Non-participants	4.574	0.984	4.234
	Participants	5.015	1.114	4.692

**** p<0.01, ** p<0.05, * p<0.1*

Source: authors' calculation





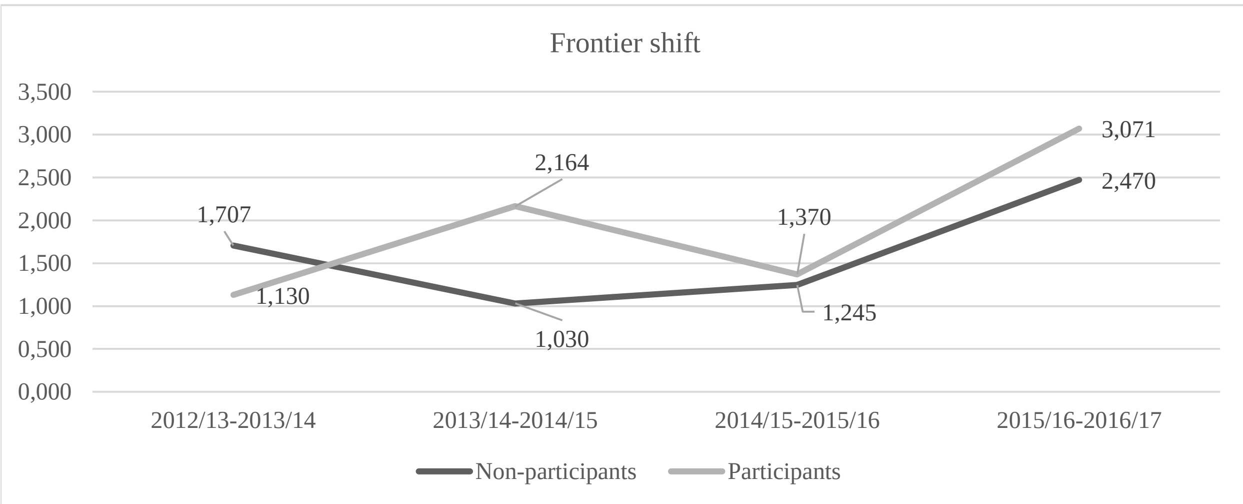


Figure 2. Malmquist index and its decompositions scores by groups in dynamics.

Source: authors' calculations

Table 11. Bootstrapped panel Tobit-regression analysis results

Dependent variable: Bootstrapped efficiency scores	Coefficient	Bootstrapped st. errors	P>|z|	95% Confidence intervals	
"Autonomous institution"	0.0956 ***	0.031	0.002	0.035	0.156
Difference between entry exam scores of state-funded enrollees and entry exam scores of tuition fee paying enrollees	-0.0098 *	0.005	0.059	-0.020	0.000
Share of R&D income	0.0055 ***	0.002	0.000	0.003	0.009
Share of full-time students	0.0012 ***	0.000	0.001	0.001	0.002
Constant term	0.57	0.080	0.000	0.410	0.725
Sigma u	0.10	0.017	0.000	0.064	0.131
Sigma e	0.14	0.017	0.000	0.102	0.171
rho	0.34	0.124		0.140	0.597

**** p<0.01, ** p<0.05, * p<0.1*

Source: authors' calculations

Annex 1. Additional details on DEA calculations

Table 1 Robustness check of DEA model: step-by-step exclusion of each variable and correlations with the original model

	Efficiency 2012-2013	Efficiency 2013-2014	Efficiency 2014-2015	Efficiency 2015-2016	Efficiency 20167-201
Total income per capita	0.6390*	0.8099*	0.6213*	0.8592*	0.8456*
sig	0.0001	0	0.0002	0	0
Assets	0.8031*	0.7357*	0.9790*	0.9832*	0.7529*
sig	0	0	0	0	0
R&D_per_capita	0.5052*	0.8099*	0.8559*	0.5945*	0.7590*
sig	0.0044	0	0	0.0005	0
Share of foreign staff	0.9388*	0.8103*	0.9603*	0.7935*	0.6002*
sig	0	0	0	0	0.0005
Share of foreign students	0.5281*	0.8164*	0.9180*	0.7357*	0.5994*
sig	0.0027	0	0	0	0.0005
Publications	0.6893*	0.7619*	0.9533*	0.7949*	0.9135*
sig	0	0	0	0	0
Citations	0.6860*	0.8104*	0.9999*	0.8082*	0.9135*
sig	0	0	0	0	0
Rankings	0.5805*	0.5946*	0.9441*	0.7827*	0.8311*
sig	0.0008	0.0005	0	0	0

Source: authors' calculations

Note: for each academic year each input or output variable used in DEA model is excluded, then the efficiency scores are recalculated. Later, the new efficiency scores are correlated with the initial ones. The "sig" stands for the p-value for correlation analysis.