



# Measuring the efficiency of public schools in Uruguay: main drivers and policy implications

Daniel Santín<sup>1</sup> · Gabriela Sicilia<sup>1</sup>

Received: 21 July 2014/Revised: 18 December 2014/Accepted: 16 April 2015/

Published online: 5 May 2015

© The Author(s) 2015. This article is published with open access at Springerlink.com

**Abstract** The aim of this research is to explore the existence of inefficient behaviors in public high schools in Uruguay and identify its potential drivers. To do so, we perform a two-stage model using PISA 2009 and 2012 databases. In the first stage, we use Data Envelopment Analysis (DEA) to estimate efficiency scores, which are then regressed on school and student contextual variables. This second stage is carried out using four alternative models: a conventional censored regression and three different regression models based on the use of bootstrapping recently proposed in the literature. Our results show that educational efficiency in Uruguayan high schools significantly dropped in nine percentage points between 2009 and 2012. In terms of educational policy recommendations, in order to reduce the inefficiencies in the evaluated public schools in Uruguay, the focus should be put on reducing grade-retention levels and promoting teaching–learning techniques that enhance student’s mathematics study skills and assessing students continuously through test and homework throughout the academic year. In this vein, our findings also show positive effects on public schools’ efficiency of providing the responsibility in the distribution of the school budget to school principals.

**Keywords** Public secondary education · Technical efficiency · Performance-based reforms · PISA · DEA

**JEL Classification** I21 · C14 · H52

---

✉ Gabriela Sicilia  
gabriels@ucm.es

<sup>1</sup> Department of Applied Economics VI, Complutense University of Madrid, Campus de Somosaguas, 28223 Pozuelo de Alarcón, Madrid, Spain

## 1 Introduction

There are basically two reasons why governments in developed countries have taken a strong interest in the determinants of educational quality over the last 50 years. First, improving academic outcomes have been proven to have a positive impact on economic growth (Barro 2001; Barro and Lee 2012; Hanushek and Kimko 2000; Hanushek and Woessmann 2012). Second, public expenditure on education is one of the largest public budget items, and the public sector is the main provider of education in most countries. Governments are not concerned solely with improving academic results, however, they mean to do so with the current educational resources, that is, through efficiency gains. The main reason is that public expenditure on education has grown over recent years in many countries, without leading to better academic results.

Particularly, the Uruguayan government has increased the country's investment in education considerably over the last decade. Public expenditure on education accounted for 3.5 % of Uruguay's GDP in 2000, whereas 10 years later it had risen to 4.5 %.<sup>1</sup> But this significant budgetary effort has not been accompanied by adequate reforms and public policies leading to better educational achievement in public schools. Conversely, the Uruguayan education system has entered into stagnation and recession in recent years, particularly at the public secondary education level, which has recorded high repetition and dropout rates as well as a steady decline in academic performance. For example, the repetition rate from 1st to 4th grades in public schools has increased between 2003 and 2012 from 21.3 to 27 % while the attainment rate was reduced from 72.7 to 67.4 % in the same period.<sup>2</sup> In addition, as evidenced by the latest results published in the PISA 2012 (Programme for International Student Assessment) Report from the OECD (Organisation for Economic Co-operation and Development), results in public schools remain steady across the first three waves in which Uruguay has participated, showing a downward trend in the last cycle (416, 420, 419 and 399 average points in 2003, 2006, 2009 and 2012, respectively).

As a consequence of these poor results, the Uruguayan public educational system problems are a recurring concern, not only for educational policymakers and the government, but also for teachers and families involved in the education process. In many cases, the discussion primarily still focuses on increasing public resources expended on education; however, there is no concluding empirical evidence in the economics of education literature to show that a higher level of resources leads per se to better results (Hanushek 2003).

These findings reveal that the solution to Uruguay's educational problem is not simply to pour additional resources into the system; instead it is necessary to review and change some existing practices and educational policies that are not effective. In this sense, the main concern of educational policy makers in Uruguay should be to improve the quality of teaching and academic outputs with the currently available

---

<sup>1</sup> The GDP grew by 37 % in real terms over this period [Uruguayan Central Bank (BCU)].

<sup>2</sup> Education Observatory, National Administration of Public Education (ANEP).

resources. To do this, it is clearly necessary to explore and address the main sources of educational inefficiencies.

Using the databases of different international programs,<sup>3</sup> many researchers have performed specific analyses of the main sources of inefficient behavior in the educational production process using student and school contextual variables (Wilson 2005; Afonso and St. Aubyn 2006; De Jorge and Santín, 2010; Cordero et al. 2011; Perelman and Santin 2011; Crespo-Cebada et al. 2014).<sup>4</sup>

Semiparametric two-stage models were popularized by Ray (1991) and McCarty and Yaisawarng (1993) and are among the best-known models for explaining the sources of inefficiency.<sup>5</sup> The first stage of this approach prescribes the use of a Data Envelopment Analysis (DEA) model to estimate a production frontier, which defines both the efficient and inefficient units. In the second stage, a regression technique is applied to explain the identified inefficient behaviors taking into account contextual variables. Two-stage models differ primarily in the regression model specified in the second stage to explain efficiency scores. The most commonly applied methodology is the censored regression model (the so-called Tobit regression), followed by ordinary least squares (OLS) and truncated regression. Recently, Simar and Wilson (2007, 2011) proposed a new estimation methodology for the second stage based on the use of bootstrapping to overcome some drawbacks of these conventional estimation models. We apply the Simar and Wilson (2007) two-stage approach as our baseline model in this research, but, as the discussion about which is the best model to be run in the second-stage regression is ongoing, we also run other second-stage specifications proposed in the literature in order to check the robustness of our conclusions.

Finally, it is noteworthy that even though there are several international educational efficiency studies for the OECD countries, research in the Latin American context is scant. To the best of our knowledge, there are no studies using this efficiency approach for the Uruguayan case. In Uruguay, interest has traditionally focused on education system coverage rates, the system's redistributive effect and its impact on poverty and growth rather than the quality of the services provided and the academic outputs (Llambí and Perera 2008; Llambí et al. 2009; Fernández 2009).

Therefore, the main aim of this paper is to explore the sources of inefficiencies in Uruguayan secondary schools in order to provide new valuable and complementary evidence for the current national debate about which educational practices and policies could contribute to improving school academic results with the current resources. For this purpose, we apply a semiparametric two-stage DEA approach to PISA 2009 and 2012 data in order to compare the results between the two periods. The paper is organized as follows. Section 2 presents the main methodological concepts. Section 3 briefly describes the Uruguayan education system, the PISA

<sup>3</sup> These programs include PISA, TIMSS (Trends in International Mathematics and Science Study), IALS (International Assessment of Literacy Survey) and PIRLS (Progress in International Reading Literacy Study).

<sup>4</sup> See Worthington (2001) and Mancebón and Muñiz (2003) for a detailed review of educational efficiency studies with other country-specific databases.

<sup>5</sup> See Simar and Wilson (2007) for a detailed review of two-stage models.

program and the variables included in the model. Section 4 reports the estimation results. Finally, Sect. 5 discusses the conclusions of this research and their implications for educational policy makers.

## 2 Methodology

### 2.1 The educational production function

The educational production function framework refers to the relationship between inputs and outputs for a given production technology. The theoretical approach used in this paper for linking resources to educational outcomes at school level is based on the well-known educational production function proposed by Levin (1974), Hanushek (1979) and Hanushek et al. (2013):

$$A_i = f(B_i, S_i), \quad (1)$$

where subindex  $i$  refers to school, and  $A_i$  represents the educational output vector for school  $i$ . This output is normally measured through the students' average scores in standardized tests. On the other hand, educational inputs are divided into  $B_i$ , which denotes average student family and socio-economic background, and  $S_i$ , which are school educational resources.

The educational production function is frequently estimated considering the possible existence of inefficient behaviors in schools. Differences in efficiency may be due to multiple factors, such as poor teacher motivation, teaching and class organization issues, teacher quality or school management. Although all these factors are not direct inputs, they may affect student performance significantly. In this case, we estimate a production frontier where fully efficient schools would belong to the educational production frontier. These relatively efficient units achieve the maximum observed result given their resource allocation. However, inefficient units do not belong to the estimated frontier, and their inefficiency level is measured by the radial distance between each school and the constructed frontier. The production frontier to be estimated at school level would be

$$A_i = f(B_i, S_i) \cdot u_i, \quad (2)$$

where  $0 < u_i \leq 1$  denotes the efficiency level of school  $i$ . Values of  $u_i = 1$  imply that the analyzed schools are fully efficient, meaning that given the initial input endowment and the existing technology, these schools are maximizing their outputs and managing correctly the available school inputs. Values  $u_i \leq 1$  would indicate that the school is inefficient, and therefore the efficiency rate,  $\theta_i = 1/u_i$  indicates the amount by which the actual output should be multiplied to reach the frontier in which case the school would be fully efficient.

In short, three types of variables are involved in the production process: educational outputs ( $A_i$ ), educational inputs ( $B_i$ ,  $S_i$ ) and the estimated efficiency level ( $u_i$ ) for each school. Ray (1991) and McCarty and Yaisawarng (1993) were the first to propose applying a semiparametric two-stage model to estimate efficiency

scores and identify the main drivers. The first stage of this approach is to apply a DEA model which measures technical efficiency, whereas a regression analysis conducted in the second stage seeks out the main explanatory factors of efficiency. A more detailed description of the two-stage methodology follows.

## 2.2 First stage: measuring efficiency through a DEA-BCC model

The measurement of efficiency is associated with Farrell's concept of technical efficiency (Farrell 1957). Farrell defines the production frontier as the maximum level of output that a decision-making unit (DMU) can achieve given its inputs and the technology (output orientation). In practice, the true production frontier and the technology is not known and should be estimated from the relative best practices observed in the sample.

There are basically two main groups of techniques for estimating the production frontier: parametric, or econometric approaches (see Battese and Coelli 1988, 1992, 1995 for a review), and nonparametric methods based on mathematical optimization models. Although the use of parametric approaches has increased in education in the last decades,<sup>6</sup> nonparametric methods have been the most extensively applied methods for measuring educational efficiency.

Since the pioneering work by Charnes et al. (1978, 1981) and Banker et al. (1984),<sup>7</sup> the DEA model has been widely used to measure efficiency in many areas of public expenditure. The main reason for its widespread application is its flexibility, and the fact that it accounts for multiple outputs and inputs, unknown production technology and missing price information, which makes it to well suited to the peculiarities of the public sector. The technique applies a linear optimization program to obtain a production frontier that includes all the efficient units and their possible linear combinations. As a result, the estimated efficiency score for each DMU is a relative measure calculated using all the production units that are compared. The formulation of the output-oriented DEA program under variable returns to scale (DEA-BCC model) for each analyzed unit is

$$\theta_i^{\text{DEA}} = \max_{\lambda, \theta} \{ \theta_i \mid \theta y_i \leq Y\lambda; x_i \geq X\lambda; n' \lambda = 1; \lambda \geq 0; \forall i = 1, \dots, n \}, \quad (3)$$

where for the  $i$ th DMU,  $\theta_i \geq 1$  is the efficiency score,  $y_i$  is the output vector ( $q \times 1$ ) and  $x_i$  is the input vector ( $p \times 1$ ), and thus  $X$  and  $Y$  are the respective input ( $p \times n$ ) and output ( $q \times n$ ) matrices. The  $(n \times 1)$  vector  $\lambda$  contains the virtual weights of each unit determined by the problem solution. When  $\theta_i = 1$  the analyzed unit belongs to the frontier (is fully efficient), whereas  $\theta_i > 1$  indicates that the  $i$ th unit is inefficient,  $\theta_i$  being the radial distance between the  $i$ th unit and the frontier. In other words,  $\theta_i$  indicates the equiproportional expansion over outputs needed to reach the frontier. Therefore, the higher the score value  $\theta_i$  is, the greater the inefficiency level is.

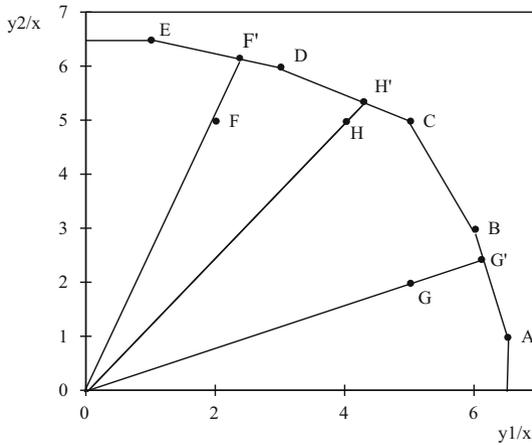
<sup>6</sup> See, for example, Perelman and Santin (2011) and Crespo-Cebada et al. (2014).

<sup>7</sup> The DEA-CCR model and DEA-BCC model, respectively, from now on.

**Table 1** Example data to illustrate an output-oriented DEA

	A	B	C	D	E	F	G	H
Output 1 ( $y_1$ )	6.5	6	5	3	1	2	5	4
Output 2 ( $y_2$ )	1	3	5	6	6.5	5	2	5
Input 1 ( $x$ )	1	1	1	1	1	1	1	1
Efficiency ( $\theta_i$ )	1	1	1	1	1	1.2273	1.2273	1.0715

Source: authors' own elaboration



**Fig. 1** An output-oriented BCC-DEA. Source: authors' own elaboration

In order to briefly illustrate a DEA model let assume the following simple two-output single-input setting where it is assumed that eight DMUs,  $A, B, C, D, E, F, G$  and  $H$  have an equal single-input  $x$  to produce outputs  $y_1$  and  $y_2$ . The data and the efficiency index measured by DEA are showed in Table 1.

DEA runs eight linear programming problems (Eq. 3), one for each of the eight DMUs contained in our example in order to construct a piecewise linear frontier with best performers which envelops the other inefficient DMUs. Figure 1 shows this production frontier where  $A, B, C, D$  and  $E$  are efficient DMUs  $\theta_A = \theta_B = \theta_C = \theta_D = \theta_E = 1$  because they lie on the boundary of the production frontier; whereas, being interior points,  $F, G$  and  $H$  are inefficient units  $\theta_F > 1; \theta_G > 1; \theta_H > 1$ .

DEA measures inefficiency as the radial distance from the inefficient unit to the frontier. For example, the performance of DMU  $F$  is measured projecting this unit upwards to point  $F'$ , a linear combination of DMUs  $D$  and  $E$ , with output 1 and output 2 equal to 2.4546 and 6.1365, respectively. DEA calculates the efficiency of DMU  $F$  as  $\theta_F = OF'/OF = 1.2273$ . This result means that DMU  $F$  could increase all its outputs proportionally multiplying its actual outputs level by 1.2273 with its input vector fixed.

### 2.3 Second stage: explaining educational efficiency scores

The estimated efficiency scores  $\hat{\theta}_i$  are regressed on a vector  $Z = (z_1, z_2, \dots, z_k)$  of school and student contextual variables, which are not inputs but are related to the learning process:

$$\hat{\theta}_i = f(Z_i, \beta_i) \quad (4)$$

The most used estimation method in this second stage is the censored regression model (Tobit), followed by ordinary least squares (OLS),<sup>8</sup> from which the main explanatory factors of the efficiency scores can be drawn<sup>9</sup>:

$$\hat{\theta}_i = f(Z_i, \hat{\beta}_i) + \varepsilon_i \quad (5)$$

Xue and Harker (1999) were the first to argue that these conventional regression models applied in the second stage yield biased results because the efficiency scores estimated in the first stage ( $\hat{\theta}_i$ ) are serially correlated. Accordingly, there has been a lively debate in recent years about which would be the most accurate model to apply in this second stage in order to provide consistent estimates. According to Simar and Wilson (2007), the efficiency rates estimated by the DEA model in the first stage are correlated by construction (as they are relative measures), and therefore estimates from conventional regression methods (Eq. 5) would be biased. Additionally, the possible correlation of the contextual variables  $Z_i$  with the error term  $\varepsilon_i$  in Eq. (5) is another source of bias.

Simar and Wilson (2007) state that bootstrapping can overcome these drawbacks. In their paper, the authors propose two algorithms<sup>10</sup> that incorporate the bootstrap procedure in a truncated regression model. They run a Monte Carlo experiment to examine and compare the performance of these two algorithms, and they prove that both bootstrap algorithms outperform conventional regression methods (Tobit and truncated regressions without bootstrapping), yielding valid inference methods. For small samples (problems with fewer than 400 units and up to three outputs and three inputs), Algorithm #1 fits results better than Algorithm #2, which is more efficient as of samples that exceed 800 units.<sup>11</sup> Since the samples analyzed in our research are made up of around a hundred schools, we apply the simple *Algorithm #1*, which is described below.<sup>12</sup>

<sup>8</sup> Some authors actually estimate both models simultaneously to verify results robustness.

<sup>9</sup> For a detailed review of estimation methods used in the second stage of semiparametric models, see Simar and Wilson (2007).

<sup>10</sup> The authors propose a simple Algorithm #1 and a double Algorithm #2. The difference lies in the fact that Algorithm #2 incorporates an additional bootstrap in the first stage, which amends the estimates of the efficiency scores.

<sup>11</sup> For a more detailed analysis of the results, see Simar and Wilson (2007, p. 45).

<sup>12</sup> According to Simar and Wilson (2007) we define  $L = 2000$  in this paper.

*Algorithm #1*

Estimate efficiency scores  $\hat{\theta}_i \forall i = 1, 2, \dots, n$  solving DEA (Equation 3).

- 1) Estimate  $\hat{\beta}_i$  and  $\hat{\sigma}_\varepsilon$  by maximum likelihood in a truncated regression of  $\hat{\theta}_i$  on  $z_i$  (Equation 5), using  $m < n$  observations, where  $\hat{\theta}_i > 1$
- 2) Iterate step (3.a to 3.c)  $L^{12}$  times by a loop, obtaining the bootstrap estimations of  $A = \left\{ \left( \hat{\beta}^*, \hat{\sigma}_\varepsilon^* \right) \right\}_{b=1}^L$ .
- 3) Compute the bootstrap estimations:
  - a) For each  $i = 1, 2, \dots, m$  extract  $\varepsilon_i$  from a normal distribution  $N(0, \hat{\sigma}_\varepsilon^2)$  left-truncated in  $(1 - z_i \hat{\beta})$ .
  - b) Again, for  $i = 1, 2, \dots, m$ , estimate  $\theta_i^* = z_i \hat{\beta} + \varepsilon_i$ .
  - c) Using maximum likelihood, estimate the truncated regression of  $\theta_i^*$  on  $z_i$ , obtaining  $\hat{\beta}^*$  and  $\hat{\sigma}_\varepsilon^*$  estimations.
- 4) Use L values of A to build the confidence intervals of  $\beta$  and  $\sigma_\varepsilon$ .

Later, Hoff (2007), McDonald (2009), Banker and Natarajan (2008) and Ramalho et al. (2010) took up the discussion about the use of OLS, Tobit and fractional regression models in the second stage. Unlike Hoff (2007), who concluded that both (Tobit and OLS) models yield consistent estimations, McDonald (2009) showed that only the Tobit produces consistent results. Meanwhile, Banker and Natarajan (2008) provided a statistical model which yields consistent second-stage OLS estimations. Simar and Wilson (2011) again took part in the ongoing debate and concluded that only the truncated regression and, under very particular and unusual assumptions, the OLS model provides consistent estimates. Further, they proved that in both cases only bootstrap methods were capable of statistical inference.

From the above, we conclude that the research community does not yet totally agree about which is(are) the most consistent regression model(s) because this conclusion depends on previous assumptions about the data generation process. For this reason, we have chosen to estimate four alternative regression models in the second stage and compare the results. First, we specify the conventional Tobit (censored regression model), as it is the most commonly used in the literature. Then, for the sake of robustness, we estimate three regression models applying the bootstrap procedure: Algorithm #1 proposed by Simar and Wilson (2007) based on a truncated regression; and a Tobit regression and an OLS model with bootstrapping.

### 3 Data and variables

#### 3.1 Brief description of the Uruguayan education system

The Uruguayan national education system is composed of four levels: 3 years of pre-primary education (3–5 years old), 6 years of primary education (6–11 years old), 6 years of secondary education (12–17 years old), and college education at the end of secondary education. Secondary education is divided into 3 years of lower secondary education (*Ciclo Básico Común*) and 3 years of upper secondary education (*Bachillerato*). Compulsory education covers 14 years from the last 2 years of pre-primary education (4 and 5 years old), through primary school, to the end of secondary education.<sup>13</sup>

In terms of public and private education production, the public sector takes absolute primacy over the private sector. In 2011, 84.5 % of high school students attended public schools (Education Observatory, National Administration of Public Education). This highlights how important the performance of public institutions is for national academic results, and therefore the need to benchmark schools and to assess both the management and the teaching practices implemented by these schools.

Uruguay has historically occupied a leading position in Latin America in terms of educational achievement, according to the main standard indicators and international studies. However, the Uruguayan education system (particularly the secondary and tertiary levels) is currently undergoing a phase of stagnation and recession. The major budgetary effort made by the government in the first decade of the twenty-first century has not been accompanied by effective reforms and policies to improve educational outcomes.

The results of PISA 2009 and 2012 corroborate that Uruguay is still in an advantageous position within the region,<sup>14</sup> but also confirm that results have not improved compared to previous waves. In addition, test scores in the three analyzed areas (mathematics, reading and science) are more highly dispersed than in other countries, which mirror the high social segmentation of the education system. Comparing student's performance by the schools socio-economic context in PISA 2012, it is noteworthy that while almost 89 % of students who attended to schools in "very unfavorable circumstances" do not reach the minimum "competence threshold" defined by the OECD in mathematics,<sup>15</sup> this figure

---

<sup>13</sup> Art. 10 of the General Education Law N. 18,437 of December 12, 2008.

<sup>14</sup> PISA 2009 showed that Uruguay was the Latin American country with the best results for mathematics and was second placed in science and reading (after Chile). In PISA 2012 Uruguay is placed in the third position in the three evaluated areas between all Latin American countries that participated in this wave.

<sup>15</sup> PISA defines six competencies levels and states that basic skills are obtained at Level 2. In the case of mathematics Level 2 threshold is described as follows: "At Level 2 students can interpret and recognize situations in contexts that require no more than direct inference. They can extract relevant information from a single source and make use of a single representational model. Students at this level can employ basic algorithms, formulae, procedures, or conventions. They are capable of direct reasoning and making literal interpretations of the results". For more details, see OECD (2013a).

drops to 13 %<sup>16</sup> for students who attend to schools in “very favorable circumstances”.<sup>17</sup> By contrast, analyzing the percentage of top-scoring students (performance levels four to six) defined by PISA analysts, we find that this proportion rises to almost 30 % of students in “very favorable circumstances”, whereas students from “very unfavorable circumstances” account for less than 1 %. This heterogeneity may be the consequence of differences not only in the initial resources endowment but also of efficiency. It is essential to explore the sources of such differences in order to improve academic outputs in more inefficient schools and to reduce inequalities in the education system.

### 3.2 PISA databases and model specification

PISA 2009 and 2012 are the fourth and fifth edition of an initiative that the OECD started up in the late 1990s to assess 15-year-old students. The assessment focuses on measuring the extent to which students are able to apply their knowledge and skills to fulfill future real-life challenges rather than evaluating how well they have mastered a specific school curriculum. The evaluation addresses three knowledge areas: reading, mathematical and scientific literacy, and each wave tests in depth a major domain. In 2000 and in 2009 the major domain was reading, in 2003 it was mathematics, in 2006 science and finally, in 2012, it is again mathematics. In addition to academic achievement data, the PISA database contains a vast amount of information about students, their households and the schools they attend. Uruguay took part in PISA 2009 (2012) assessing 5927 (5315) students from (232) 180 public and private schools.

To perform the DEA model one of the main requirements is that the evaluated decision-making units should be as homogeneous as possible (Dyson et al. 2001). To estimate the production frontier and the efficiency indexes the technique assumes that all units operate under the same production technology and therefore under similar context and circumstances. In order to analyze homogenous schools, the original PISA databases were refined. Firstly, we assume that technologies in public and private sector are different due to different legal, organizational and curricular contexts and therefore, management drivers also differ in the two schools type. Two frontiers should be estimated, one for each sector. Unfortunately, the sample size of private schools is small to carry out a specific analysis of this sector so, we only analyze public schools. Secondly, we eliminate schools which only offer basic secondary education (1st, 2nd and 3rd year of high school) or only offer upper secondary education (4th, 5th and 6th year of high school). The cut-off age between the two cycles in Uruguay is just 15 years old and, since PISA evaluates students of this age, those students attending schools where only basic secondary education is offered are inevitably repeaters and, on the contrary, students attending schools

---

<sup>16</sup> National Administration of Public Education (ANEP), “Informe Ejecutivo Preliminar Uruguay en PISA 2012”. Available at <http://www.anep.edu.uy/anep/index.php/presentaciones-2012>.

<sup>17</sup> Schools are classified into five levels of socio-economic context based on the quintile distribution of the average socio-economic background of the students who attend to these schools (the average ESCS PISA index for each school). Levels are defined as “Very unfavorable” (the bottom quintile), “Unfavorable”, “Medium”, “Favorable” and “Very favorable” (the top quintile).

where only upper secondary education is offered are all non-repeaters. As a result, in schools where only basic secondary education is imparted, 100 % of the assessed students in PISA are repeaters in at least one previous course and, in those schools where only upper secondary education is imparted, 100 % of the assessed students are on the right course. Therefore, these institutions are not comparable when estimating the production frontier.

In sum, this analysis is carried out for 169 mixed public schools (98 from PISA 2009 and 71 from PISA 2012) which provide both cycles of secondary education. For comparative and robustness purposes, we perform the same analysis for PISA 2009 and PISA 2012 waves separately. Additionally, we run the model for both databases in a pool<sup>18</sup> including contextual variables simultaneously available in the two waves in order to check whether or not technical efficiency has changed significantly over the two periods.

### 3.3 Outputs, inputs and contextual variables

#### 3.3.1 Outputs

It is difficult to empirically quantify the education received by an individual, especially when the focus is on analyzing its quality beyond the years of education. However, there is a consensus in the literature about considering the results from a standardized test as educational outputs, as they are difficult to forge and, above all, they are taken into account by parents and politicians when making decisions on education. In this research, we selected two variables as outputs of the educational process: the average results in reading (Read\_mean) and mathematics (Maths\_mean).<sup>19</sup>

#### 3.3.2 Inputs

Regarding educational inputs, three variables were selected taking into account the educational production function in Eq. (1). They represent the classical inputs in education economics required to carry out the learning process: students (raw material), teachers (human capital) and infrastructure (physical capital).<sup>20</sup> A previous requirement for a variable to be considered as an input in an efficiency analysis is that it has to be positively correlated with all outputs. This monotonicity assumption (Coelli et al. 2005) implies that if we give more input to a DMU then we will expect to obtain equal or more quantity of outputs than in the previous situation or, in other words, additional units of an input will not decrease output. The following inputs were included in the first stage of the DEA:

---

<sup>18</sup> This analysis was suggested by a referee to explore extra sources of variation (especially temporal).

<sup>19</sup> The result for science has been omitted since it provides little additional information to the reading and mathematical results. Besides, DEA becomes less discriminative as more dimensions are added to the problem (curse of dimensionality); therefore, we prioritize parsimony by choosing only two outputs.

<sup>20</sup> We focus on the quality rather than just the quantity of these inputs.

- Parental education (PARED): is an index that reflects the higher parental education expressed by the number of years of schooling according to the International Standard Classification of Education (ISCED-1997, OECD).<sup>21</sup> It therefore represents the quality of the ‘raw material’ to be transformed through the learning process.
- School educational resources (SCHRES): is an index of the quality of the school resources constructed from the school’s principal responses. It is therefore associated with the physical and human capital. The index was computed from the responses by principals to several questions related to the scarcity or lack of ten educational resources<sup>22</sup> including teachers, educational material and infrastructures. The school receives one point for each item for which the principal’s answer is that the school is not deficient ‘at all’. The maximum (minimum) score for each school is ten (zero) points, which indicates an excellent (dreadful) educational input.<sup>23</sup>
- Proportion of fully certified teachers (PROPCERT): this index reflects the quality of teachers, and therefore the school’s human capital. The index is constructed by dividing the total number of certified teachers (with a teaching degree)<sup>24</sup> by the total number of teachers. This variable is especially relevant in the case of Uruguay since not all teachers have received the teaching training required to qualify as teachers.

As mentioned above, it is necessary to check the monotonicity assumption in order to ensure a correct DEA model specification. Table 2 presents the bivariate correlations of the selected outputs and inputs where all correlations are positive.

### 3.3.3 Contextual variables

The distinction between an educational input (first stage) and an explanatory variable of inefficiency or contextual variable (second stage) can be confusing. In this research we have considered that a variable is an explanatory factor of efficiency, and not an educational input, when it is not strictly essential to produce education but it can affect academic results through the efficiency term. These variables fulfill one or some of these conditions:

1. The variable reflects some key aspect of school management and organization and/or the teaching–learning processes enacted in the classrooms.

<sup>21</sup> In the case of Uruguay the equivalent scale used to compute the years of schooling is the following: ISCED 1 equals to 6 years; ISCED 2 equals to 9 years; ISCED Level 3A, 3B, 3C or 4 equals to 12 years; ISCED 5B equals to 15 years; and ISCED 5A or 6 equals to 17 years of schooling.

<sup>22</sup> The item included are: ‘Qualified science teachers’, ‘Qualified mathematics teachers’, ‘Qualified reading teachers’, ‘Any other personal support’, ‘Science laboratory equipment’, ‘Instructional materials’, ‘Computers’, ‘Internet connectivity’, ‘Software’, ‘Library materials’.

<sup>23</sup> This variable has been rescaled so the minimum value is one in order to avoid zero values in the empirical analysis.

<sup>24</sup> Certified teachers in Uruguay are required to complete a 4-year degree at the Instituto de Profesores Artigas (IPA), a higher education institution which provides specialized secondary teacher training.

**Table 2** Bivariate correlations between outputs and inputs

		PARED	SCHRES	PROPCERT
2009	Maths_mean	0.588**	0.014	0.373**
	Read_mean	0.633**	0.023	0.393**
2012	Maths_mean	0.585**	0.122	0.116
	Read_mean	0.479**	0.188	0.180
Pool	Maths_mean	0.544**	0.048	0.263**
	Read_mean	0.545**	0.091	0.297**

\*\* Statistical significance at 1 % (bilateral). Sample size PISA 2009 = 71; PISA 2012 = 98; pool = 169. Source: own elaboration based on PISA 2009 and PISA 2012 data

2. The variable is dichotomous, categorical or does not have a continuous measurement scale.
3. The monotonicity assumption does not hold in practice, i.e., the selected variable does not show a positive correlation with academic outcomes.
4. The variable is an indicator based on opinions with a high degree of subjectivity and difficult to contrast.

Building upon this criteria, we select fifteen contextual variables<sup>25</sup> ( $Z$  vector in Eqs. 4 and 5) associated with students and schools. Most of contextual variables appear in PISA 2009 and 2012; however, there are some variables that are only available in one wave. To be more precise we employ 12 (13) variables with PISA 2012 (2009) to run the second stage regression analysis. Finally, we only use the ten contextual variables that were collected in both waves.

### 3.3.3.1 Contextual variables included only in PISA 2009

- TEST: a dummy variable that takes the value one when students are assessed by teachers through tests, quizzes or exams more often than once a month.
- HOMEWORK: a dummy variable which refers to the assessment tools as well as the frequency with which they are applied. In this case, the variable takes value one when the students are assessed by means of homework every month. Both Tests and Homework are expected to have a positive effect on school efficiency.
- EXTREADING: the percentage of students in the school who spend between one and 2 h per day reading for pleasure after school. It is understood that reading contributes to the student learning process, as it helps to improve spelling, reading comprehension and understanding skills. It is expected therefore to have a positive effect on school efficiency.

<sup>25</sup> Not all contextual variables appear in both PISA waves.

### 3.3.3.2 Contextual variables included only in PISA 2012

- STUCHECK: percentage of students in the school that have answered yes to the statement ‘When I study mathematics, I make myself check to see if I remember the work I have already done’. This variable reflects the learning skills acquired along the student’s academic life.
- STUIMPORT: percentage of students in the school that have answered yes to the statement ‘When I study for a mathematics test, I try to work out what the most important parts to learn are’. As in the previous case, this variable also reflects the learning skills acquired along the student’s academic life.

### 3.3.3.3 Contextual variables included in both databases and in the pool

- PERIOD: dummy variable that takes value one if the student belongs to PISA 2012. This variable is only included in the pool.
- PCTCORRECT: percentage of students assessed in the school who are in the academic year that a 15-year student should really be in. This variable reflects the grade-retention policy, and is another focus of attention in current educational discussions because there is no consensus about its net effect on educational results.
- TEACHVOC: dummy variable that takes value one if the institution is a vocational technical school.
- RURAL: dummy variable that takes value one if the institution is located in a town with less than 3000 inhabitants.
- CITY: dummy variable that takes value one if the institution is located in a town with more than 100,000 inhabitants.
- TEACHSTU: the number of teachers per hundred students. Some research includes class size as an educational input in the first stage, but we have decided to use it as an explanatory variable of efficiency since there is still no conclusive evidence about the real effect of this variable on student results.<sup>26</sup> Furthermore, this variable does not show a positive correlation with the analyzed outputs.

Finally, we incorporate a number of variables associated with school autonomy and management in terms of budget allocation, curriculum development, disciplinary policies and student assessment practices. There are no expected a priori positive or negative relationships between these variables and school efficiency, since empirical evidence emerging from international comparisons does not provide conclusions that are applicable to all education systems (OECD 2013b).

- Curr\_author: a dummy variable which takes the value one when the national authorities have a considerable responsibility for determining the content of the courses.

---

<sup>26</sup> For a more detailed review, see Hanushek (2003) and Hoxby (2000).

- *Disc\_author*: a dummy variable which takes the value one when the national authorities have a considerable responsibility for establishing student disciplinary policies.
- *Budget\_ppal*: a dummy variable which takes the value one when the school principal has a considerable responsibility for distributing the school budget.
- *Budget\_author*: a dummy variable which takes the value one when the national authorities have a considerable responsibility for distributing the school budget.
- *Asses\_author*: a dummy variable which takes the value one when the national authorities have a considerable responsibility for establishing student assessment policies.

Table 3 shows the main descriptive statistics of all selected variables: outputs, inputs and contextual variables.

## 4 Results

### 4.1 First stage analysis

Figure 2 illustrates the distribution of efficiency scores,  $\theta_i$  in Eq. (1), estimated by the output-oriented DEA-BCC model for the three data sets. Results show that between 2009 and 2012 the percentage of fully efficient schools in Uruguay dropped from 27 to 21 %. The average estimated efficiency score also decreases over this period from 1.087 in PISA 2009 to 1.116 in PISA 2012. This means that, on average, educational results in public schools could be increased in 2012 by 11.6 % given the available resources. Moreover, more than 16 % of the schools in 2012 could improve their results by over 20 % to reach the frontier while this percentage represented 13 % in 2009. Finally half of schools could improve outcomes by over 10 % with their current inputs in 2012 what represent an important decrease in results taking into account that barely one out of three schools belonged to this inefficient group in 2009.

Table 4 shows, as an example of the potential benefits of measuring inefficiency, the most inefficient schools with estimated efficiency scores greater than 1.20 for PISA 2012, i.e., where outputs could be increased by more than 20 % if they were fully efficient. All these schools perform far below the Uruguayan means in PISA 2012 (409 in mathematics and 411 in reading); whereas, if they were fully efficient, most of them would be above these national averages. In fact, some schools would even significantly exceed these averages in both subjects (e.g., schools 2, 4, 6 and 10).

An alternative way to evaluate the potential gains due to inefficiencies reduction is by comparing the actual distribution of students into PISA proficiency levels defined by the OECD with the potential distribution that would be observed once inefficient schools reached the frontier. Table 5 shows the percentage of students under (over) proficiency level 1 (level 4) and level 2 (level 5) in mathematics, reading and in at least one of the two evaluated disciplines in the two periods. Results show great potential gains derived from improving efficiency.

**Table 3** Descriptive statistics of outputs, inputs and explanatory variables of efficiency

Variable	Description	2009		2012		Pool	
		Mean	SD	Mean	SD	Mean	SD
<b>Outputs</b>							
Maths_mean	School average score in mathematics	398.0	49.0	382.1	44.7	391.3	47.8
Read_mean	School average score in reading	390.7	53.6	380.5	54.3	386.4	54.0
<b>Inputs</b>							
PARED	Students' highest parental education level expressed as years of schooling	9.65	1.56	10.15	1.40	9.86	1.51
SCHRES	School Educational Resources Index	4.94	2.71	5.31	3.04	5.09	2.85
PROPERT	Proportion of fully certified teachers in the school	0.53	0.18	0.52	0.20	0.52	0.19
<b>Explanatory variables</b>							
TEST	Student's assessment through tests with frequency more than one per month	0.13	0.34	–	–	–	–
HOMEWORK	Student's assessment through monthly homework	0.14	0.35	–	–	–	–
EXTREADING	Percentage of students in the appropriate year	0.09	0.06	–	–	–	–
STUCHECK	Percentage of students who check to see if I remember the work I have already done (mathematics)	–	–	0.25	0.09	–	–
STUIMPOR	Percentage of students who try to work out what the most important parts to learn are (mathematics)	–	–	0.32	0.09	–	–
PCTCORRECT	Percentage of students in the appropriate year	0.55	0.26	0.55	0.24	0.55	0.25
TEACHVOC <sup>a</sup>	Vocational technical school	0.31	0.46	0.32	0.47	0.31	0.47
RURAL <sup>a</sup>	School located in a town with less than 3000 inhabitants	0.14	0.35	0.13	0.34	0.14	0.34
CITY <sup>a</sup>	School located in a city with more than 100,000 inhabitants	0.09	0.29	0.10	0.30	0.09	0.29
TEACHSTU	Number of teachers per 100 students	7.67	3.27	7.48	2.49	7.59	2.96
Curr_author <sup>a</sup>	National authorities have a considerable responsibility for determining the content of the courses	0.84	0.37	0.72	0.45	0.79	0.41
Disc_author <sup>a</sup>	National authorities has a considerable responsibility for establishing student disciplinary policies	0.63	0.48	0.65	0.48	0.64	0.48
Budget_ppal <sup>a</sup>	The school's principal has a considerable responsibility for distributing the school budget	0.53	0.50	0.30	0.46	0.43	0.50
Budget_author <sup>a</sup>	National authorities have a considerable responsibility for distributing the school budget	0.57	0.50	0.66	0.48	0.61	0.49
Asses_author <sup>a</sup>	National authorities have a considerable responsibility for establishing the students assessment policies	0.80	0.41	0.68	0.47	0.75	0.44

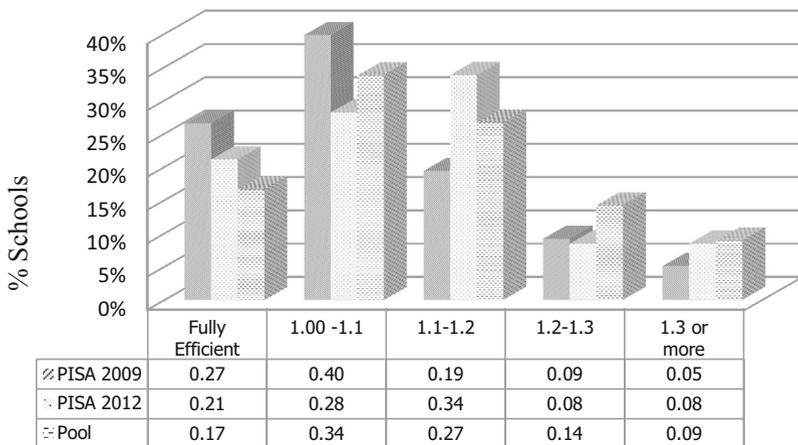
Source: own elaboration based on PISA 2009 and PISA 2012

<sup>a</sup> For dummy variables, the mean represents the proportion of schools in each category

If all evaluated public schools were efficient in PISA 2012, the percentage of students below proficiency level 2 (the minimum ‘competence threshold’ defined by OECD) in mathematics (reading) could be reduced from 67.2 to 48.5 % in mathematics and from 58.8 to 42.9 % in reading. Moreover, the actual percentage of students who is below proficiency level 2 in at least one of the two evaluated areas would decline from 73.5 to 57.4 %. By contrast, analyzing the percentage of top-scoring students (performance levels four to six) defined by PISA analysts, we find that this proportion could be doubled from 12.6 to 24.5 % and from 13.6 to 26.9 % in mathematics and reading tests, respectively. Indeed, these figures would be close to those actually observed in some OECD countries (e.g., United States 24.6 %, Sweden 24.4 % or Italy 26.7 % in mathematics). It is also important to note here the decline of PISA 2012 results with respect PISA 2009. Table 5 clearly shows how the actual number of top performers in at least one area was almost 11 percentage points higher in 2009 than in 2012. As a consequence, the potential percentage of students that could become top performers decreased too from 42.9 % in PISA 2009 to 35 % in PISA 2012. Results are closely related when we turn our attention to the comparison of students under the minimum competence threshold in mathematics or reading. In this case, we observe that the actual percentage of students without the minimum level of competences rose from 63.9 % in PISA 2009 to 73.5 % in PISA 2012 while the potential to lift students out of poor results decreased between the 2 years.

### 4.2 Second-stage analysis

In a second stage, the estimated efficiency scores are regressed over the contextual variables using four model specifications: the truncated regression with bootstrap proposed by Simar and Wilson (2007), the conventional Tobit, the Tobit regression with bootstrap and, finally, the OLS model with bootstrap. Results are shown for



**Fig. 2** Estimated efficiency scores distribution (DEA-BBC). Values equal to one represent full efficient units. Higher values of the estimated score imply more inefficiency. Source: authors’ estimates using PISA 2009 and 2012 databases

**Table 4** Output targets for the most inefficient schools

School ID	Estimated efficiency	Actual mathematics	Actual reading	Target mathematics	Target reading
1	1.52	291	224	442	341
2	1.47	308	305	452	448
3	1.39	307	271	427	377
4	1.38	324	313	447	433
5	1.36	302	273	411	371
6	1.34	323	328	433	440
7	1.28	344	309	440	395
8	1.27	318	321	404	408
9	1.27	319	290	405	368
10	1.26	365	355	460	448
11	1.22	318	315	388	385
12	1.21	349	347	422	420

Source: authors' estimations using PISA 2012 data

**Table 5** Actual and potential percentage of students into PISA proficiency levels

Description	2009		2012	
	Actual %	Potential %	Actual %	Potential %
Students under the minimum "competence threshold" in mathematics	54.7	41.5	67.2	48.5
Top-performing students in mathematics	20.7	30.7	12.6	24.5
Students under the minimum "competence threshold" in reading	49.7	38.2	58.8	42.9
Top-performing students in reading	21.9	31.3	13.6	26.9
Students under the minimum "competence threshold" in at least one area	63.9	52.4	73.5	57.4
Top-performing students in at least one area	29.8	42.9	19.0	35.0

Source: authors' estimations based on PISA 2009 and PISA 2012 databases

cross-sections of PISA 2009, PISA 2012 and the pool in Tables 6, 7 and 8, respectively.

From the comparative analysis of the four specified models we can conclude that there are no major discrepancies between the results. The sign, magnitude and significance of almost all variables are similar in all models and databases, implying that any educational policy recommendations derived from them would be basically the same regardless the second-stage regression model finally chosen adding robustness to these findings. Taking into account this general conclusion, we will consider the specification proposed by Simar and Wilson (2007) as the baseline for discussing the results.

**Table 6** Efficiency drivers: second-stage estimations (PISA 2009)

Explanatory variables	Conventional Tobit <sup>a</sup>			Truncated + bootstrap <sup>b</sup>			Tobit + bootstrap <sup>a</sup>			OLS + bootstrap		
	Coef.	SE	t	Coef.	SE	z	Coef.	SE	z	Coef.	SE	z
PCTCORRECT	-0.20	0.05	-4.28***	-0.257	0.066	-3.91***	-0.196	0.053	-3.72***	-0.149	0.040	-3.70***
TEST <sup>c</sup>	-0.04	0.03	-1.42	-0.100	0.056	-1.79*	-0.039	0.032	-1.23	-0.040	0.023	-1.75*
HOMEWORK <sup>c</sup>	-0.07	0.03	-2.41**	-0.060	0.033	-1.82*	-0.069	0.034	-2.02***	-0.037	0.019	-1.95*
EXTREADING	0.01	0.19	0.05	-0.146	0.265	-0.55	0.010	0.220	0.05	-0.049	0.150	-0.33
TEACHVOC <sup>c</sup>	0.04	0.03	1.56	0.105	0.041	2.58***	0.043	0.032	1.34	0.042	0.024	1.76*
RURAL <sup>c</sup>	0.01	0.03	0.27	0.057	0.056	1.02	0.009	0.037	0.25	0.004	0.027	0.14
CITY <sup>c</sup>	0.00	0.04	-0.05	-0.002	0.054	-0.04	-0.002	0.050	-0.04	-0.011	0.043	-0.25
TEACHSTU	0.00	0.00	-0.04	-0.004	0.009	-0.47	0.000	0.005	-0.03	0.000	0.004	0.04
Curr_author <sup>c</sup>	0.06	0.03	1.76*	0.019	0.059	0.32	0.060	0.042	1.43	0.028	0.025	1.11
Disc_author <sup>c</sup>	0.00	0.03	0.07	-0.020	0.046	-0.44	0.002	0.037	0.06	-0.004	0.029	-0.15
Asses_author <sup>c</sup>	-0.04	0.02	-1.64*	-0.021	0.035	-0.61	-0.040	0.027	-1.47	-0.025	0.020	-1.21
Budget_author <sup>c</sup>	-0.03	0.03	-1.25	-0.009	0.037	-0.23	-0.033	0.030	-1.08	-0.022	0.023	-0.93
Budget_ppal <sup>c</sup>	-0.01	0.03	-0.20	0.019	0.040	0.49	-0.005	0.029	-0.18	0.008	0.021	0.38
Constant	1.17	0.05	21.34***	1.242	0.074	16.86***	1.167	0.069	17.03***	1.173	0.049	24.11***

**Table 6** continued

Explanatory variables	Conventional Tobit <sup>a</sup>		Truncated + bootstrap <sup>b</sup>		Tobit + bootstrap <sup>a</sup>		OLS + bootstrap	
	Coef.	SE	Coef.	SE	Coef.	SE	Coef.	SE
/sigma	0.100	0.010	0.086	0.012	0.100	0.010	0.085	0.010

Coef.: is the estimated coefficient. SE is the robust standard error of the estimated coefficient. Number of schools = 98

Source: authors' estimations using PISA 2009 data

\*  $p < 0.10$

\*\*  $p < 0.05$

\*\*\*  $p < 0.01$

<sup>a</sup> Censored regression model with 26 left-censored observations at value 1

<sup>b</sup> Truncated regression model with 26 left-truncated observations at value 1

<sup>c</sup> Dummy variables. Reference categories for school dummies are: non-vocational technical school; school located in a town with more than 3000 and less than 100,000 inhabitants; students are assessed through test with monthly frequency or less; students are assessed through homework with frequency less than monthly; the national authorities do not have a considerable responsibility for determining the content of the courses or establishing student disciplinary policies or distributing the school budget or establishing the students assessment policies; the school principal does not have a considerable responsibility for distributing the school budget; the school board does not have a considerable responsibility for distributing the school budget

**Table 7** Efficiency drivers: second-stage estimations (PISA 2012)

Explanatory variables	Conventional Tobit <sup>a</sup>			Truncated + bootstrap <sup>b</sup>			Tobit + bootstrap <sup>a</sup>			OLS + bootstrap		
	Coef.	SE	t	Coef.	SE	z	Coef.	SE	z	Coef.	SE	z
PCTCORRECT	-0.152	0.068	-2.25**	-0.176	0.108	-1.65*	-0.152	0.075	-2.04**	-0.142	0.066	-2.14**
STUcheck	-0.598	0.144	-4.16***	-0.735	0.280	-2.62**	-0.598	0.181	-3.31***	-0.534	0.157	-3.40***
STUimport	-0.285	0.170	-1.68*	-0.478	0.248	-1.97**	-0.285	0.195	-1.46	-0.279	0.172	-1.62*
TEACHVOC <sup>c</sup>	0.013	0.039	0.33	0.022	0.057	0.39	0.013	0.046	0.27	0.007	0.040	0.17
RURAL <sup>c</sup>	-0.075	0.046	-1.63	-0.012	0.075	-0.16	-0.075	0.078	-0.97	-0.055	0.039	-1.43
CITY <sup>c</sup>	0.001	0.045	0.02	0.080	0.069	1.16	0.001	0.065	0.01	0.004	0.045	0.08
TEACH_stu	-0.001	0.007	-0.10	-0.009	0.011	-0.82	-0.001	0.008	-0.09	0.000	0.007	-0.02
Curr_author <sup>c</sup>	0.030	0.033	0.92	0.026	0.053	0.49	0.030	0.039	0.77	0.022	0.034	0.64
Disc_author <sup>c</sup>	0.006	0.036	0.17	-0.016	0.064	-0.25	0.006	0.044	0.14	-0.003	0.033	-0.09
Asses_author <sup>c</sup>	0.025	0.039	0.65	0.026	0.072	0.36	0.025	0.047	0.53	0.033	0.037	0.89
Budget_author <sup>f</sup>	-0.074	0.032	-2.36**	-0.074	0.048	-1.56	-0.074	0.037	-1.99**	-0.067	0.030	-2.22**
Budget_ppal <sup>c</sup>	-0.082	0.030	-2.70***	-0.125	0.053	-2.36**	-0.082	0.035	-2.35**	-0.079	0.028	-2.80***
Constant	1.467	0.089	16.46***	1.634	0.132	12.37***	1.467	0.102	14.34***	1.453	0.089	16.35***

**Table 7** continued

Explanatory variables	Conventional Tobit <sup>a</sup>		Truncated + bootstrap <sup>b</sup>		Tobit + bootstrap <sup>a</sup>		OLS + bootstrap	
	Coef.	SE	Coef.	SE	Coef.	SE	Coef.	SE
/sigma	0.097	0.008	0.096	0.012	0.097	0.008	0.089	

<sup>a</sup>Coef.: is the estimated coefficient. SE is the robust standard error of the estimated coefficient. Number of schools = 71

Source: authors' estimations using PISA 2012 data

\*  $p < 0.10$

\*\*  $p < 0.05$

\*\*\*  $p < 0.01$

<sup>a</sup> Censored regression model with 16 left-censored observations at value 1

<sup>b</sup> Truncated regression model with 16 left-truncated observations at value 1

<sup>c</sup> Dummy variables. Reference categories for school dummies are: non-vocational technical school; school located in a town with more than 3000 and less than 100,000 inhabitants; the national authorities do not have a considerable responsibility for determining the content of the courses or establishing student disciplinary policies or distributing the school budget or establishing the students assessment policies; the school principal does not have a considerable responsibility for distributing the school budget

**Table 8** Efficiency drivers: second-stage estimations (pool)

Explanatory variables	Conventional Tobit <sup>a</sup>			Truncated + bootstrap <sup>b</sup>			Tobit + bootstrap <sup>a</sup>			OLS + bootstrap		
	Coef.	SE	t	Coef.	SE	z	Coef.	SE	z	Coef.	SE	z
Period <sup>c</sup>	0.067	0.017	3.88***	0.093	0.025	3.76***	0.067	0.018	3.76***	0.016	3.82***	3.82***
PCTCORRECT	-0.246	0.042	-5.91***	-0.351	0.062	-5.66***	-0.246	0.043	-5.72***	-0.216	0.037	-5.79***
TEACHVOC <sup>c</sup>	0.053	0.023	2.25***	0.080	0.034	2.37**	0.053	0.024	2.18**	0.051	0.022	2.31**
RURAL <sup>c</sup>	-0.031	0.028	-1.12	0.003	0.042	0.07	-0.031	0.029	-1.09	-0.028	0.022	-1.27
CITY <sup>c</sup>	-0.014	0.033	-0.42	0.018	0.044	0.40	-0.014	0.034	-0.40	-0.017	0.031	-0.56
TEACH_stu	-0.001	0.003	-0.29	-0.005	0.005	-1.00	-0.001	0.004	-0.25	-0.001	0.003	-0.18
Curr_author <sup>c</sup>	0.048	0.023	2.05**	0.053	0.033	1.61	0.048	0.025	1.95*	0.035	0.020	1.73*
Disc_author <sup>c</sup>	-0.019	0.021	-0.89	0.013	0.033	0.40	-0.019	0.023	-0.84	-0.011	0.020	-0.54
Asses_author <sup>c</sup>	0.006	0.025	0.23	-0.004	0.036	-0.10	0.006	0.027	0.21	0.000	0.023	-0.01
Budget_author <sup>c</sup>	-0.024	0.020	-1.15	-0.051	0.032	-1.62	-0.024	0.022	-1.09	-0.021	0.018	-1.13
Budget_ppal <sup>c</sup>	-0.034	0.019	-1.85*	-0.068	0.029	-2.34**	-0.034	0.020	-1.76*	-0.035	0.016	-2.17**
Constant	1.217	0.042	29.25***	1.272	0.054	23.58***	1.217	0.044	27.75***	1.221	0.039	31.24***
/sigma	0.103	0.008		0.104	0.010		0.103	0.007		0.094		

Source: authors' estimations using PISA 2009 and PISA 2012 databases

'Coef.' is the estimated coefficient. SE is the robust standard error of the estimated coefficient. Number of schools = 169

\*  $p < 0.10$

\*\*  $p < 0.05$

\*\*\*  $p < 0.01$

<sup>a</sup> Censored regression model with 16 left-censored observations at value 1

<sup>b</sup> Truncated regression model with 16 left-truncated observations at value 1

<sup>c</sup> Dummy variables. Reference categories for school dummies are: PISA 2009 database; non-vocational technical school; school located in a town with more than 3000 and less than 100,000 inhabitants; the national authorities do not have a considerable responsibility for determining the content of the courses or establishing student disciplinary policies or distributing the school budget or establishing the students assessment policies; the school principal does not have a considerable responsibility for distributing the school budget

First, there is a set of variables that do not affect efficiency scores in any estimation. First, school location does not seem to affect the efficiency (RURAL and CITY). On average, schools in rural areas or small villages have worse educational outcomes than those located in bigger cities. The fact that the town size does not affect significantly the efficiency implies that the higher results are due to a greater allocation of educational resources and not to a better use of them. Likewise, the teacher–student ratio (TEACHSTU) does not affect either school's efficiency.

Second, hardly any of the variables associated with school autonomy are significant (except for Budget\_ppal). Decentralizing the responsibility of establishing the disciplinary policies (Disc\_author) and assessment practices (Asses\_author) or determining the content of the courses (Curr\_author) does not seem to affect school efficiency. This is an interesting finding, since the decentralization issue is part of most current education discussions. International evidence shows that decentralization is successful in countries where there is also a school accountability practice properly regulated and with standardized criteria (Hanushek et al. 2013; OECD 2013b). This is not the case of Uruguay, where there is great heterogeneity in accountabilities and where, in many cases, there is not even a systematic way of presenting them.

Therefore, the results of this research could be associated with this international evidence, which points out that decentralization would only have positive effects on improving academic results if it is carried out accompanied by an appropriate accountability system. Another possible interpretation of this result lies in the fact that the autonomy indexes were computed from the principals' responses and their perceived autonomy and therefore might not be reflecting the true degree of autonomy they actually have. In Uruguay, public high schools generally have low levels of autonomy; however, the variables included in this analysis show certain degree of variance (Table 3). This fact could suggest some distortion between reality and principals' perceptions regarding their responsibility and autonomy.

By contrast, the fact that the school's principal has a considerable responsibility for distributing the school budget (Budget\_ppal) has a strong significant positive effect on efficiency in PISA 2012 and in pool estimations. Therefore, this result would suggest that to give the responsibility of allocating the school budget to the school's principal would be an appropriate policy, at least in the case of secondary schools in Uruguay.

Third, there is a group of variables associated with students and teaching practices that are systematically significant and show the expected sign. Firstly, the percentage of students that are in the right year (PCTCORRECT) appears to be a positive and significant driver of efficiency in both databases separately and in the pool. This result calls into question the adequacy of current Uruguayan grade-retention policies at all levels of the education system. Uruguay has one of the highest repetition rates in the region, which contrasts with international test results which show this country to be one of the region's top performers. Therefore, it would perhaps be better to attempt to identify younger (primary education) students who are at risk of repeating and provide them with additional support early on in order to prevent grad retention. Secondly, the dummy variable that indicates whether the school is a secondary high school or a technical school (TECHVOC) is

statistically significant in PISA 2009 and the pool pointing out that technical schools are more inefficient. Uruguayan high schools have on average better average academic results than technical schools. This result seems to point out that secondary high schools perform better due to a better management and not only because they have higher initial input endowments.

Thirdly, other interesting variables only appear in one PISA. On one hand, according to PISA 2009 estimations, student assessment methods and their frequency appear to positively influence efficiency. Indeed, schools where teachers assess their students continuously by setting conventional tests or exams (TEST) more often than once a month or by means of the homework made monthly (HOMEWORK) perform better than schools that do not make use of this tool or do so with a frequency other than once a month. At early ages homework needs to be set daily to establish students' study habits, but 15-year-olds should be set homework at less regular intervals to complement regular individual study. So, monthly homework to assess learning seems to positively affect students' results.

On the other hand, regarding PISA 2012 both variables associated with student's study skills in mathematics (STUcheck and STUimport) have a positive impact on efficiency. These variables reflect the students skills acquired over their academic life and thus, this ability could be associated with classroom teaching techniques adopted by teachers. Thus, it would be desirable to promote these learning techniques both in the classroom and at home. This means, not only to work at school but also to foster families' commitment to support students work at home. Although this research is focused in secondary education, such practices should be encouraged from the beginning of the student's academic life in previous cycles, when students are assimilating the learning techniques to be used throughout their academic life and when it is most effective to impact on their non-cognitive skills (Heckman and Kautz 2013).

Finally, it is worth to highlight that the coefficient associated to the time period variable (PERIOD) points out to a significant drop in efficiency results in 2012 with respect to 2009 even after controlling for other contextual covariates also related with efficiency. From Table 3 it is straightforward to conclude that over this period mean outputs significantly decreased while mean inputs clearly increased (PARED and SCHRES) or remained almost constant (PROPCERT). This decline in performance cannot be easily explained but should alert the Uruguayan educational system how to invert this result to gain efficiency.

## 5 Discussion and conclusions

Modern countries agree about the need and importance of having a more and better educated population in order to ensure economic growth based on the high productivity of a skilled labor force. The high percentage of public spending on education is a reflection of this conviction. During the last decade the Uruguayan government has made a huge effort to increase educational resources; however, academic results have not improved. On the contrary, public education system (especially public secondary education) is in a deep crisis and the current

educational national debate mainly focuses on the need to put more resources into the system instead of exploring how to make better use of available inputs, i.e., how to achieve a more efficient education system. This situation raises two open questions. Are Uruguayan public secondary schools efficient? Which policies and practices should be promoted in order to increase school efficiency? As far as we know, however, this issue has yet to be analyzed for the Uruguayan education system. This is the main aim of this research.

Our findings corroborate the presence of inefficient behaviors in public secondary schools. According to PISA 2012 results we conclude that with the current inputs schools could have increased their academic results on average by 11.6 % if adequate educational policies and practices had been designed by national authorities and implemented by schools. Furthermore, if schools were fully efficient, the percentage of students below proficiency level 2 (the minimum 'competence threshold' defined by OECD) could be reduced from 67 to 49 % in mathematics and from 59 to 43 % in reading. By contrast, the percentage of top-scoring students (performance levels four to six), could be doubled from 12 to 24 % and from 14 to 27 % in mathematics and reading tests, respectively.

In addition, the second-stage analysis yields interesting evidence for planning and implementing effective policies to improve the efficiency of the Uruguayan public secondary education. The first noteworthy conclusion is that just increasing educational resources (e.g., reducing class size through recruiting more teachers) does not appear to be an appropriate policy because it does not have a positive and significant effect on school efficiency. By contrast, the results suggest that the national discussion and action on increasing education system efficiency should focus on reviewing the current grade-retention policies and the teaching techniques.

Second, this research evidences that inefficiency is higher where there is a higher percentage of repeating students. So, students at risk of repetition should be identified at an early age and provided with extra support with the aim of preventing future school failure. Third, promoting teaching and learning techniques to enhance students' study skills evidences positive effects on results. In addition, student assessment methods and their frequency appear to positively influence efficiency. Indeed, schools where teachers assess their students continuously by setting monthly homework or through test or exams more than once a month perform better than schools that do not make use of this tool or do so with other frequency. So, continuous monthly assessments seems to positively affect students' results. Fourth, the fact that the school principal had a considerable responsibility for distributing the school budget (Budget\_ppal) has a strong significant positive effect on efficiency. Therefore, this result suggests that it is a good practice to deliver the responsibility of allocating the school budget to the school principal.

Finally, we find a significant decline in efficiency results in Uruguay between the two analyzed periods. In other words, educational outputs in the last years have decreased despite the effort that Uruguayan authorities made putting more public expenditure in the system. Therefore it seems necessary to reach a large commitment from all stakeholders involved in the educational process in order to effectively seek and remove the inefficiencies.

In conclusion, this research offers a new perspective on how to tackle the current educational problem in public high schools in Uruguay from an efficiency viewpoint, providing some potential practices and policies that positively affect academic results. In this respect, this paper reports preliminary findings, and more research is, of course, still needed. For example, a qualitative and in depth analysis of the most efficient and inefficient schools could provide additional useful information about how to implement efficient practices and avoid the inefficient ones.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

- Afonso A, St Aubyn M (2006) Cross-country efficiency of secondary education provision: a semi-parametric analysis with non-discretionary inputs. *Econ Model* 23(3):476–491
- Banker RD, Natarajan R (2008) Evaluating contextual variables affecting productivity using data envelopment analysis. *Oper Res* 56(1):48–58
- Banker RD, Charnes A, Cooper WW (1984) Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Manag Sci* 30(9):1078–1092
- Barro RJ (2001) Human capital and growth. *Am Econ Rev* 91(2):12–17
- Barro RJ, Lee JW (2012) A new data set of educational attainment in the world, 1950–2010. *J Dev Econ* 104:184–198
- Battese GE, Coelli TJ (1988) Prediction of firm-level technical efficiencies with a generalized frontier production function and panel data. *J Econom* 38(3):387–399
- Battese GE, Coelli TJ (1992) Frontier production functions, technical efficiency and panel data: with application to paddy farmers in India. *J Prod Anal* 3(1–2):153–169
- Battese GE, Coelli TJ (1995) A model for technical inefficiency effects in a stochastic frontier production function for panel data. *Empir Econ* 20(2):325–332
- Charnes A, Cooper WW, Rhodes E (1978) Measuring the efficiency of decision making units. *Eur J Oper Res* 2(6):429–444
- Charnes A, Cooper WW, Rhodes E (1981) Evaluating program and managerial efficiency: an application of data envelopment analysis to program follow through. *Manag Sci* 27(6):668–697
- Coelli T, Rao D, O’Donnell C, Battese G (2005) *An introduction to efficiency and productivity analysis*. Springer, New York
- Cordero JM, Crespo-Cebada E, Pedraja F, Santín D (2011) Exploring educational efficiency divergences across Spanish regions in PISA 2006. *Revista de economía aplicada* 19(57):117–146
- Crespo-Cebada E, Pedraja-Chaparro F, Santín D (2014) Does school ownership matter? An unbiased efficiency comparison for regions of Spain. *J Prod Anal* 41(1):153–172
- De Jorge J, Santín D (2010) Determinantes de la eficiencia educativa en la Unión Europea. *Hacienda Pública Esp* 193:131–155
- Dyson RG, Allen R, Camanho AS, Podinovski VV, Sarrico CS, Shale EA (2001) Pitfalls and protocols in DEA. *Eur J Oper Res* 132(2):245–259
- Farrell MJ (1957) The measurement of productive efficiency. *J R Stat Soc S A (Gen)* 120(3):253–290
- Fernández T (2009) La desafiliación en la educación media en Uruguay. Una aproximación con base en el panel de estudiantes evaluados por PISA 2003. *Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación (REICE)* 7(4):32–49
- Hanushek EA (1979) Conceptual and empirical issues in the estimation of educational production functions. *J Hum Resour* 14(3):351–388
- Hanushek EA (2003) The failure of input-based schooling policies. *Econ J* 113(485):64–98

- Hanushek EA, Kimko DD (2000) Schooling, labor-force quality, and the growth of nations. *Am Econ Rev* 90(5):1184–1208
- Hanushek EA, Woessmann L (2012) Do better schools lead to more growth? Cognitive skills, economic outcomes, and causation. *J Econ Growth* 17(4):267–321
- Hanushek EA, Link S, Woessmann L (2013) Does school autonomy make sense everywhere? Panel estimates from PISA. *J Dev Econ* 104:212–232
- Heckman JJ, Kautz T (2013) Fostering and measuring skills: interventions that improve character and cognition. National Bureau of Economic Research, No. w19656
- Hoff A (2007) Second stage DEA: comparison of approaches for modelling the DEA score. *Eur J Oper Res* 181(1):425–435
- Hoxby CM (2000) The effects of class size on student achievement: new evidence from population variation. *Q J Econ* 115(4):1239–1285
- Levin HM (1974) Measuring efficiency in educational production. *Publ Financ Q* 2(1):3–24
- Llambí C, Perera M (2008) La función de producción educativa: el posible sesgo en la estimación de efectos “institucionales” con los datos pisa. El caso de las escuelas de tiempo completo, Montevideo, Centro de Investigaciones Económicas (cinve)
- Llambí C, Perera M, Messina P (2009) Desigualdad de oportunidades y el rol del sistema educativo en los logros de los jóvenes uruguayos. Centro de Investigaciones Económicas, Working paper, No. 4
- Mancebón MJ, Muñoz MA (2003) Aspectos clave de la evaluación de la eficiencia productiva en la educación secundaria. *Papeles de Economía Esp* 95:162–187
- McCarty TA, Yaisawang S (1993) Technical efficiency in New Jersey school districts. In: Fried HO, Knox Lovell S, Smith CA (eds) *The measurement of productive efficiency: techniques and applications*. Oxford University Press, New York, pp 271–287
- McDonald J (2009) Using least squares and tobit in second stage DEA efficiency analyses. *Eur J Oper Res* 197(2):792–798
- OECD (2013a) PISA 2012 assessment and analytical framework: mathematics, reading, science, problem solving and financial literacy. PISA, OECD Publishing, Paris
- OECD (2013b) PISA 2012 results: what makes schools successful? Resources, policies and practices, vol IV. PISA, OECD Publishing, Paris
- Perelman S, Santin D (2011) Measuring educational efficiency at student level with parametric stochastic distance functions: an application to Spanish PISA results. *Educ Econ* 19(1):29–49
- Ramalho EA, Ramalho JJ, Henriques PD (2010) Fractional regression models for second stage DEA efficiency analyses. *J Product Anal* 34(3):239–255
- Ray SC (1991) Resource-use efficiency in public schools: a study of Connecticut data. *Manag Sci* 37(12):1620–1628
- Simar L, Wilson PW (2007) Estimation and inference in two-stage, semi-parametric models of production processes. *J Econom* 136(1):31–64
- Simar L, Wilson PW (2011) Two-stage DEA: caveat emptor. *J Product Anal* 36(2):205–218
- Wilson PW (2005) Efficiency in education production among PISA countries with emphasis on transitioning economies. Mimeo, Austin
- Worthington AC (2001) An empirical survey of frontier efficiency measurement techniques in education. *Educ Econ* 9(3):245–268
- Xue M, Harker PT (1999) Overcoming the inherent dependency of DEA efficiency scores: a bootstrap approach. Mimeo