The educational efficiency drivers in Uruguay: Findings from PISA 2009.

Santín, Daniel and Sicilia, Gabriela

Complutense University of Madrid, Spain.

March 2012

The educational efficiency drivers in Uruguay:

Findings from PISA 2009.

Daniel Santín

Gabriela Sicilia*

Complutense University of Madrid, Spain

Department of Applied Economics VI

April, 2012

Abstract

The aim of this research is to identify the main drivers of secondary school efficiency in Uruguay. We are particularly interested in identifying which variables could be influenced by the design of public policies in order to improve academic outcomes with the current resource allocation. To do this, we build a two-stage semiparametric model using PISA 2009 database. 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 censured regression (Tobit) and three different regression models based on the use of bootstrapping recently proposed in the literature. The results show an average inefficiency of 7.5% for the evaluated Uruguayan schools, suggesting that there is room for improving academic outcomes by adopting appropriate educational policies. Following on from this, the findings of the second stage demonstrate that increasing educational resources, such as reducing class size, has no significant effects on efficiency. In contrast, educational policies should focus on reviewing grade-retention policies, teaching-learning techniques, assessment systems and, most importantly, encouraging students to spend more time reading after school in order to reduce inefficiencies.

Key words: Educational production, efficiency, data envelopment analysis, bootstrap, PISA

JEL Classification: C61, D61, I2.

---

* Corresponding author.
E-mail addresses:
Daniel Santín: dsantin@ccee.ucm.es
Gabriela Sicilia: gabriels@ucm.es

1. Introduction

The interest in improving school performance and educational attainment through efficiency gains is growing basically in response to two main findings. First, improved academic outcomes have been proven to have a positive impact on economic growth [Barro and Lee (1993); 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[1]. In fact, the level of educational expenditure and its percentage share of GDP are indicators commonly used to measure a country's educational investment.

Public expenditure on education accounted for 3.53% of Uruguay's GDP in 2000, whereas ten years later it had risen to 4.5%[2]. But this significant budgetary effort has not been accompanied by adequate reforms and public policies leading to better educational achievement, as evidenced by the latest results published in the PISA 2009 (Programme for International Student Assessment) Report from the OECD (Organisation for Economic Co-operation and Development)[3].

Again, although the average PISA 2009 results confirm that Uruguay holds a foremost position within the Latin American region, the education system has entered into stagnation and recession in recent years, particularly at secondary education level which has recorded high repetition, dropout and low attainment rates. Thus, the main concern of educational policy makers in Uruguay is not only to expand the education system coverage, but also to improve the quality of teaching and academic outputs. To do this, it is clearly necessary to explore and address the sources of educational inefficiencies and improve the management of growing educational resources.

---

[1] In 2011, 84.5% of secondary students attended public schools in Uruguay (*Education Observatory, National Administration of Public Education (*ANEP*)).*

[2] The GDP grew by 37% in real terms over this period (*Uruguayan Central Bank* (BCU)).

[3] Results remain steady across the three waves in which Uruguay has participated (428, 423 and 426 average points in mathematics in 2003, 2006 and 2009, respectively).

In many cases, the discussion focuses exclusively on increasing public resources expended on education, but there is no empirical evidence to show that a higher level of resources leads *per se* to better results (Hanushek, 2003). To analyse how to boost educational performance rates, there are several international programs[4] (including PISA), which provide rich databases containing an array of information about students and schools. Using these international databases researchers can perform more specific analyses of the main drivers of academic performance[5] and also identify the sources of inefficient behaviors in the production process using student and school contextual variables (Grosskopf *et. al.* (1997); Afonso and St. Aubyn (2005); Cordero et al. (2005); De Jorge and Santín (2010), Cordero et al. (2011))[6]. Such studies are important because the presence of inefficiencies in an educational system would imply that results can be improved without increasing the current levels of resources, which is one of governments' main targets.

Two-stage models popularized by Ray (1991) and McCarty and Yaisawarng (1993) are among the best-known models for explaining the sources of inefficiency[7]. Under this approach, in a first stage we use a 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 student and school contextual variables. Two-stage models differ primarily as to the regression model specified in the second stage to explain efficiency rates. The most commonly applied methodology is the censored regression model (the so-called Tobit regression), followed by ordinary least squares (OLS) and truncated regression.

---

[4] These programs include TIMSS (Trends in International Mathematics and Science Study), IALS (International Assessment of Literacy Survey) and PIRLS (Progress in International Reading Literacy Study).

[5] See Cordero *et al.* (2011a) for a review of methodological options and research analyzing the relationship between educational resources and school performance using PISA data for the Spanish case.

[6] See Worthington (2001) and Mancebón and Muñiz (2003) for a detailed review of educational efficiency studies.

[7] See Simar and Wilson (2007) for a detailed review of two-stage models.

Xue and Harker (1999) were the first to point out the main drawback of the two-stage approach. They underline that two-stage model results are bound to be biased due to the fact that the radial efficiency scores estimated in the first stage (the dependent variable in the second stage) depend on each other. Hence, conventional inference methods are invalid in this context because the error term is serially correlated, and this violates the basic econometric assumption of independence within the sample. The cause is that the computation of the DEA efficiency score for one DMU in the first stage necessarily involves all other DMUs in the observation set. To overcome this drawback, Simar and Wilson (2007) proposed a new estimation methodology for the second stage based on the use of bootstrapping.

Finally, note that even though there are several international educational efficiency studies for the OECD countries, there is scant research in the Latin American context. To the best of our knowledge, there are no studies using this efficiency approach for the Uruguayan case in particular. In Uruguay, interest has traditionally focused on education system coverage rates, the system's redistributive effect and its impact on poverty and growth instead of analyzing the quality of the services provided and the resulting 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 inefficiency in Uruguayan secondary schools in order to provide clear evidence for the discussion of which are likely to be the best educational practices and policies to strengthen the education system and, thus, improve academic outcomes. For this purpose, we apply the semiparametric two-stage DEA approach to PISA 2009 data. The second aim of the paper is to use different regression model specifications in order to check the robustness of the results and to explore whether differences between them really matter. To do this, we explain the DEA estimated efficiency scores using four approaches available in the literature: the conventional Tobit regression and three other specifications (truncated regression, OLS and Tobit regression) based on the bootstrap procedure proposed by Simar and Wilson (2007).

The paper is organized as follows. Section 2 presents the main methodological concepts, introducing the DEA model and the alternative models implemented in the second stage. Section 3 briefly describes the Uruguayan education system and the PISA program, explaining the variables selected in both stages. Section 4 reports the results. Finally, Section 5 discusses the conclusions of this research and their implications for educational policy.

2. Methodology

2.1. The educational production function

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

$$A_s = f(B_s, S_s), \qquad \textit{Equation 1}$$

where subindex s refers to school, and A_s represents the educational output vector for school s, normally measured through the average student score on standardized tests. On the other hand, educational inputs are divided into B_s, which denotes the average student family background, and S_s, 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. All these factors may affect student performance significantly. In this case, we estimate a production frontier where fully efficient schools would belong to the frontier. These relatively efficient units achieve the maximum observed result given their resources allocation. 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:

$$A_s = f(B_s, S_s,) - u_i, \qquad \textit{Equation 2}$$

where u_s denotes the school efficiency level. Null values of u_s imply that the analyzed schools are fully efficient, meaning that given the initial input endowment and the existing technology, these schools are maximizing and correctly managing the resulting outputs. Positive u_s values would indicate that the school is inefficient, and therefore the inefficiency rate indicates by how much the output could be increased up to 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_s), educational inputs (B_s, S_s), and the estimated efficiency level (u_s) 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. This approach uses a DEA model in the first stage which measures the 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 estimation of efficiency is associated with Farrel's concept of technical efficiency (Farrel, 1957); who 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 available 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 non-parametric methods based on mathematical optimization models. Although the use of the parametric approaches has increased in the last decades[8], nonparameteric methods have been the most extensively applied for measuring educational technical efficiency.

Since the pioneering work by Charnes, Cooper and Rhodes (1981) and Banker, Charnes and Cooper (1984)[9], the DEA[10] model has been widely used to measure efficiency in several areas of public expenditure. The main reason for its widespread application is its flexibility, and the fact that it accounts for multiple outputs, the uncertainty about the true production technology and the lack of price information, making it 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-BBC model) for each analyzed unit is:

$$\begin{aligned} & Max_{\lambda,\theta_i}\theta_i \\ s.t.\ \ & \theta y_i \leq Y\lambda \\ & x_i \geq X\lambda \\ & n1'\lambda = 1 \\ & \lambda \geq 0\ ;\quad i = 1, \ldots\ldots, N \end{aligned} \qquad \textit{Equation 3}$$

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 λ 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, θ_i being the radial distance between the *i*th unit and the frontier. In other words, θ_i indicates the equiproportional expansion over outputs needed to reach the frontier. Therefore, the higher the score value θ_i , the greater the inefficiency level.

---

[8] See, for example, Perelman and Santín (2011) and Crespo-Cebada et al. (2013).

[9] The DEA-CCR model and DEA-BBC model, respectively.

[10] See Worthington (2001, p. 253f) for a detailed review of available research that measures efficiency in education through frontier techniques and mostly DEA models.

2.3. Second Stage: Explaining educational efficiency

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:

$$\hat{\theta}_i, = f(Z_i, \beta_i). \qquad \text{Equation 4}$$

The most used estimation method in this second stage is the censored regression model (Tobit), followed by ordinary least squares (OLS)[11], from which the main explanatory factors of the efficiency scores can be drawn[12]:

$$\hat{\theta}_i, = Z_i\hat{\beta}_i + \varepsilon_i \,. \qquad \textit{Equation 5}$$

Xue and Harker (1999) argued 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 perform in this second stage in order to provide consistent estimates. According to Simar and Wilson (2007) (hereinafter referred to as SW2007), 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 (Equation 5) would be biased. Additionally, the possible correlation of the contextual variables Z_i with the error term ε_i in Equation 5 is another source of bias.

SW2007 state the need for bootstrapping to overcome these drawbacks. In their paper, SW2007 propose two algorithms[13] which incorporate the bootstrap procedure in a truncated regression model. They run a Monte Carlo experiment to examine and compare the performace of these two algorithms, and they prove that both bootstrap

[11] Some authors actually estimate both models simultaneously to verify results robustness.

[12] For a detailed review of estimation methods used in the second stage of semiparametric models, see Simar and Wilson (2007).

[13] 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.

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[14]. Since the sample analyzed in our research is made up of 132 schools, we apply the simple *Algorithm #1* proposed by SW2007, which is described below.

Algorithm #1

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

1) Estimate $\hat{\beta}$ *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 [15] times by a loop, obtaining the bootstap estimations of $A = \left\{ \left(\hat{\beta}^*, \hat{\sigma}_\varepsilon{}^*\right)_b \right\}_{b=1}^{L}$.

3) Compute the bootstrap estimations:

   a) For each $i = 1, \dots, m$, extract ε_i from a normal distribution $N(0, \hat{\sigma}_\varepsilon^2)$ left-truncated in $\left(1 - z_i\hat{\beta}\right)$.

   b) Again, for $i = 1, \dots, m$, estimate $\theta_i^* = z_i\hat{\beta} + \varepsilon_i$.

   c) Using maximum likelihood, estimate the truncated regression of θ_i^* on z_i, obtaining $\hat{\beta}^*$ *and* $\hat{\sigma}_\varepsilon{}^*$ estimations.

4) Use L values of A to build the confidence intervals of β and σ_ε

Later, Hoff (2007), McDonald (2009) 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) shows that only the Tobit produces consistent results. Meanwhile, Banker and Natarajan (2008) (BN2008) provide a statistical model which yields consistent second-stage OLS estimations. Simar and Wilson (2011) again took part in the ongoing debate and compared the consistency between truncated

[14] For a more detailed analysis of the results, see Simar and Wilson (2007, p. 45f.).

[15] According to Simar and Wilson (2007), we define $L = 2000$ in this paper.

regressions and the BN2008 OLS specification. They conclude that only the truncated regression model proposed by SW2007 and, under very particular and unusual assumptions, the OLS model presented by BN2008 provide consistent estimates. Further, they prove that in both cases only bootstrap methods were capable of statistical inference.

From the above, we conclude that there is as yet no agreement in the published literature about which is (are) the most consistent regression model(s). For this reason, two-stage model practitioners find the selection of the second-stage regression model baffling, as they are unsure about whether or not results will vary significantly with their choice of specification. To clarify this point, 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, we estimate three regression models applying the bootstrap procedure: *Algorithm #1* proposed by SW2007 based on a truncated regression; and a Tobit regression and an OLS model with bootstrapping. The aim here is to explore the real implications of this methodological discussion for policy recommendations derived from an empirical analysis of real educational data.

3. Data and variables

3.1. Brief description of the Uruguayan education system

The Uruguayan national education system is composed of four levels: three years of infant education (3-5 years old), six years of primary education (6-11 years old), six years of secondary education (12-17 years old), and tertiary education as of age 18 years. Secondary education is divided into three years of basic secondary education (*Ciclo Básico Cómun*) and three years of upper secondary education (*Bachillerato).*

Compulsory education covers 14 years: the two last years of infant education (4-5 years old), primary and secondary education[16].

In terms of public and private education coverage, the public sector takes absolute primacy over the private sector. In 2011, 84.5% of high school students attended public schools[17]. This highlights how important the performance of public institutions is for national academic results, and therefore the need 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 last decade has not been accompanied by effective reforms and policies that achieve improved educational outcomes.

The results of PISA 2009 corroborate that Uruguay is still in an advantageous position within the region[18], but also confirm that results have not improved compared to previous waves. Test scores in the three analyzed areas are more highly dispersed than in other countries, which mirror the high social polarization of the education system. Comparing student performance by socioeconomic background, it is noteworthy that while almost 70% of students in very unfavorable circumstances do not reach the "competence threshold" defined by the OECD in reading[19], this figure drops to 7.7% for students from a very favorable background. By contrast, analyzing the percentage of top-scoring students (performance levels 4-6) defined by PISA analysts, we find that this proportion rises to almost 40% of students in very favorable circumstances whereas students from unfavorable backgrounds account for less than 2%.

---

[16] Art. 10 of the *General Education Law* N.18.437 of December 12, 2008.

[17] *Education Observatory, National Administration of Public Education (*ANEP*).*

[18] Uruguay is the Latin American country with the best results for science and is second placed in mathematics and reading (after Chile).

[19] For more details, see PISA 2009. Technical report (OECD, 2011).

3.2 PISA database

PISA 2009 is the fourth edition of an initiative promoted by the OECD as of the late 1990s assessing 15-year-old students. The evaluation addresses three knowledge areas: reading, mathematical and scientific literacy. 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 they have mastered a specific school curriculum.

In addition to academic achievement data, the PISA database contains a vast amount of information about students, their households and the schools they attend. Additionally, the database provides information through synthetic indexes, elaborated by OECD experts, by clustering responses to related questions provided by students and school authorities. The advantage of working with these indexes is that they have been constructed considering both theoretical considerations and empirical studies and have therefore been extensively tested at international level (OECD 2011).

The 2009 PISA cycle is the third wave in which Uruguay has taken part, and assessed 5,927 students from 232 public and private schools. For the purposes of this research, this database was refined. 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) were omitted, since one of the main DEA requirements is that the assessed units should be homogeneous. To sum up, this study analyzes 132 secondary schools (73.5% public and 26.5% private).

3.2. Outputs, inputs and contextual variables

Outputs

It is very 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, 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*)[20].

Inputs

Regarding educational inputs, three variables were selected taking into account the educational production function in Equation 1, which represent the classical inputs required to carry out the learning process (raw material, physical and human capital):

- ESCS (economic, social and cultural status): is an index developed by the PISA analysts to indicate the student socioeconomic status. It therefore represents the "raw material" to be transformed through the learning process[21]. It is the result of running a categorical principal component analysis with three variables: the highest occupational status of either parent (HISEI), the highest educational level of either parent measured in years of education (PARED), and finally an index of home possessions (HOMEPOS) [22].
- SCMATEDU (school educational resources): is an index of the quality of educational resources in the school. It is therefore associated with the physical capital. It is elaborated from the responses by principals to seven questions related to the scarcity or lack of laboratory equipment, institutional materials, computers, internet, educational software, library materials and, finally, audiovisual resources. The higher the index, the better the quality of the school's material resources.
- PROPCERT (proportion of fully certified teachers): this index reflects the quality of teachers, and therefore the school's human capital. The index is constructed by

[20] 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.

[21] Both the ESCS index and the clustered variables are standardized with mean to zero and standard deviation equal to one across equally weighted OECD countries.

[22] For further details, see PISA 2009. Technical Report, OECD (2011).

dividing the total number of certified teachers (with a teaching degree)[23] 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.

To ensure a correct DEA model specification, it is necessary to verify the monotonicity assumption, that is, all selected inputs must have a significant positive correlation with all outputs. Table 1 presents the bivariate correlations of the selected outputs and inputs where all correlations are positive and statistically significant.

Table 1. Bivariate correlations between outputs and inputs.

	ESCS		SCAMATEDU		PROPCERT	
Maths_mean	0.826	**	0.362	**	0.348	**
Read_mean	0.842	**	0.335	**	0.353	**

$^{**}p < 0.01$. Source: Own elaboration based on PISA 2009 data.

Contextual variables

Regarding the explanatory variables (Z vector in Equations 4 and 5) of the efficiency scores considered in the second stage, we select thirteen variables associated with students and schools. These variables reflect some key aspects of management and school organization and the teaching-learning processes conducted in the classroom:

- *Ownership*: this variable is a major focus of the educational debate in several countries, and there is still no consensus about its influence on educational efficiency. In the case of Uruguay, there is no previous evidence to contrast with the results of this research, whereas several studies have highlighted the limited influence of school ownership on educational efficiency in other countries like Spain[24]. The reference category in the estimated models is the private school.

[23] Certified teachers in Uruguay are required to complete a four-year degree at the *Instituto de Profesores Artigas* (IPA), a higher education institution which provides specialized secondary teacher training.
[24] See Perelman and Santín (2011) and Cordero *et al.* (2011).

- *Metasum:* an index developed by PISA analysts from student responses regarding the usefulness of five different strategies for writing a summary of a long and difficult text. These strategies have been ranked by PISA experts on a scale where higher index values imply that students select better strategies for summarizing texts, that is, the index could be considered as an approximation of the students' synthesis ability, and would therefore be associated with their reading results. This variable is expected to have a positive impact on efficiency.

- *Right_year:* the 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.

- *Extra_Reading:* the percentage of students in the school who spend between one and two hours per day reading for pleasure after school. It is understood that reading contributes to the student's 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.

- *Teach_stu*: 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[25].

- *Tcshort:* an index developed by PISA analysts that reflects how short the school is of of qualified teachers[26]. It is based on the responses provided by the school principal regarding the shortage of teachers of mathematics, science, reading and other subjects. The higher the rate, the greater the shortage of teachers. The *a priori* expected relationship is that the greater the shortage, the lower efficiency.

- *Test:* a dummy variable that takes the value one when students are assessed by teachers through tests, quizzes or exams more often than monthly.

---

[25] For a more detailed review, see Hanushek (2003) and Hoxby (2000).

[26] According to OECD, qualified teachers have a higher education degree.

- *Homework*: a dummy variable which refers to the assesment 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 possitive effect on efficiency.

Finally, we incorporate a number of variables associated with school autonomy in terms of budget allocation, curriculum development, disciplinary policies and student assessment practices. Unlike the above variables, there is no expected *a priori* positive or negative relationships between these variables and efficiency in this case, since empirical evidence emerging from international comparisons does not provide conclusions that are applicable to all education systems (OCDE 2010, *Vol IV*).

- *Budget_director*: dummy variable which takes the value one when the school principal is mainly responsible for distributing the school budget.

- *Budget_author:* dummy variable which takes the value one when the national authorities are mainly responsible for distributing the school budget.

- *Curr_board*: dummy variable which takes the value one when the school board is the ultimately responsible for determining the content of the courses.

- *Disc_board:* dummy variable which takes the value one when the school board is ultimately responsible for developing student disciplinary policies.

- *Asses_auth*: dummy variable which takes the value one when the national authorities are the mainly responsible for setting student assessment policies.

3.3. Descriptive Analysis

Table 2 presents the main descriptive statistics of outputs, inputs and contextual variables. Results are presented for all schools and by school ownership. Statistics evidence a notable heterogeneity in the results, especially when comparing public and private schools. It appears from Table 2 that private schools perform better in both tests, possess greater educational resources and have higher average levels of the explanatory factors of efficiency. For most variables, public schools are more heterogeneous than private institutions.

Finally, there are some noteworthy differences. First, the average results for both tests in private schools are approximately ninety points (two standard deviations) greater than the scores observed in public schools. Additionally, the socioeconomic index for private institutions almost doubles the figures for public schools. This highlights the unequal initial endowment received by each school type. Thirdly, whereas 88% of 15-year-old students are in the right year at private schools, this applies to only 56% at public schools. Finally, private schools are more autonomous than public institutions in three out of the four selected dimensions (assessment policies, budget allocation and disciplinary policies).

Table 2. Descriptive statistics of outputs, inputs and explanatory variables of efficiency.

		All schools				Private schools				Public schools			
Variable	**Description**	**Min**	**Max**	**Mean**	**S.D.**	**Min**	**Max**	**Mean**	**S.D.**	**Min**	**Max**	**Mean**	**S.D.**
Outputs													
Maths_mean	Schools average score in mathematics	268.9	554.3	424.1	58.8	393.8	554.3	489.0	40.4	268.9	507.2	400.7	45.3
Read_mean	School average score in reading	266.9	560.4	421.5	63.5	369.4	560.4	494.6	39.4	266.9	505.7	395.1	48.0
Inputs													
ESCS	Economic, Social and Cultural Status index	1.90	5.60	3.35	0.88	3.53	5.60	4.57	0.54	1.90	4.25	2.91	0.47
SCAMATEDU	School Educational Resources index	0.001	5.32	3.51	1.04	2.42	5.32	4.23	0.86	0.00	5.32	3.25	0.98
PROPCERT	Proportion of fully certified teachers	0.001	1	0.57	0.18	0	1	0.65	0.17	0	1	0.54	0.17
Explanatory variables													
Ownership	School ownership	0	1	0.73	0.44	0	0	0.00	0.00	1	1	1.00	0.00
METASUM	Student's use of synthesis strategies index	1.07	2.58	1.80	0.36	1.50	2.58	2.16	0.30	1.07	2.31	1.67	0.29
Test	Student's assessment through test with frequency more than one per month	0	1	0.14	0.34	0	1	0.11	0.32	0	1	0.14	0.35
Homework	Student's assessment through monthly homework	0	1	0.17	0.37	0	1	0.20	0.41	0	1	0.15	0.36
Right_year	Percentage of students in the appropriate year	0	1	0.65	0.26	1	1	0.88	0.11	0	1	0.56	0.25
Extra_reading	Percentage of students who spend between 1 and 2 hours a day reading for pleasure after school	0	0.25	0.10	0.06	0	0.25	0.12	0.07	0	0.23	0.09	0.06
TEACH_stu	Number of teachers per 100 students	1.93	50.00	8.07	4.86	1.93	18.85	8.37	3.54	2.75	50.00	7.96	5.26
TCSHORT	Shortage of qualified teachers index	0.86	4.34	2.04	0.93	0.86	2.76	1.46	0.59	0.86	4.34	2.25	0.94
Curr_board	The School Board is the last responsible for determining the content of the courses	0	1	0.03	0.17	0	1	0.06	0.24	0	1	0.02	0.14
Asses_auth	National Authorities are the mainly responsible for setting the students assessment policies	0	1	0.69	0.46	0	1	0.46	0.51	0	1	0.77	0.42
Budget_director	The school head is mainly responsible for distributing the school budget	0	1	0.48	0.50	0	1	0.31	0.47	0	1	0.54	0.50
Disc_board	The School Board is the last responsible for disciple policies for the students	0	1	0.30	0.46	0	1	0.57	0.50	0	1	0.21	0.41
Budget_auth	National Authorities are mainly responsible for distributing the school budget	0	1	0.41	0.49	0	0	0.00	0.00	0	1	0.56	0.50

For dummies variables, the mean represents the proportion of schools in that category.
Source: Own elaboration based on PISA 2009 data.

4. Results

Figure 1 illustrates the distribution of efficiency scores, δ_i, estimated by the output-oriented DEA-BCC model. Results show that 17% of the schools behave efficiently. On average, educational results could be increased by 7.5% given the available resources. Moreover, nearly one in ten schools could improve their results by over 20% to reach the frontier and 15% could improve outcomes by 10% to 20% with their current inputs.

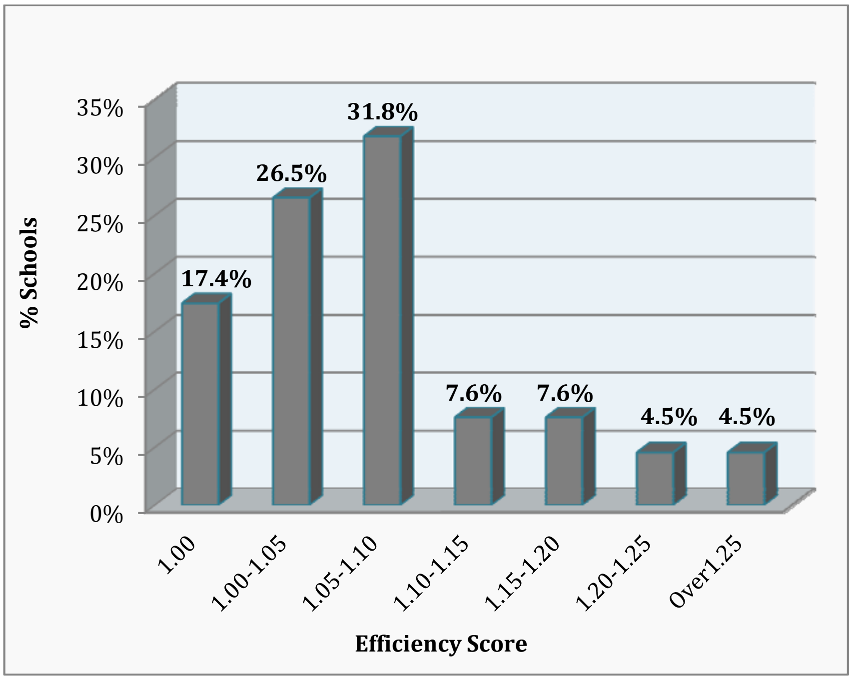

Figure 1. Efficiency scores distribution (DEA-BBC) based on PISA 2009 data. N=132. Mean estimated efficiency score 1.075 (sd. 0.075).

The estimated efficiency scores are regressed over the contextual variables. Results are shown in Table 3. The first conclusion from the comparative analysis of the four specified models is that there are no major discrepancies between the results. The largest appreciable differences are between results from the conventional Tobit model without bootstrapping and the three models that employ *bootstrap* procedures. Particularly, if we observe the effect of the controversial teacher-student ratio variable ("*Teach_stu*"), we could arrive at inappropriate policy recommendations if we do not perform bootstrapping to tackle the potential bias. This highlights the importance of applying bootstrapping on top of the selected specification during the second stage. This finding is consistent with the evidence reported by Simar and Wilson (2007, 2011). Taking into account this general conclusion we will from now on discuss the results of the three estimated models (censured, truncated and OLS) with bootstrapping only. Particularly, we will consider the specification proposed by Simar and Wilson (2007) as a reference.

Firstly, there is a set of variables that do not affect efficiency scores. The first is school ownership suggesting that private schools achieve better results because of a higher initial endowment of inputs and a more conducive learning environment. In addition, hardly any of the variables associated with school autonomy are significant in any of models (with the exception of "*Curr_board*"). Decentralizing the budgetary allocation decision or the design of disciplinary policies and assessment practices does not affect school efficiency. These variables are closely related to school ownership in the Uruguayan education system, and this is therefore consistent with the non-significance of ownership discussed above. On the other hand, there are two variables whose significance does differ slightly between Tobit and OLS with bootstrapping and the SW2007 model: *Test* and *Tcshort*. This would be the only case where there would be no conclusive evidence about whether the increase of qualified teachers in schools or the evaluation of students by tests set more often than once a month may lead to improvements in efficiency[28].

By contrast, there is a group of variables associated with students and teaching practices that are significant and with the expected sign in all specifications, corroborating the robustness of the estimations. First, the *METASUM* index, which reflects the students' synthesis ability, has a positive impact on efficiency. This ability could be associated with classroom teaching techniques adopted by teachers and is thus a factor to be considered by school managers, especially in the early stages of the learning process when students are assimilating the learning techniques to be used throughout their academic life.

Second, extracurricular reading is the variable that most affects efficiency. Thus, schools with the highest proportion of students who read between one and two hours per day after school prove to be more efficient. This variable could be interpreted as a result of student motivation received at home and their attitude towards learning, but also as a consequence of motivation from school. If this were the case, it would be apposite to take actions to encourage students to read after school in order to boost educational efficiency. However, these results should be interpreted

[28] Since the real data generator process is unknown, it is not possible *a priori* to say which model is the best. Running several Monte Carlo simulations in order to analyze which second-stage specification yields the best result would be a very interesting line of future research.

with caution, since the time spent on reading should not replace the time students spend on their homework; it should be a complementary not a substitute activity.

Thirdly, the percentage of students that are in the right year appears to be a positive and significant driver of efficiency. This result calls into question the adequacy of current Uruguayan grade retention policies at all levels of the education system. Therefore, it would perhaps be better to attempt to identify younger (primary) students who are at risk of repeating and provide them with additional support in order to prevent their repetition.

Finally, student assessment methods and their frequency appear to positively influence efficiency. Indeed, schools where teachers set their students homework on a monthly basis perform better than schools that do not make use of this tool or do so with a different frequency. Finally, the school board being mainly responsible for determining the *curricula* has a significant negative effect on efficiency. In the Uruguayan education system, national authorities are mainly responsible for designing academic programs, irrespective of school ownership and educational level. Therefore, the result of this research would suggest that decentralizing this responsibility would not be an appropriate policy, at least in the case of Uruguay.

5. **Concluding remarks**

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. Although PISA 2009 results place Uruguay in a prominent position within Latin America, it is still far removed from the average for the OECD countries. In turn, educational policies have over the last decades focused on allocating more resources to the educational system, but academic outcomes have not significantly improved. This can be associated with the presence of inefficiencies in the education system, implying that current results could be improved without further increasing the present resource allocation. Quantifying and identifying the main sources of these inefficiencies would appear to be crucial for assessing which educational policies to promote and therefore ensure efficient resource allocation. This is the main focus of the present research.

Table 3. Efficiency drivers: second-stage estimates.

	Model I: Conventional *Tobit*				**Model II: *Algorithm* #1 (*Trunc*)**				**Model III: *Algorithm* #1 (*Tobit*)**				**Model IV: *Algorithm* #1 (*OLS*)**			
Variables	**Coef.**	**Std. Err.**	**t**	**P>t**	**Coef.**	**Std. Err.**	**z**	**P>z**	**Coef.**	**Std. Err.**	**z**	**P>z**	**Coef.**	**Std. Err.**	**z**	**P>z**
Ownership	0.000	0.0237	-0.01	0.993	0.003	0.0388	0.07	0.942	0.000	0.0259	-0.01	0.994	0.003	0.0203	0.17	0.866
METASUM	-0.054 **	0.0252	-2.13	0.035	-0.080 **	0.0387	-2.06	0.039	-0.054 **	0.0280	-1.92	0.050	-0.046 **	0.0227	-2.04	0.042
Test	-0.031 *	0.0165	-1.86	0.065	-0.041	0.0313	-1.31	0.191	-0.031 *	0.0185	-1.66	0.096	-0.029 **	0.0147	-1.96	0.050
Homework	-0.041 ***	0.0157	-2.65	0.009	-0.050 *	0.0296	-1.70	0.089	-0.041 **	0.0169	-2.45	0.014	-0.034 ***	0.0131	-2.62	0.009
Right_year	-0.089 **	0.0347	-2.56	0.012	-0.178 ***	0.0566	-3.14	0.002	-0.089 **	0.0383	-2.32	0.020	-0.087 ***	0.0320	-2.73	0.006
Extra_reading	-0.275 ***	0.1025	-2.68	0.008	-0.280 *	0.1585	-1.76	0.078	-0.275 **	0.1102	-2.49	0.013	-0.229 **	0.0949	-2.42	0.016
TEACH_stu	-0.002 *	0.0011	-1.83	0.070	-0.003	0.0032	-0.85	0.396	-0.002	0.0018	-1.14	0.256	-0.002	0.0012	-1.43	0.153
TCSHORT	-0.018 **	0.0086	-2.14	0.034	-0.015	0.0127	-1.21	0.225	-0.018 **	0.0091	-2.02	0.043	-0.016 **	0.0075	-2.19	0.029
Curr_board	0.085 ***	0.0281	3.02	0.003	0.092 **	0.0460	1.99	0.046	0.085 **	0.0351	2.42	0.016	0.085 **	0.0335	2.54	0.011
Disc_board	0.028 *	0.0149	1.86	0.066	0.013	0.0235	0.56	0.576	0.028 *	0.0162	1.71	0.088	0.022 *	0.0133	1.65	0.099
Budget_director	-0.021 *	0.0113	-1.87	0.064	-0.028	0.0200	-1.41	0.158	-0.021	0.0129	-1.63	0.103	-0.018 *	0.0103	-1.71	0.087
Budget_auth	-0.016	0.0153	-1.07	0.285	-0.038	0.0242	-1.59	0.113	-0.016	0.0164	-1.00	0.315	-0.016	0.0136	-1.17	0.241
Asses_auth	-0.009	0.0165	-0.52	0.602	-0.035	0.0236	-1.49	0.135	-0.009	0.0181	-0.48	0.633	-0.013	0.0148	-0.86	0.389
Constant	1.327	0.0586	22.64	0.000	1.451	0.0929	15.62	0.000	1.327	0.0645	20.57	0.000	1.307	0.0506	25.84	0.000
$\hat{\sigma}_u$	0.0662	0.0050			0.0682	0.0064			0.0662	0.0050			0.0606			

***p< 0.01 ; **p< 0.05 ; *p< 0.10

Source: Own estimations based on PISA 2009 data.

Our findings corroborate the presence of inefficiencies in the secondary education production by the evaluated schools, which with the current inputs could increase their results by 7.5% if adequate education policies were designed by national authorities and implemented by schools. In this respect, the second-stage analysis yields interesting responses for planning and implementing effective policies in the Uruguayan education system. First, increasing educational resources *per se*, like reducing class size, appears to be an inappropriate policy because it has an insignificant effect on school efficiency. Neither is school ownership significant, whereby we can conclude that the better mean results achieved by private schools are a consequence of higher initial resources endowment (student socioeconomic status, school educational resources and better trained teachers) and therefore to a better learning environment. By contrast, this research suggests that the debate and action in order to improve education system efficiency should focus on reviewing repetition policies, teaching techniques and assessment systems rather than just increasing educational resources. Promoting teaching techniques that enhance students' synthesis ability, providing support for students with a higher risk of repetition at an early age and assessing students by setting homework continously throughout the academic year are some of the recommended practices that would improve Uruguayan school efficiency. Furthermore, encouraging students to engage in extracurricular activities that complement the classroom learning process, such as spending some more time reading after school, would lead to considerable improvements in academic results with the currently available resources.

Finally, note that the conclusions drawn by the SW2007 truncated regression model and the BN2008 OLS model (with bootstrap) show similar results at a 95% significance level. The sign, magnitude and significance of almost all variables are the same in both models, implying that the educational policy recommendations derived from them would be basically the same, adding robustness to the findings discussed above. From this result we can conclude too that practitioners should be more concerned about performing bootstrapping in the second stage than about the final choice of the second-stage regression model.

References

Afonso, A. and St. Aubyn, M. (2005): "Cross-country efficiency of secondary education provision. A semi-parametric analysis with nondiscretionary inputs". Working Paper Series No. 494. *European Central Bank*.

Barro, R. and Lee, J.W.; (1993), "International Comparisons of Educational Attainment". *Journal of Monetary Economics*, 32(3), 363-94

Banker, R. D.; Charnes, A. and Cooper, W. W. (1984): "Models for estimating technical and scale efficiencies in data envelopment analysis". *Management Science,* 30 (9), 1078-92.

Banker, R. D. and Natarajan, R. (2008) Evaluating contextual variables affecting productivity using data envelopment analysis. *Operations Research,* 56(1)**,** 48-58.

Calero, J. and Escardíbul, J. O. (2007). "Evaluación de servicios educativos: el rendimiento en los centros públicos y privados medido en PISA 2003". *Hacienda Pública Española / Revista de Economía Pública*, 183 (4), 33-66.

Charnes, A.; Cooper, W. W. and Rhodes, E. (1981): "Evaluating program and managerial efficiency: an application of data envelopment analysis to program follow through". *Management Science*, 27, 668-697.

Charnes, A.; Cooper, W.W.; Lewin, A.Y. and Seiford, L.M. (1994): "Data Envelopment Analysis: Theory, Methodology and Applications". *Kluwer Academic Publishers*, New York.

Coleman, J.S.; Campbell, E.Q.; Hobson, C.J.; McPartland, J.; Mood, A.M.; Weinfeld, F.D. and York, R.L. (1966). Equality of Educational Opportunity. *Washington, D.C.: U.S. Government Printing Office.*

Cordero, J. M., Pedraja, F. and Salinas, J. (2005). "Eficiencia en educación secundaria e inputs no controlables: sensibilidad de los resultados ante modelos alternativos". *Hacienda Pública Española*, 173 (2), 61-83.

Cordero, J.M.; Crespo, E.; Pedraja, F and Santín, D. (2011): "Exploring Educational Efficiency Divergences Across Spanish Regions in PISA 2006", *Revista de Economía Aplicada* 57, (vol. XIX), 117-145.

Crespo-Cebada, E., Pedraja-Chaparro and Santín, D. (2013). "Does school ownership matter? An unbiased efficiency comparison for regions of Spain". *Journal of Productivity Analysis.* DOI 10.1007/s11123-013-0338-y

De Jorge, J and Santín, D. (2010): “Los determinantes de la eficiencia educativa en la Unión Europea”. *Hacienda Pública Española*, 193, 131-156.

Duncombe, W. (1997): “Why is it so hard to help central city schools? *Journal of Policy Analysis and Management,* 16(1), 85-113.

Farrel, M.J. (1957): “The measurement of productive efficiency”. *Journal of the Royal Statistical Society*, 120, 253-266.

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 Electrónica Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, 7 (4), 164-179.

Grosskopf, S., Hayes, K.; Taylor, L. y Weber, W. (1997): “Budget-constrained frontier measures of fiscal equality and efficiency in schooling”. *Review of Economics and Statistics*, 79 (1), 116-124.

Hanushek, E.A. (1979): 'Conceptual and empirical issues in the estimation of educational production functions'. *Journal of Human Resources*, 14 (3), 351-88.

Hanushek, E. A. (1986): “The economics of schooling”. *Journal of Economic Literature*, 24 (3), 1141-1171.

Hanushek, E.A. and Kimko, D.D. (2000): “Schooling, labor-force quality, and the growth of nations”, *The American Economic Review*, 90 (5), 1184-1208.

Hanushek, E.A. (2003): “The failure of input based schooling policies”, *The Economic Journal*, 113, 64-98.

Hoff, A. (2007): “Second stage DEA: Comparison of approaches for modeling the DEA score”. *European Journal of Operation Research*, 181, 425-435.

Levin, H. M. (1974): “Measuring efficiency in educational production”. *Public Finance Quarterly*. 2, 3-24.

Llambí, C. and 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”. Working paper, *Centro de Investigaciones Económicas*.

Llambí, C.; Perera, M. and Messina, P. (2009): “Desigualdad de oportunidades y el rol del sistema educativo en los logros de los jóvenes uruguayos”. Working paper, *Centro de Investigaciones Económicas*.

Mancebón, Mª.J. and Muñiz, M. (2003): “Aspectos clave de la evaluación de la eficiencia productiva en la educación secundaria”, *Papeles de Economía Española*, 95, 162-187.

McDonald, J. (2009): "Using least squares and tobit in second stage DEA efficiency analyses". *European Journal of Operational Research*, 197, 792-798.

OECD (2010): "PISA 2009 Results: What Makes a schools successful? Resources, policies and practices". (Volume IV).

OECD (2011): PISA 2009. Technical Report.

Perelman, S. and Santín, D. (2011): "Measuring educational efficiency at student level with parametric stochastic distance functions: an application to Spanish PISA results". *Education Economics*, 19 (1), 29-49.

Ramalho, E. A., Ramalho, J. J. and Henriques, P. D. (2010) Fractional regression models for second stage dea efficiency analyses. *Journal of Productivity Analysis,* 34(3)**,** 239-255.

Simar, L. and Wilson, P.W. (2007): "Estimation and Inference in Two-Stage, Semiparametric Models of Production Processes", *Journal of Econometrics*, 136, 31-64.

Simar, L. and Wilson, P. W. (2011) Two-stage DEA: Caveat emptor. *Journal of Productivity Analysis,* 36(2)**,** 205-218.

UNESCO (2009): "Desafíos de la educación uruguaya. Aportes para un diagnostico del sistema educativo".

Worthington, A. C. (2001): "An Empirical Survey of Frontier Efficiency Measurement techniques in Education". *Education Economics*, 9 (3), 245-268.

Xue, M., and Harker, P. T. (1999): "Overcoming the inherent dependency of DEA efficiency scores: a bootstrap approach". Working Paper, *Wharton Financial Institutions Center, University of Pennsylvania*.