Latin American Economic Review (2024), volume 32, article 12

RESEARCH ARTICLE

The productivity of Colombian public schools and its determinants from a value-added perspective, 2014–2019: An empirical analysis

Tommaso Agasisti * Juliana Arias † Gloria Bernal ‡ Luz Karime Abadía §

Submission received: January 23, 2023
Final version received: August 25, 2023
Accepted: August 30, 2023
Published: April 7, 2024

Abstract

Identifying how efficient public schools are – that is, the academic performance they can generate with the available inputs – is relevant information for designing public policies to improve education quality. Using a two-stage data envelopment analysis (DEA) and the Malmquist index, this study measures the efficiency of public schools in Colombia and their productivity evolution, from a value-added perspective. The sample includes 3,854 public schools and standardized test records for students in 9th and 11th grade for the period 2014–2019. The findings reveal that on average schools could have increased outputs by 18.5% with the inputs available to them. The highest level of inefficiency is observed in 2014, and the lowest in 2017. The increase in productivity can be associated with two policy initiatives: the nationwide scholarship *Ser Pilo Paga* (which encouraged high-achieving low-income students to devote more effort to the exit examinations) and the Synthetic Index of Educational Quality (which provided rewards to schools that showed academic progress).

Keywords: efficiency, productivity, public school, Colombia, value-added, data envelopment analysis (DEA), Malmquist index.
JEL codes: H21, H52, I21

*Politecnico di Milano School of Management, Milano, Italy. Via Raffaele Lambruschini, 4. Post code: 20156. E-mail address: tommaso.agasisti@polimi.it

†Instituto Tecnológico Metropolitano de Medellín, Colombia. Calle 54a #30-01. Post code: 050012. E-mail adress: julianaarias@itm.edu.co

‡Corresponding author. Pontificia Universidad Javeriana, Bogotá, Colombia. Cra. 7 #40 – 62. Post code: 11001000. E-mail address: gbernal@javeriana.edu.co. Phone: +57 3112323661

§Pontificia Universidad Javeriana, Bogotá, Colombia. Cra. 7 #40 – 62. Post code: 11001000. E-mail address: labadia@javeriana.edu.co

1. Introduction

Studying the efficiency of public schools is of paramount importance in the present times. Public schools are funded with public budgets, and this implies accountability, that is, the demonstration that the resources are being used productively (Hanushek, 1986). In this vein, schools must provide evidence that they are pursuing and achieving their goals (educating the students) at the highest possible level, given the available resources. Doing so is even more necessary in current circumstances, in which tight public budgets must address a variety of pressing concerns, such as health care or social welfare (Rubin, 2019). Moreover, studies across countries demonstrate that student achievement in some countries is lower than in others despite similar levels of funding, so the former countries are requested to find more efficient ways to improve their performance with the given resources (Woessmann, 2016).[1]

These concerns about educational efficiency are of striking relevance for the case of Colombia. While education spending as a proportion of GDP is similar to that of countries like Chile and Poland, the latest data reveal significantly worse academic achievement than in the other two. For example, in 2014 government expenditure in Colombia on education as a percentage of GDP was 4.7%, comparable to that in Chile (4.8%) and in Poland (4.9%) (The World Bank, 2020). However, according to OECD (2019), Colombia ranked 57^{th} out of 77 participants countries in PISA reading score, with a lag of 40 points compared to Chile (the best-performing country in Latin America) and 100 points behind Poland, which ranked 10^{th}. The underperformance of Colombian 15-year-old students is equivalent to over one year of schooling behind Chile and almost two years behind Poland. This evidence suggests a lack of efficiency in the Colombian educational sector.

Within such a context, this research addresses four main research questions related to public schools in Colombia. First, what has the Colombian public schools' efficiency level been in recent years, from a value-added perspective? Second, how has the productivity of Colombian public schools evolved over time in recent years? Third, what are the main factors associated with the efficiency of Colombian schools? Fourth, to what extent are the efficiency scores heterogeneous across Colombian departments? To answer these questions, this paper employs a two-stage data envelopment analysis (DEA) and the Malmquist index in considering a sample of 3,854 public schools (distributed across all departments in Colombia) for the period 2014–2019. The output is measured by the scores on the high school exit examination (Saber11). The inputs include several variables of school resources and students´ prior achievement (standardized test scores from 9^{th} grade -Saber9).

This paper contributes to the literature about the efficiency of schools in Colombia. This is the first study in Latin American to include in the empirical analysis prior (school-average) test scores as inputs, thus modelling the efficiency of schools in a VA perspective using a large administrative data set. By controlling for students' prior academic performance, it is possible to isolate the school's contribution to student outcomes from the impact of pre-existing student abilities. This is essential because students with higher prior academic performance may naturally perform better in subsequent assessments, regardless of the school's quality or effectiveness in fostering students' skills. On the other hand, students with lower prior performance might face additional challenges in catching up, even if they attend an excellent school. Therefore, the value-added perspective used in this study allows us to obtain robust measurements of school efficiency scores. Second, the determinants of efficiency are explicitly considered in order to identify specific school-level factors that can be actioned by decision-makers. Third, the productivity change of public schools from 2015 to 2019 is decomposed between school-specific modifications of efficiency and system-level technological change, thus enabling understanding of the evolution of educational performance over time – in line with other studies in the field since Hanushek (1986).

[1] See, for the case of developing countries, Agasisti and Zoido (2019)

The present paper finds that, on average, the inefficiency level of Colombian schools is 18.5% with important variations across regions. We find evidence of a "productivity peak" between 2016 and 2018 due to a rise in technical efficiency. In addition, we find that schools, where the academic performance and socioeconomic status of students are more homogenous, with a lower proportion of female students, higher proportion of teacher with postgraduate studies, lower dropout rates, lower proportion of female teachers, and located in urban areas, are positively associated with school efficiency.

The positive changes in productivity in these years coincided with the introduction of two educational national policy initiatives that might have influenced the intrinsic motivation of the students to perform better on the exit examination exams and extrinsic incentives for schools to support students in this direction. On the one hand, the introduction of the Ser Pilo Paga (SPP) full scholarship for college resulted in the improved performance of high-achieving low-income students (mostly placed in public schools), who sought to meet the scholarship score requirements (Rattini, 2014). That is, the possibility of obtaining full financing for attending a high-quality college (even the most expensive private universities) increased the motivation of low-income students to do well on the exit examination. On the other hand, the creation of the *Índice Sintético de la Calidad Educativa* (ISCE; Synthetic Index of Educational Quality) might have incentivized schools to perform better. This initiative introduced rewards for public schools that met specific goals related to progress, performance, efficiency, school environment, and the use of infrastructure. Both the SPP scholarship and the index were discontinued beginning in the third quarter of 2018, when a new government came to power. The period of the introduction and dismantling of both programs matches the increase and decrease in the efficiency frontier identified in this paper.

The remainder of the paper is organized as follows: Section 2 describes the theoretical framework. Section 3 reports the main findings from the existing literature about the efficiency of public schools in Colombia and other Latin American countries. Section 4 describes the context and data sets and provides some descriptive statistics. The methodological approach is described in section 5. Section 6 presents the main results obtained from the empirical analysis. Section 7 concludes and sets forth some key policy implications.

2. Theoretical framework

Microeconomic theory uses the concept of the production function to describe the maximum output that an organization can feasibly produce with a set of inputs. This concept is relevant to the empirical analysis of the quality and efficiency of education, specifically the identification of their determinants, in developing and developed countries (Chakraborty et al., 2001).

In this vein, the educational process can be described by an educational production function (EPF), such that schools are understood as organizations that use inputs and transform them into outputs. In previous empirical analyses, the output has generally been measured by the academic performance of students and inputs are all observed variables that can influence the learning environment, such as characteristics of students and their families, school resources, characteristics of teachers, and institutional aspects (Hanushek, 1979; Deutsch et al., 2013; De Witte and López-Torres, 2017). Efficiency can be defined as the ability of schools to maximize the outcome (students' academic performance) with the available inputs (financial and human resources), that is, to achieve its production frontier. In the present paper, the EPF tool enables the measurement of the technical efficiency of the productive process of a school over time. In formal terms,

$$Y_{jt} = f(Y_{j(t-1)}, x_{jt1}, x_{jt2}, x_{jt3}, ..., z_{jt1}, z_{jt2}, z_{jt3}, ..., h_{jt1}, h_{jt2}, h_{jt3}, ...) \quad (1)$$

where Y_{jt} is the academic performance of school j at time t. Moreover, $Y_{j(t-1)}$ is the prior achievement of school j at time t and *x, z*, and h are additional inputs, for example, characteristics of the students and school, at the school level over time.

It is important to note here that $Y_{j(t-1)}$ is included in order to account for the progress that pupils make that is not associated with school practices alone, in line with the value-added (VA) literature (Doran and Izumi, 2004). Controlling for students' previous academic performance provides a more robust estimation of efficiency and its determinants than simply measuring the level of academic achievement at a given point in time. Lastly, including students' prior achievement – averaged at the school level – in the production function enables the efficiency analysis to account for the portion of the variation in the students' performance that cannot be attributed to their individual characteristics.

3. Related literature

The most efficient use of resources in education has been a source of discussion in the literature of the economics of education. This literature has for the most part relied on data from developed Western countries, such as the USA, Italy, Australia, and the UK (De Witte and López-Torres, 2017) in drawing its conclusions. Studies of the efficiency of schools in Latin American countries are limited but growing; some existing studies explore the comparison of schools' efficiency across different countries in the area. For example, Deutsch et al. (2013) estimate the educational production functions for Brazil, Chile, Colombia, Mexico, and Uruguay with OECD-PISA 2006 data. Using a Shapley decomposition assessing the educational production process, the authors estimate the impact of each input on the schools' individual efficiency. They find that individual efficiency probably depends on the strength of the intergenerational link, and for countries such as Mexico and Colombia, female pupils have lower efficiency, signaling a gender gap or discrimination. Agasisti and Zoido (2019) use a much larger sample in analyzing the efficiency of more than 6,800 schools in 28 developing countries, including 7 from Latin America, using a two-stage data envelopment analysis (DEA) and data from OECD PISA 2012. These authors find that achievement scores can be raised by 30% on average through more effective use of available resources and that high heterogeneity is present both between and within countries. Of the factors associated with schools' efficiency, the most important are the characteristics of the student population, such as the average socioeconomic background. Despite their interesting findings, both papers limit the time frame of their data to a single year; a recent strand of studies on school efficiency in Latin America incorporates analysis over time, thus overcoming the possible accumulative impact of inputs that might be neglected in a cross-sectional analysis. Arias Ciro and Torres García (2018) and Salazar Cuéllar (2014) investigate the efficiency of public secondary education expenditure and are similar in that they employ DEA methodology and data from the PISA test, although they consider different time windows. The former finds for the 2012–2015 period that developing countries could improve their PISA test scores using the same level of current resources, whereas the latter finds for the period between 2000 and 2009 that holding expenditure constant, secondary schools could increase their students' PISA scores by about 10%. In line with cross-sectional studies, both of these recent contributions highlight that schools should be able to get closer to the efficiency frontier.

In Latin America, Brazil and Colombia have seen the most studies assessing school efficiency through frontier (mostly DEA) methods. For the former, Queiroz et al. (2020)) evaluate the efficiency of primary education in Brazilian schools using a dynamic DEA model in light of school differences in terms of the socioeconomic levels of students. The paper reports almost no progress in school efficiency between 2007 and 2015 but does find evidence of possible efficiency improvements as a result of investments in school infrastructure. Cardoso et al. (2021) analyze the technical efficiency of municipal educational systems in cities in the state of Rio Grande do Sul, Brazil, by using DEA and identifying benchmark cities in terms of efficiency. Finally, Bernardo et al. (2021) use the stochastic frontier analysis method to measure the efficiency of Brazilian municipalities in relation to the application of public resources in education. Their findings indicate that higher levels of education among the population contribute to efficient school management because better-educated citizens realize the importance of social control. In Colombia, school efficiency studies have been focused on comparisons of private and public schools. Notably, close to 79% of all students in the country are enrolled in the latter; these students have disproportionately fewer resources and score significantly lower on academic assessments such as the PISA and national exit examination exams. De Jorge-Moreno et al. (2018) estimate that public schools are 4.3% more inefficient than private. However, Iregui et al. (2006) and Arbona et al. (2021) do not find remarkable differences in efficiency between public and private schools.

In terms of focus and methodology, Arbona et al. (2021) is the closest to the present study, but they do not consider prior academic performance as a control. Using a metafrontier and the Malmquist-Luenberger index, along with exit examinations tests (Saber11) as outputs, the authors measure changes in the productivity of 4,587 schools in the Colombian education system between 2014 and 2017, with an emphasis on differences across administrative regions. Two outcomes are investigated in the study, performance and inequality. The general results indicate a deterioration of efficiency in both public and private schools, due to the change in best practices and the change in efficiency. Another important finding is the large gap in the efficiency scores of schools operating in different administrative regions.

This paper adds to the existing evidence about the efficiency of schools in the Latin American region. This is the first to investigate the efficiency of Colombian schools over a relatively long span (the six years between 2014 and 2019). The main contribution of this paper to the literature, therefore, is the modelling of the efficiency of schools from a value-added (VA) perspective. As mentioned before, this paper includes 9th-grade test scores as an input, with the output being 11th-grade test scores. Adding prior test scores as inputs yields better output measures for efficiency analysis than does relying on levels measures of performance (Gronberg et al., 2012). Gronberg et al., 2012). Not including the prior attainment of students, due to the usual data limitations, has been argued to be a major shortcoming of most of the efficiency of education studies (De Witte and López-Torres, 2017). In the educational efficiency literature, only a few studies have included VA in school efficiency estimations. Portela and Camanho (2010) and Portela et al. (2013), who consider test scores upon entry into and exit from secondary education in Portugal; and Gort et al. (2019),), who incorporate school VA in student learning outcomes as a measure of the effectiveness of Australian schools. To the best of our knowledge, there are no studies that make use of the VA perspective in estimating the efficiency of Latin American schools, so the present paper fills this gap in the literature and contributes to a still narrow strand of studies which adopt this methodology.

4. Context, data, and descriptive statistics

4.1. Context – The Colombian educational system at glance

In Colombia, elementary education spans five grades and secondary education six. The exit examination, Saber11, is given in grade 11; in a typical year, nearly half million students across the country take it, with the compliance rate being approximately 97%. Except for those in expensive private schools, on average, students do not achieve satisfactory levels of performance in this examination. For example, in 2019 the Global score on the Saber11 of students in public schools was on average 25 points lower than that of their private-school counterparts. Education in Colombia is provided by public and private schools, and the system is very segregated. While private schools have flexibility in their management, public schools are strictly regulated by the government. In addition, public schools are entirely subsidized by the government, enroll 79% of the country's students in primary and secondary education, and have (as stated above) most of the low-income pupils.

Within the studied period, two salient national education-related programs aimed at improving educational outcomes – specifically, Saber11 performance – were initiated. The first, the *Ser Pilo Paga* grant program (which began in October 2014), offered to pay the college costs for 10,000 high-achieving low-income high school students who scored above a certain level on the exit examination every year.

The first cohort of awarded students started college in 2015 (Decree 1075). Close to 300,000 students were eligible based on their family's income, but only the 10,000 highest achievers could benefit from it. The scholarship, which was widely publicized, made it possible for students to enroll in highly ranked private universities previously inaccessible to those students due to the prohibitive cost. The program lasted for four cohorts, ending in September 2018. Although students in the program could finish their programs, new students were given other credit possibilities that were not as generous as the scholarship. The second, the ISCE, was introduced in March 2016 (Decree 501).[2] This tool monitors the results of the educational process of public schools in four domains (progress, performance, efficiency, and school environment) and the use of infrastructure. Importantly, schools that meet their annual goal of excellence in the ISCE, which is set by a formula, receive an in-kind reward – as a further extrinsic economic incentive. The ISCE was discontinued in the third semester of 2018. Both policies represented incentives in the period 2015–2018 for students and schools to perform better than before.

4.2. Data and descriptive statistics

Three sources of information are used in this study. First, the data set contains the standardized databases Saber9 and Saber11, administered by the Colombian Institute for the Evaluation of Education (Instituto Colombiano para la Evaluación de la Educación, ICFES), as well as the C600 school database from the Ministry of Education. Saber9 is a standardized exam taken by 9th-grade students that evaluates competence in the areas of science, mathematics, language, and citizenship. The exam was given annually between 2012 and 2017 to all schools. Within each school, a representative sample of 9th-grade students takes the Saber9 exam (the only year for which Saber9 was censual is 2017). Only competency in mathematics and language was tested in the study period. Unlike Saber9, Saber11, the Colombian high school exit examination, is mandatory. Furthermore, taking this exam is a requirement for enrolling in higher education. Saber11 evaluates competency in five areas: mathematics, language, natural science, social science, and citizenship. Due to comparability purposes between the two exams, this analysis can only use information about the mathematics and language scores, those on Saber9 for 2012 to 2017 and on Saber11 for 2014 to 2019. For example, students who were in 9th grade in 2012 must have been in 11th grade in 2014[3]. Using information from no later than 2019 enabled the isolation of the impact of the COVID-19 pandemic on academic performance (Abadía and Bernal, 2017). In addition, due to methodological changes, the Saber11 test scores are not comparable before and after the year 2014. Therefore, we use the information from Saber11 since 2014[4]. Thus, the school-average Saber9 is the input for prior academic performance (included among the other school inputs), whereas Saber11 is the

[2] In Decree 501, it was also recommended that public schools that offered two shifts (morning and afternoon, 4 hours each) switch to a long single shift instead (approx. 8 hours per day), without reducing the number of attending students. This policy is not included in our analysis for two reasons: first, it was not mandatory, and second, there were significant limitations in schools' infrastructure that restricted schools' ability to make this change (Ikoya and Onoyase, 2008).

[3] Saber9 and Saber11 do not have unique identifiers at the student level that would make it possible to track them over time. Furthermore, while Saber11 is nearly comprehensive, the Saber9 database is a representative sample of students each year, except for 2017 when it was comprehensive. As a result, we cannot guarantee that the corresponding cohort from Saber9 is exactly what we are observing in Saber11. Additionally, students who repeat a grade since 9th grade and those who change schools can also affect the feasibility of tracking the same students in both databases. On average, during the analyzed period, 1.2% of students in upper secondary repeated a scholar grade. Overall, the limited dimension of the phenomenon is unlikely to affect the estimation of the schools' efficiency in a substantial way.

[4] The Saber11 test was designed with two main goals i) to estimate the performance of the students and ii) to make comparisons across time to monitor the learning quality (Instituto Colombiano para la Evaluación de la Educación (ICFES)., 2019a). According to Decree 1075 of 2015, the structure of the Saber11 exam must be maintained for a minimum of 12 years to ensure the comparability of the results over time. ICFES uses the methodologies of Balanced Incomplete Block Designs (BIBs) and Item Response Theory (IRT) with equating to ensure comparability on Saber9 and Saber11 tests. These methodologies guarantee the comparability of test measurements among students who take the Saber11 in a specific year and across student cohorts over time. These methodologies guarantee the comparability of test measurements among students who take the Saber11 in a specific year and across student cohorts over time. For a detailed explanation of the Sabe11 test methodology see https://www.icfes.gov.co/documents/39286/2231027/Edicion+3+-+boletin+saber+al+detalle+.pdf/9e086de0-eeff-bf05-dcd8-5738cae9969e?version=1.3&t=1678150141307

output. Beyond scores, the Saber9 and Saber11 databases also contain a rich set of characteristics of the pupils, their families, and their schools.

The C600 database contains administrative data at the school level, such as the share of female teachers, the number of teachers and their level of education, etc. The C600 survey is administrated annually by the Departamento Administrativo Nacional de Estadística (DANE, the national statistics agency). The data set used in this paper merges all three sources (i.e., it was collapsed at the mean the Saber9 and Saber11 individual variables) in order to build a pool of data at the school level.

The sample is restricted to public schools, given that this is where public resources are allocated and as noted above most children in primary and secondary education are enrolled. We decided to perform the study only with public schools due to the Colombian educational system is very segregated. Private and public schools are so different in the students´ family composition (mainly in the socioeconomic status of students and therefore in their prior academic performance), the curriculums implemented, the way they are administrated, and the availability of resources they have. Private schools are autonomous in how they can administrate their resources and spend their budget whereas public ones are not. Therefore, very different drivers could explain their efficiency levels and possible changes in productivity.

The final sample includes 3,854 schools, 68.9% of total secondary public schools (see Table A.1 for more detail). These are the schools for which complete information for the period of study was available across all the variables used.[5] For those schools for which no merging was possible, observable characteristics were compared with those in the sample. It turns out that missing schools were more likely to be in rural areas and be attended by students who scored lower on academic tests and come from a lower socioeconomic family background. In this vein, the results of the present study must be interpreted accordingly – the results hold strong internal validity, but their application to more-disadvantaged schools is not straightforward.

The inputs used in the estimations of schools' efficiency scores were teacher/student ratio, computer/student ratio, socioeconomic status-SES (school-average) and the results on previous tests (*Saber9_math* and *Saber9_language*, respectively), the latter conceived as a proxy for students' abilities. The measures of outputs were math and language scores in 11th grade (*Saber11_math* and *Saber11_language*, respectively). *SES* is an index computed by ICFES using the Item Response Theory that summarizes three dimensions of students´ characteristics: educational attainment of parents, occupation, and family income. It ranges between 0 and 100, where the higher the index, the better the living conditions the individual has[6]. The inclusion of this input dimension allows to compare schools' performance net of the effects due to the socioeconomic composition of students.

Table 1 shows that, Saber 11 and Saber 9 scores as well as the socioeconomic status index (SES) in general decreased between 2014 and 2016, however, these indicators improved in 2017. The teacher/student and computer/student ratios improved during the analyzed period. The SES indicator was on average 45.71, which implies a low socioeconomic status but it had a peak of improvement in 2017. In 2014 and 2015, mathematics scores were higher than language scores, but from that point on the reverse was the case. At the beginning of the period under study the average mathematics and language score for Colombian schools was 48.67 and 48.32 (on a 0–100 scale), respectively; as of 2019, the former had only increased by 2.2% and the latter by 4.2%. The ratio of teachers per student was stable throughout the period of study, with around 20 students per teacher and 0.04 computers per student.

[5]For most of the database merging, we use School DANE ID, by "sede." Those schools that could not be merged by ID, probably due in some cases to migration to new IDs numbers, were merged manually using the names and location.

[6]For a detailed explanation of how ICFES computes the SES index, see Instituto Colombiano para la Evaluación de la Educación (ICFES). (2019b)

Table 1: *Descriptive statistics of schools by years*

		Outputs		Inputs				
Year	Variable	(1) Math test	(2) Language test	(3) Teacher / Student	(4) Computer / Student	(5) SES	(6) Math test t-2* (9°)	6) Language test t-2* (9°)
2014	Mean	48.67	48.32	0.04	0.13	45.01	287.23	293.14
	sd.	3.93	4.28	0.01	0.32	5.92	40.3	39.66
2015	Mean	48.58	48.38	0.04	0.2	44.92	285.7	284.17
	sd.	5.14	3.82	0.01	0.17	6.04	38.05	39.01
2016	Mean	49.49	51.34	0.05	0.25	44.98	284.87	284.31
	sd.	5.49	4.16	0.01	0.21	5.62	41.25	39.29
2017	Mean	48.78	51.93	0.05	0.28	46.69	283.93	279.6
	sd	5.64	4.32	0.01	0.23	4.74	39.2	39.04
2018	Mean	49.17	51.25	0.05	0.29	46.15	302.8	297.11
	sd	5.65	4.42	0.01	0.24	5.23	31.63	30.63
2019	Mean	49.73	50.99	0.05	0.29	46.52	297.28	303.54
	sd.	5.67	4.66	0.01	0.24	5.24	27.95	26.47
Total	Mean	49.07	50.37	0.05	0.24	45.71	290.3	290.31
	sd	5.31	4.52	0.01	0.24	5.53	37.41	37.01
	Min	27.57	33.64	0.01	0	21.24	167	157
	Max	78.97	69.69	0.25	17.25	66.74	500	500
	Obs	23124	23124	23124	23124	23124	23124	23124

*t-2 cohort taking a test in 9th grade (i.e. 2 years before the output test)
Source: Own elaboration.

In addition to inputs and outputs, a list of nondiscretionary variables was used with the aim of exploring factors that can be statistically associated to different levels of school efficiency. Incorporating non-discretionary variables—factors explaining inefficiency scores—guards against misinterpreting efficiency (for example as a result of inefficient management). In the analysis, the variables selected for this purpose can be classified in five groups: student's characteristics, school's characteristics, teacher's characteristics, regional dummies, and time dummies. The student´s characteristics include the standard deviation of student´s SES as a proxy for homogeneity of student body (*sd_SES*) and the standard deviation of mathematics and language scores as proxies for homogeneity of student´s academic level (*sd_math11* and *sd_lang11*, respectively). The school´s characteristics include the number of students who had dropped out during the previous year (*dropouts*), the proportion of female students (*female_stud*), and the track of the school (e.g., academic, technical, or teaching-vocational[7]). With regard to the last characteristic, two dummies for academic and teacher training schools were constructed that were compared against schools with a technical orientation. In addition, the data set considers the school's location, i.e., urban vs. rural. Finally, the list of conditional variables includes the size of the school, measured by number of students enrolled (*size*). Teacher´s characteristics include the proportion of teachers with postgraduate degrees (*tc_postgrade*) and the number of female teachers (*tc_female*).). Finally, we construct dummies for each time period (*year2014, year2015, year2016, year2017, year2018, year2019*) and each geographical department of Colombia, with the exception of Guaviare and Vaupés, because the schools in both departments do not have complete data for the study period. Including year and regional fixed effects allows for the control of time trends and unobserved characteristics of regions that do not

[7]In Colombia, high schools have one of three types of orientation, depending on the track schools want their students to take in the last two years of secondary education: academic, technical, and teacher training. In teacher training schools, students get the high school diploma at the same time they are being educated as teachers of preschool or basic primary school.

change over time (i.e., regional fixed-effects capture the consistent public spending on education at the department level).

The descriptive statistics for the previously mentioned nondiscretionary variables are presented in Table A.2. On average, the heterogeneity in language test scores in 11^{th} grade is lower than in mathematics test scores. There has been an increase in the heterogeneity of students in terms of their socioeconomic status and a decrease in the number of dropouts. On average, 82% of schools have an academic orientation, schools have around 1,200 students, 93% of teachers have completed postgraduate studies, and 33% of teachers and 45% of students are female in each school.

5. Methodological approach

The methodological approach used in this paper consists of two steps. In the first step, the Colombian public schools' efficiency level is estimated through DEA. The measurement of education efficiency assumes that education is a production process in which schools transform inputs (such as students, teachers, and school resources) into outputs (such as test scores). DEA methodology consists of estimating an education "frontier" based on the inputs and outputs used by each decision-making unit (DMU; in the case of this paper, a school). This frontier is constructed nonparametrically based on observed data following an optimization program without any specification for the production function's functional form and as a reference point, each school is classified as efficient or inefficient, given their relative distance from the estimated efficiency frontier (Cooper et al., 2002). The main advantages of this methodology are that (1) it can be applied using multiple inputs and outputs, (2) it is not necessary to specify a previous relationship between inputs and outputs, and (3) it is not required to assume any distribution about the statistical error. The main critical aspect is the selection of the inputs and outputs (Agasisti and Wolszczak-Derlacz, 2016); however previous literature and the available data have guided the choices in this paper (De Witte and López-Torres, 2017).

Mathematically, the efficiency of each j-th school in the sample is estimated based on the following ratio:

$$Eff_i = \frac{\sum_{n=1}^{N} y_{nj}\alpha_n}{\sum_{m=1}^{M} x_{mj}\beta_m} \tag{2}$$

where y_{nj} is the vector of N outputs, x_{mj} is the vector of M inputs, and α_n and β_m are the weights of each output and input. The DEA technique consists on maximizing this ratio, defining the weights that make Eff_i the highest possible (Agasisti and Zoido, 2019). A school is efficient when it is on the frontier. It is inefficient when $Eff_i < 1$ and the distance from 1 measure the level of inefficiency.

We follow the two-stage procedure of Simar and Wilson (2007) to estimate efficiency scores, which enables the capture of the relationship with nondiscretionary variables throughout the estimation of a truncated regression. Additionally, using bootstrap estimations enables the correcting for serial correlation issues due to the correlation between the estimated efficiency scores in the first stage with the non-discretionary variables. These estimated efficiency scores are called "robust" scores, because they are bias corrected using bootstrapping. Specifically, we use the second algorithm method with bootstrapping following subsequent steps with $L1 = 1,000$ and $L2 = 2,000$ replications (see Simar and Wilson (2007), pp. 42-43 for details).

In summary, after estimating the DEA efficiency frontier for all schools, the analysis estimates a semiparametric model for the education production process that includes schools', teachers', and students' characteristics as external factors that affect educational performance Z_i. The estimated regression is as follows:

$$\hat{\delta}_i = \Psi(Z_i\beta) + \epsilon_i \geq 1 \tag{3}$$

where the dependent variable is the bias-corrected score $\hat{\delta}$, Ψ is a smooth function, β is the vector of parameters, and ϵ_i is a truncated normal random variable with $N(0, \sigma_\epsilon^2)$ distribution and left truncation at $(1 - Z_i\beta)$. Here, 2,000 replications are used to estimate the marginal effect of environmental (nondiscretionary) variables Z_i.

In the second step of the methodology, a dynamic approach, in this case the Malmquist index (MI), is employed to analyze the evolution of the productivity of Colombian public schools from 2014 to 2019 (Equation 4). The index enables the evaluation of changes in total factor productivity (TFP) of each school between two points of time. Any change will be due to two complementary effects: the "catching" effect (a change in technical efficiency, ϵ_i, for example when schools approach the efficiency frontier) or a "frontier shift," as the result of a technological change, τ_i. MI is the product of the changes in technical efficiency and technological change, and it can be mathematically expressed as follows:

$$M_{i,(t1,t2)} = \underbrace{\frac{D_i^{t2}(x_i^{t2}, y_i^{t2})}{D_i^{t1}(x_i^{t1}, y_i^{t1})}}_{\epsilon_i} \underbrace{\left[\frac{D_i^{t1}(x_i^{t1}, y_i^{t1})}{D_i^{t2}(x_i^{t1}, y_i^{t1})} \cdot \frac{D_i^{t1}(x_i^{t2}, y_i^{t2})}{D_i^{t2}(x_i^{t2}, y_i^{t2})}\right]^{1/2}}_{\tau_i} \tag{4}$$

where D_i is the efficiency distance function, and x are the inputs and y the outputs in periods $t1$ and $t2$. In the Equation (4), $D_i^{t1}(x_i^{t1}, y_i^{t1})$is the distance in the ith school from the period $t1$ using as reference the technology of the same period $t1$; $D_i^{t2}(x_i^{t2}, y_i^{t2})$ is the distance using $t2$ as reference; $D_i^{t1}(x_i^{t2}, y_i^{t2})$ is the distance of the school from period 1 using technology of period 2; and $D_i^{t2}(x_i^{t1}, y_i^{t1})$ is the distance from period 2 using technology of period 1. The first component, the technical efficiency rate (ϵ_i) is the ratio of the efficiency score at time t_2 to the efficiency score at time t_1. It represents the change in technical efficiency over the two time periods, the improvement or deterioration in the use of inputs to produce a given level of output. It measures the change in efficiency over time and captures the extent to which a school has moved closer to or farther away from the production frontier. A higher value of ϵ_i indicates an improvement in efficiency, while a lower value indicates a decrease in efficiency. It can be attributed to the capacity of a school to improve in terms of management, organization, and coordination. The second component is called technological change (τ_i). It is the geometric mean of the distance functions at time t_1 and t_2,repectively. This represents the shift in the production frontier over time. It captures the progress made in producing more output with the same level of inputs over time. Positive values indicate technological progress, while negative values indicate a decrease in productivity due to a decrease in the efficiency of technology utilization (Arbona et al. 2021; Oh 2010; & Pastor and Lovell 2005).

The efficiency distance function is estimated as a linear programming problem with bootstrap (Agasisti and Wolszczak-Derlacz, 2016). The values of $MI > 1$ indicate an increase in TFP between periods t_1 and t_2, which means that the distance from the school to the frontier is less in period t_2 than the period t_1; while values of $MI < 1$ indicate a decrease in the school´s productivity. If there is no change in productivity between t_1 and t_2, $MI = 1$.

DEA methodology and the Malmquist index (MI) rely in some key assumptions to ensure their accurate and reliability. In particular, MI assumes that changes in technical efficiency and technology are exogenous, meaning they are not affected by other factors. However, unobservable factors could bias the results. For this reason, we are caution in assuming causal relationship in our findings. In addition, we are assuming that the education production function has constant returns to scale and there is not substitution between inputs, in order to interpret MI as TFP change (Førsund and Ove Kalhagen 1999; Agasisti and Wolszczak-Derlacz 2016; Wolszczak-Derlacz 2018).

6. Results of the empirical analysis

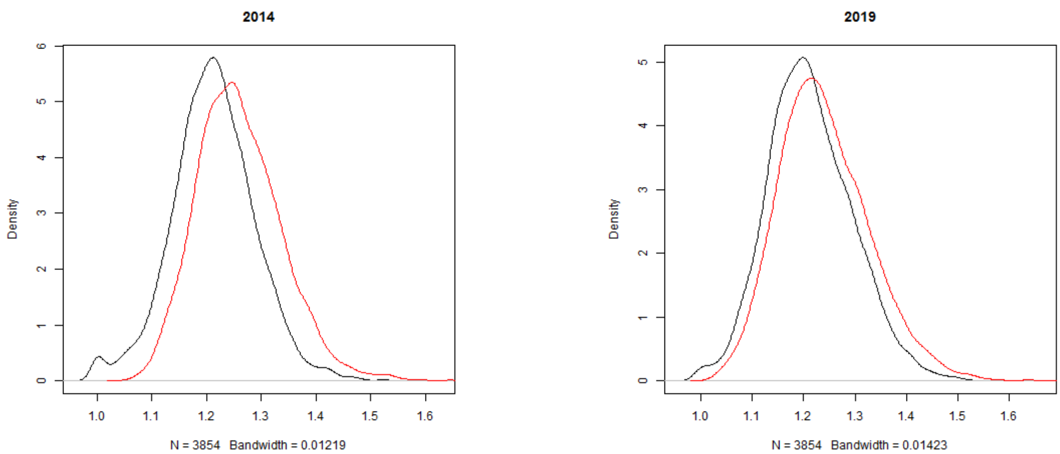
6.1. The efficiency of Colombian schools in 2014 and 2019

Table 2 shows the average levels of efficiency scores by year and the percentage of the school inefficiency. Findings reveal that the inefficiency of Colombian schools was around 15.3% (the lowest) in 2017 and 20.6% (the highest) in 2014. On average, schools could have increased their outputs (Saber11 scores in mathematics and language) by 18.5% while keeping their current inputs (teacher-student ratio, etc.) level. Although inefficiency is low compared to what other studies have found (example Agasisti and Zoido (2019)), the standard deviation is high (0.08), with some schools having a level of inefficiency of 30%. Figure 1 shows the density of schools' efficiency scores for 2014 and 2019 and Figure 2 the distribution of efficiency across the whole period. The analysis of these figures enables the description of the heterogeneity of efficiency across schools. For example, on the left tail of the distribution of efficiency scores are some schools with efficiency levels around 60% both in 2014 and 2019, even after including environmental variables in the efficiency measurement.

Table 2: *Efficiency scores between 2014-2019*

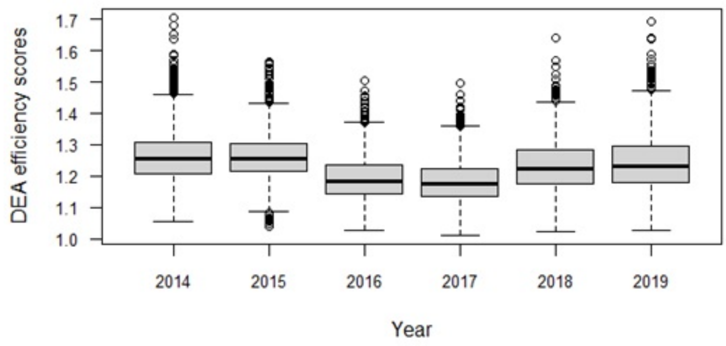
Year	Robust DEA score	School Inefficiency
2014	1.2600	20,63%
2015	1.2594	20,60%
2016	1.1913	16,06%
2017	1.1804	15,28%
2018	1.2311	18,77%
2019	1.2401	19,36%
Mean efficiency	1.2271	18,51%
Mean Inv_efficiency	0.8150	
Max	1.6998	
Min	1.0123	
Sd	0.0808	

Source: Own elaboration. The percentage of school inefficiency are computed as 1 – (1/DEA score).



Source: Own elaboration.

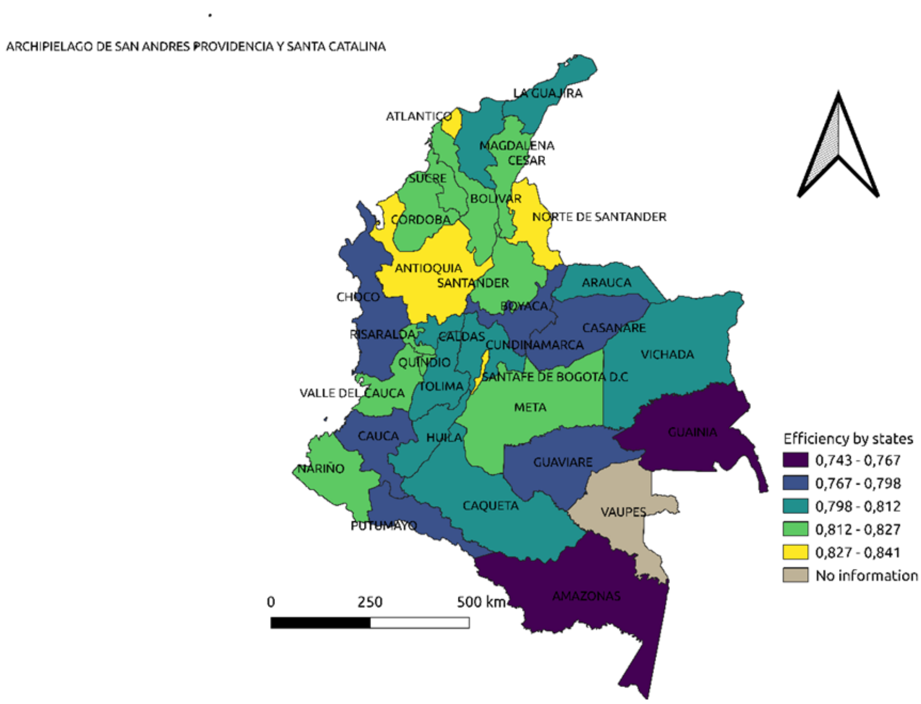
Figure 1: *Density of scores. Black: Naïve – Red: Robust*



Source: Own elaboration.

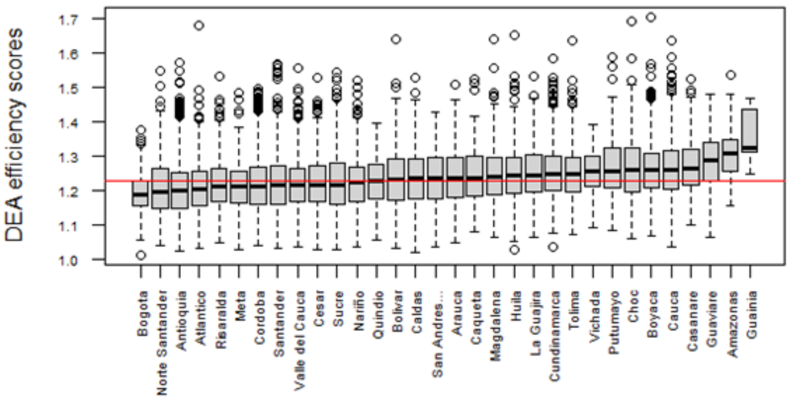
Figure 2: *DEA Efficiency scores 2014-2019*

There are also interesting differences across departments in Colombia. Figuer 3 shows that schools in Antioquia, Atlántico, and Norte de Santander are on average more efficient than in other departments (inefficiency is around 16–17%), and schools in departments such as Amazonas and Guainía present the worst performance in terms of efficiency (inefficiency is around 26%), although Guainía presents less dispersion, as the boxplots of scores show in Figure 4. However, within each department there are substantial divergences. For example, more outliers can be seen to the right of the distribution (high inefficiency values) and very few extreme values to the left (extraordinarily low efficiency scores, associated to schools only in Bogotá, Cundinamarca, and Huila). Bogotá is the state with the lowest dispersion in the score between quartiles 1 and 3. Finally, it can be seen that in most departments there is right asymmetry (a substantial number of schools with high inefficiency values).



Source: Own elaboration.

Figure 3: *Distribution of efficiency scores by departments*



Source: Own elaboration.

Figure 4: *Heterogeneity of efficiency scores by departments*

6.2. The factors associated with schools' efficiency

The second part of the schools' efficiency analysis presents the results from the second stage of DEA methodology. The robust efficiency scores of a school are those that consider some variables (out of the school's control) in the estimation as factors affecting the school's transformation (efficient or not) of its resources into the output of interest (academic scores). The following equation shows the potential determinants of previously estimated bias-corrected efficiency scores from 2014 to 2019 to be estimated (section 6.1):

$$\begin{aligned}\hat{\delta}_{i,t} = \alpha + \beta_1 sdSES + \beta_2 sdMATH11 + \beta_3 Dropouts + \beta_4 Academic + \beta_5 teachertraining \\ + \beta_6 Urban + \beta_7 Size + \beta_8 FemaleStud + \beta_9 tcPosgrade + \beta_1 0tcFemale + State \\ + \nu_t + \epsilon_{i,t}\end{aligned} \tag{5}$$

[8]

Table 3, column 1, shows the results of this second stage estimation, the aim of which is to derive schools' efficiency scores net of the influence exerted by factors that are beyond their control. Almost all the conditional variables included in the regression are significant at the 5% level. The results show that the following characteristics of schools: being more heterogeneous in terms of academic performance (i.e., with a higher standard deviation of mathematics scores), and in terms of socioeconomic status (i.e., with a higher standard deviation of SES), with a higher percentage of female students, higher dropout rates, higher proportion of female teachers and located in rural areas are negatively correlated with school efficiency. By contrast, bigger schools (with size measured by number of students) and those with a higher proportion of teachers who have completed postgraduate studies are more efficient. The academic orientation of the institution is shown to not be statistically significant, though in previous studies in different contexts these two variables have been identified as relevant to the explanation of the (in)efficiency of schools. For example, Agasisti and Zoido (2019) found, in a sample of 28 developing countries, that those schools with academic orientation are more efficient.

When looking at time dummies, schools were more efficient in 2016, 2017, 2018, and 2019 compared to 2014. Moreover, most of the regional dummies are significant. For example, comparing the

[8]See Table A.3 for the definition of the variables.

efficiency of the schools located in the 30 departments of Colombia with those operating in Bogotá, schools in Antioquia, Atlántico, Córdoba, Nariño, Norte de Santander, Santander y Sucre were, on average, more efficient than those in the capital. Schools in the remaining departments, excluding San Andres and Providencia, Bolivar, Cesar, Meta (which are not significant), are more inefficient that those in Bogotá. Once again, it is important to recall that these factors are correlated with efficiency, that is, with the ability to use resources to maximize test scores – and not with the test scores themselves. In other words, these factors are associated with the operation of a school as it supports its students in achieving their final scores, and not with the absolute level of performance.

The findings presented in this paragraph are in line with those of previous studies. In particular, Agasisti (2013) finds in the case of Italy that lower efficiency is associated to higher dropout rates and a larger proportion of female students. Deutsch et al. (2013) also find that at the individual level in Colombia the efficiency of female students is lower. Agasisti (2013) and Agasisti and Vittadini (2012) using data at the country level also conclude that rural location has a negative effect on efficiency. In contrast to our results, these authors conclude that schools with an academic orientation are more efficient than those with technical or vocational orientation. Regarding the correlation between the quality of teachers and efficiency, Agasisti (2013), using the PISA database, finds that in Italy schools with a greater proportion of qualified teachers are more efficient. This author also concludes that the size of the school is positively associated to efficiency.

An interesting result that departs from the findings of previous studies is the negative association we find between the proportion of female teachers and efficiency. Delprato and Antequera (2021) at the country level, using data from OECD PISA tests, do not find the results related to this input to be statistically significant. These results confirm that in Colombia gender is an important factor affecting school efficiency. Possibly, gender stereotypes could explain why greater proportions of female students and female teachers reduce a school´s efficiency (Abadía and Bernal 2017; Guiso et al. 2008; Gomez Soler et al. 2020).

Table 3: *Out of control school variables associated with school´s efficiency*

	Second stage DEA	
VARIABLES	Original Model	Model without Saber9
sd_inse	0.002***	0.001**
	(0.000)	(0.001)
sd_math11	0.002***	-0.006***
	(0.000)	(0.000)
dropouts	0.000***	0.000***
	(0.000)	(0.000)
academic	0.001	-0.001
	(0.001)	(0.002)
normal	-0.015***	-0.039***
	(0.004)	(0.005)
urban	-0.020***	-0.032***
	(0.001)	(0.002)
size	-0.000***	-0.000***
	(0.000)	(0.000)
tc_posgrade	-0.041***	-0.077***
	(0.005)	(0.006)
tc_female	0.001***	0.002***
	(0.000)	(0.000)
female_std	0.018***	0.074***
	(0.004)	(0.005)
Amazonas	0.039***	0.158***

	(0.011)	(0.015)
Antioquia	-0.011***	0.004
	(0.002)	(0.003)
Arauca	0.014***	0.003
	(0.005)	(0.006)
San Andres y Providencia	0.010	0.115***
	(0.012)	(0.016)
Atlantico	-0.011***	0.029***
	(0.003)	(0.004)
Bolivar	0.001	0.075***
	(0.003)	(0.004)
Boyaca	0.019***	-0.006*
	(0.003)	(0.003)
Cordoba	-0.021***	0.021***
	(0.003)	(0.003)
Caldas	0.010***	0.018***
	(0.003)	(0.004)
Caqueta	0.021***	0.017***
	(0.005)	(0.007)
Casanare	0.039***	0.028***
	(0.004)	(0.005)
Cauca	0.010***	0.035***
	(0.003)	(0.004)
Cesar	0.005	0.039***
	(0.003)	(0.004)
Choco	0.018***	0.139***
	(0.005)	(0.006)
Cundinamarca	0.027***	0.014***
	(0.002)	(0.003)
Guainia	0.062**	0.086**
	(0.026)	(0.035)
Guaviare	0.030**	0.049***
	(0.012)	(0.016)
Huila	0.008***	-0.013***
	(0.003)	(0.004)
Guajira	0.025***	0.068***
	(0.004)	(0.006)
Magdalena	0.017***	0.087***
	(0.003)	(0.004)
Meta	-0.005	-0.019***
	(0.003)	(0.004)
Nariño	-0.022***	-0.048***
	(0.003)	(0.004)
Norte de Santander	-0.012***	-0.023***
	(0.003)	(0.004)
Putumayo	0.020***	0.008
	(0.004)	(0.005)
Quindio	0.010***	0.028***
	(0.004)	(0.005)
Risaralda	0.006*	0.006
	(0.003)	(0.005)

Santander	-0.013***	-0.039***
	(0.002)	(0.003)
Sucre	-0.011***	0.030***
	(0.003)	(0.004)
Tolima	0.017***	0.032***
	(0.003)	(0.004)
Valle del Cauca	0.006**	0.025***
	(0.002)	(0.003)
year2015	0.002	0.023***
	(0.002)	(0.002)
year2016	-0.059***	-0.044***
	(0.002)	(0.002)
year2017	-0.067***	-0.036***
	(0.002)	(0.002)
year2018	-0.018***	-0.027***
	(0.002)	(0.002)
year2019	-0.008***	-0.014***
	(0.002)	(0.002)
σ	0.065***	0.086***
	(0.000)	(0.000)
Constant	1.247***	1.429***
	(0.006)	(0.007)
Observations	23124	23124

Notes: *** significant at 1%; ** significant at 5%, and * significant at 10%.
Source: Own elaboration.

6.3. Schools' efficiency and productivity changes, 2015–2019

As described in section 5, the MI is an analytical instrument for decomposing the productivity change of a school in a given period, specifically into two components: (1) change in the technical efficiency of the school (i.e., its ability to improve efficiency of operations) and (2) technological advancement, meaning the ability to take advantage of beneficial changes occurring throughout the school system. The MI of each school can then be aggregated to derive interesting insights about the productivity evolution of the entire school system.

Table 4 shows that the progress in terms of the productivity of Colombian schools in the period under analysis was mixed. There was deterioration in productivity between 2015 and 2016, then improvement from 2016 to 2018, and deterioration again from 2018 to 2019. In all of the time periods, the productivity change is not due to appreciable modifications in the schools' efficiency scores, instead being driven by a technological change. The productivity growth rates were 3.1% from 2016 to 2017 and 4.3% from 2017 to 2018. On average, the annual productivity growth rate was -0.68%, where the annual efficiency change was 1.13%, but the annual technological change was negative (-1.70%). This means that the Malmquist index remained almost constant over the period and the overall productivity was very stable. As highlighted, the managerial efficiency of schools throughout the period did not experience major changes (the level of efficiency remains stable), but the frontier moved quite substantially. It seems as though schools were subjected to a common set of circumstances that affected their productivity during 2016–2018 more than individual schools' experiencing specific changes in their efficiency.

Table 4: *Productivity changes of schools between 2014-2019*

Period	MI = Malmquist Index	EC = Efficiency Change	TC = Technological Change
2014-2015	0.9558	1.0159	0.9411
2015-2016	0.9830	1.0190	0.9647
2016-2017	1.0310	0.9965	1.0347
2017-2018	1.0431	0.9900	1.0540
2018-2019	0.9528	1.0351	0.9208
Average	0.9932	1.0113	0.9830

Notes: The Malmquist Index represents the evolution of productivity based on annual changes. If $MI > 1$ indicates an increase in the total factor productivity. This change can be due to a technical or a technological change. If $EC > 1$ indicates that efficiency increased in the period. If $TC > 1$ then a positive innovation occurred.
Source: Own elaboration.

These results can be critically read in light of national policy affecting mainly public schools. In the period between 2015 and 2018, two nationwide initiatives were implemented with the aim of improving the quality of education: the SPP program, which applied to secondary public schools, and the ISCE, which applied to primary and secondary public schools. SPP was launched in October 2014 (students registered for Saber 11 in August 2014 and SPP began distributing awards in 2015) and was terminated in September 2018. The SPP program granted full scholarships and stipends to attend the best private and public higher education institutions in Colombia to students who scored high on Saber 11 and came from a disadvantaged socioeconomic background. Bernal and Penney (2019) argue that the introduction of the scholarship incentivized these students to better prepare for the Saber 11: in their empirical analysis, the authors find that students who qualified for the scholarship scored about 0.09 test score standard deviations higher than those who did not. The ISCE (which remains in effect) was launched in 2016 by the Ministry of Education with the goal of creating incentives to boost schools' performance. The index evaluates schools' academic performance and their progress based on indicators such as scores on state tests like Saber 11, efficiency (using the student pass rate), and school environment. Schools that achieve the highest levels of progress receive in-kind and, in some cases, monetary awards in addition to public recognition. These two programs might be considered as having an important influence in the educational productivity results of the country, due to their implementation's coinciding with the increases in productivity.

Table 5 shows the proportion of schools in the sample that increased their efficiency. The lowest value corresponds to the period 2018 to 2019, which could be related to the termination of the education programs mentioned above. In contrast, between 2017 and 2018 (when both programs were in place), around 75% of schools improved their productivity. While we do not have instruments for testing the causal effect of the two programs in a robust way, the data at hand indicate a possible effect on the Colombian educational system's efficiency.

Table 5: *Percentage of schools where productivity improved*

Year	Percentage of schools with improvements in productivity
2014-2015	21.01713
2015-2016	44.08407
2016-2017	61.36482
2017-2018	75.27244
2018-2019	19.14894

Source: Own elaboration.

6.4. Robustness checks

One key aspect of DEA methodology is the choice of inputs. In our analysis, we based the choice of inputs based on previous empirical literature and data availability. One of our main contributions is the inclusion of the previous student´s academic performance as an input in the estimation of schools´ efficiency scores. Therefore, as a robustness checks for the efficiency of schools presented in the previous section, an alternative model were estimated in which we remove each one of the four inputs considered, in order to check possible changes in the efficiency score and the relative importance of the inputs respect to baseline findings. In the first model, the teacher-student ratio variable was dropped from the set of inputs; in the second, the computer-student ratio variable was dropped; in the third model the prior academic achievement measured in the 9th grade (Saber 9) variable was dropped, and in the last model the socioeconomic status. The efficiency scores from original model and alternative models 1, 2, and 4 are similar and the correlation is significant and strong, ranging from 0.97 to 0.98 (Table 6), suggesting that the results are robust across these specifications of the efficiency analysis. However, the correlation between the original model and the model without the previous academic performance of students is the lowest (0.78), showing the usefulness of measuring the ability of schools to employ their resources to produce gains in academic performance from a VA perspective. The results of the second stage DEA estimations of the alternative model without the previous student´s academic performance is present in the second column of Table 3. In this model, the average efficiency score is 1.29 (Table 7), which means the percentage of school inefficiency is higher than in the complete model (22.9%). These results, allows concluding that do not control for prior academic achievement produce bias measurements of school efficiency, due to overestimate the school inefficiency. This methodological choice brings much more insights about the performance (efficiency) of schools, depurating this evaluation from factors that are not under control and thus not (fully) manageable.

Table 6: *Correlation between different DEA models*

	Original model	Model 1	Model 2	Model 3	Model 4
Original model	1	0.98	0.97	0.78	0,96
Model 1	0.98	1	0.95	0.77	0,95
Model 2	0.97	0.95	1	0.81	0,94
Model 3	0.78	0.77	0.81	1	0,71
Model 4	0,96	0,95	0,94	0,71	1

Source: Own elaboration. Model 1 does not include the input teacher-student ratio. Model 2 excludes the input computer-student ratio. Model 3 does not control for prior academic achievement and Model 4 excludes socioeconomic status.

Table 7: *Efficiency scores between different DEA models*

	Score	School inefficiency
Original model	1,22710	18,5%
Model 1	1,200780	16,7%
Model 2	1,208404	17,2%
Model 3	1,297012	22,9%
Model 4	1,206138	17,1%

Source: Own elaboration. Model 1 does not include the input teacher-student ratio. Model 2 excludes the input computer-student ratio. Model 3 does not control for prior academic achievement and Model 4 excludes socioeconomic status.

7. Conclusions and Discussion

Improving the quality of education is a key strategy to promote economic development and reduce inequality, especially in countries such as Colombia where school performance is poor even in comparison to that of countries with similar socioeconomic characteristics like population size and public spending in education (for example, in the OECD PISA data). Collecting quantitative evidence periodically about the level of schools' efficiency and the determinants might be relevant to the designing of effective public policy. Since 2017 education is the sector with the highest budget in Colombia; the budget exceeds those for health and defense. However, these extra resources have not translated into better academic performance.

This paper traces the evolution of public schools´ productivity, as well as their efficiency and its determinants from a value-added perspective, during the years 2014–2019 in Colombia. The efficiency scores and their determinants are explored using DEA. The results confirm that, on average, public schools could have increased their efficiency, that is, with the same inputs they could have increased the outputs (Saber11 scores in math and language) by 18.5%. This study is particularly innovative: by controlling for academic previous achievement (Saber9 score) among inputs, a more robust measure of efficiency (due to its capture of the previous academic achievements of students) is generated. Through a second-stage regression, the study identifies the variables that contribute to higher efficiency scores: school size, the proportion of teachers who have completed postgraduate studies, operation in urban areas, and schools with a teacher training orientation.

The positive sign of the coefficients of the standard deviation of the mathematics scores and the SES variables indicate that the more heterogeneous a school is in terms of academic ability and socioeconomic status, the lower the school's efficiency. One potential explanation is that schools focus their efforts on trying to narrow the gap between low and high achievers and are less concerned with improving the performance of the latter. Similarly, to Agasisti (2011), lower efficiency is associated with a higher dropout rate and a larger proportion of female students. Surprisingly, this study finds that inefficiency is also associated to a larger number of female teachers. These gender gaps could be explained by social norms and gender stereotypes. For example, some studies have found that in less egalitarian societies, girls do not have incentives to improve academic performance, because they have less expectation of working in related areas in the future (Guiso et al. 2008; Baker and Jones 1993). Colombia's low enrollment rates of female students in STEM programs has been related to significant and permanent underperformance in math and science on national standardized tests (Abadía and Bernal, 2017). In fact, Gomez Soler et al. (2020) conclude that in Colombia women entering higher education have lower academic competence than their male peers and at the conclusion of higher education for both, the academic gap in math and reading has increased, especially for those enrolled in either STEM programs, public, or accredited universities. Moreover, compared to other countries, Colombia is marked by the greatest gender gap in math and science. One explanation for the higher level of efficiency of bigger schools (with size measured by number of students) is that such schools facilitate interaction and competitiveness between students. With regard to schools that have a greater proportion of teachers who have completed postgraduate studies, such teachers, being the best prepared and qualified, could have a positive impact on academic performance, which would be reflected in greater efficiency. In the same way, schools located in urban areas (vs. rural) and teaching-vocational orientation of schools (vs. a technical orientation) are more efficient. Rural location has a negative effect on efficiency, which has been shown in previous studies (e.g., Cordero et al. (2020)). However, the result regarding school orientation contrasts with those found by Agasisti and Zoido (2019) in developing countries and Agasisti (2013) in Italy, where students attending technical schools perform better than those attending vocational ones. The standard deviation of students' socioeconomic status (SES) and the academic orientation of the institution were not relevant statistically. These latter results contrast with the evidence presented in the literature, which usually indicates that schools with more disadvantaged students are less efficient than other schools and that academic schools are more efficient than technical and vocational ones (Agasisti and Zoido (2019)).

Using the Malmquist index, the paper finds that the annual productivity growth rate was -0.68% for the period of study. Only in 2016 and 2018 did public schools in Colombia see increases in productivity (3.1% and 4.3%, respectively). Although, with the methodological approach performed in this study, it is not possible to establish causal mechanisms that explain the changes observed in school productivity, we highlight that increases in productivity coincide with the period in which the SPP was in effect (2015–2018). This program increased the motivation of students to perform better on the exit examination Saber11 (Bernal and Penney 2019; Laajaj et al. 2022), which is the main outcome considered here. In particular, Bernal and Penney (2019) conclude that the introduction of SPP increased the desire to obtain a high score and the desire to score well. Therefore, scores in the Saber11 test of 2016 improved, mainly for those eligible students on the top of the score distribution. In addition, Authors Laajaj et al. (2022) conclude that the introduction of SPP scholarship also increase the Saber9 scores for eligible students.

Besides students´ motivation, two additional mechanisms could explain the effects of SPP in productivity of public schools. The first possibility, is that schools could have had incentive to allocate more resources to prepare students for the test. However, this is unlikely due to the fact that in Colombia the budget of public schools is very rigid. The main source of resources is the money allocate from the general budget of the nation to the certificated secretariat of education. With this monetary sources, secretariats of education cover the functioning of schools (water, electricity, and light, among others) and pay the payroll. Public schools are not autonomous in spending the money: 90% of the public education budget is destined to cover the payroll of teachers and administrative staff (Bonet et al. 2014; Villar et al. 2016). The amount of money allocated depends mainly on the number of students and teachers in the region. An additional potential mechanism, is that schools devoted time to prepare students for the test. In this line, authors Bernal and Penney (2019) find that the introduction of SPP scholarship does not change statistically significant the hours of study used to prepare the Saber11 test. However, they suggest that teachers may have shifted their focus from preparing students for the exam outside of class hours to doing so during school hours. Therefore, SPP scholarship increased the motivation of eligible students to perform better, and it could also change the students´ compositions of schools over time, aspects that affects the school productivity. While we control for socio-economic composition of the student body – as previously described – it could be the case that students differ in different cohorts due to motivation, intrinsic ability etc. The resulting estimation of the schools' efficiency, then, must be interpreted as a valid measure based on observable features of the schools' resources and activities.

Another explanation for the increase in productivity is the introduction of the ISCE in 2016 and the associated extrinsic incentives that were created for schools. The ISCE monitored public schools in terms of several dimensions and offered in-kind rewards to those that achieve certain performance targets (relative to the schools´ past performance in Saber 11 test). Both the SPP and the ISCE programs were discontinued beginning in the third quarter of 2018 when a new government took power. The period of the implementation and disbanding of these programs matches the increase and decrease in the efficiency frontier shown in this paper, which is evidence of one dynamic associated with the performance of schools.

From the findings presented in this study, it is important to monitor and consider the level of schools´ efficiency to guide educational policies as well as the resource allocation. While one of the main objectives of a country's educational system is to improve academic results, priority should be given to using the currently available resources in the most efficient way. Accordingly, policies can even be differentiated, with more efficient schools being granted additional resources and less efficient ones receiving assistance to improve the management of their operations. In addition, the more-efficient schools can serve as a reference from other schools to transfer good practices.

In Colombia, the educational system is highly segregated. Students with better socioeconomic conditions enroll in private schools (around 30%), which are, on average, of higher quality. This paper focuses the analysis of efficiency solely in public schools. However, futures research could center on measuring the level of efficiency in schools from a value-added perspective. In addition, for evaluating the efficiency of schools in a more complete way, it would be useful to consider later outcomes, like employability and salary of students, and/or enrolment to university.

References

Abadía, Luz Karime and Gloria Bernal (2017), "A widening gap? a gender-based analysis of performance on the colombian high school exit examination." *Revista de economía del Rosario*, 20 (1), 5–31.

Agasisti, Tommaso (2011), "Performances and spending efficiency in higher education: a european comparison through non-parametric approaches." *Education Economics*, 19 (2), 199–224.

Agasisti, Tommaso (2013), "The efficiency of italian secondary schools and the potential role of competition: a data envelopment analysis using oecd-pisa2006 data." *Education Economics*, 21, 520–544.

Agasisti, Tommaso and Giorgio Vittadini (2012), "Regional economic disparities as determinants of student's achievement in italy." *Research in Applied Economics*, 4, 33.

Agasisti, Tommaso and Joanna Wolszczak-Derlacz (2016), "Exploring efficiency differentials between italian and polish universities, 2001–11." *Science and Public Policy*, 43, 128–142.

Agasisti, Tommaso and Pablo Zoido (2019), "The efficiency of schools in developing countries, analysed through PISA 2012 data." *Socio-Economic Planning Sciences*, 68, 100711.

Arbona, Alexei, Víctor Giménez, Sebastián López-Estrada, and Diego Prior (2021), "Efficiency and quality in colombian education: An application of the metafrontier malmquist-luenberger productivity index." *Socio-Economic Planning Sciences*, 101122.

Arias Ciro, Juliana and Alejandro Torres García (2018), "Economic efficiency of public secondary education expenditure: How different are developed and developing countries?" *Desarrollo y Sociedad*, 119–154.

Baker, David P and Deborah Perkins Jones (1993), "Creating gender equality: Cross-national gender stratification and mathematical performance." *Sociology of education*, 66 (2), 91–103.

Bernal, Gloria L. and Jeffrey Penney (2019), "Scholarships and student effort: Evidence from colombia's ser pilo paga program." *Economics of Education Review*, 72, 121–130.

Bernardo, Joyce Santana, Fernanda Maria de Almeida, and Ana Carolina Campana Nascimento (2021), "Educational (in) efficiency from the perspective of disaggregated public expenditures." *Educação & Realidade*, 46.

Bonet, Jaime, Gerson Javier Pérez, and Jhorland Ayala (2014), "Contexto histórico y evolución del SGP en colombia." *Documentos de trabajo sobre economía regional*, 205, 1–72.

Cardoso, Ingrid Gois, Daniel Pacheco Lacerda, and Fabio Sartori Piran (2021), "A DEA model to assess the educational system in brazil." *International Journal of Management in Education*, 15 (3), 245–275.

Chakraborty, Kalyan, Basudeb Biswas, and W Cris Lewis (2001), "Measurement of technical efficiency in public education: A stochastic and nonstochastic production function approach." *Southern economic journal*, 67 (4), 889–905.

Cooper, William W, Honghui Deng, Zhimin M Huang, and Susan X Li (2002), "A one-model approach to congestion in data envelopment analysis." *Socio-Economic Planning Sciences*, 36 (4), 231–238.

Cordero, Jose Manuel, Cristina Polo, and Rosa Simancas (2020), "Assessing the efficiency of secondary schools: Evidence from OECD countries participating in PISA 2015." *Socio-Economic Planning Sciences*, 100927.

De Jorge-Moreno, Justo, Javier Díaz Castro, Diana Victoria Rodríguez Vega, and José Miguel Segura Gutiérrez (2018), "Análisis de la eficiencia educativa y sus factores explicativos considerando el efecto de la titularidad en colombia con datos Pisa 2012." *Revista Desarrollo y Sociedad*, 89–118.

De Witte, Kristof and Laura López-Torres (2017), "Efficiency in education: A review of literature and a way forward." *Journal of the operational research society*, 68, 339–363.

Delprato, Marcos and Germán Antequera (2021), "Public and private school efficiency and equity in latin america: New evidence based on pisa for development." *International Journal of Educational Development*, 84, 102404.

Deutsch, Joseph, Audrey Dumas, and Jacques Silber (2013), "Estimating an educational production function for five countries of latin america on the basis of the PISA data." *Economics of Education Review*, 36, 245–262.

Doran, Harold C and Lance T Izumi (2004), *Putting education to the test: A value-added model for California*. San Francisco: Pacific Research Institute.

Førsund, Finn R and Kjell Ove Kalhagen (1999), *Efficiency and productivity of Norwegian Colleges*. Memorandum No. 11/99. Oslo: Department of Economics, University of Oslo.

Gomez Soler, Silvia C, Luz Karime Abadía Alvarado, and Gloria Lucia Bernal Nisperuza (2020), "Women in STEM: does college boost their performance?" *Higher Education*, 79 (5), 849–866.

Gort, William, Rebecca King, and Alan Weiss (2019), "Measuring school efficiency and effectiveness to inform the design of a school funding formula." *Economic Record*, 95, 65–78.

Gronberg, Timothy J, Dennis W Jansen, and Lori L Taylor (2012), "The relative efficiency of charter schools: A cost frontier approach." *Economics of Education Review*, 31 (2), 302–317.

Guiso, Luigi, Ferdinando Monte, Paola Sapienza, and Luigi Zingales (2008), "Culture, gender, and math." *Science*, 320 (5880), 1164–1165.

Hanushek, E. A. (1986), "The economics of schooling: Production and efficiency in public schools." *Journal of economic literature*, 24 (3), 1141–1177.

Hanushek, Eric A (1979), "Conceptual and empirical issues in the estimation of educational production functions." *Journal of human Resources*, 351–388.

Ikoya, Peter O and D. Onoyase (2008), "Universal basic education in Nigeria: Availability of schools' infrastructure for effective program implementation." *Educational studies*, 34 (1), 11–24.

Instituto Colombiano para la Evaluación de la Educación (ICFES). (2019a), *Saber al Detalle - ¿Qué garantiza la comparabilidad de los resultados en las pruebas Saber realizadas por el ICFES?* 3rd Edition.

Instituto Colombiano para la Evaluación de la Educación (ICFES). (2019b), *Saber al Detalle - ¿Qué se construye el índice de nivel socioeconómico (INSE) en el contexto de las pruebas Saber?*. 4rd Edition.

Iregui, Ana M., Ligia Melo, and Jorge Ramos (2006), "La educación en colombia: análisis del marco normativo y de los indicadores sectoriales." *Revista de economía del Rosario*, 9 (2), 175–238.

Laajaj, Rachid, Andrés Moya, and Fabio Sánchez (2022), "Equality of opportunity and human capital accumulation: Motivational effect of a nationwide scholarship in colombia." *Journal of Development Economics*, 154, 102754.

OECD (2019), "PISA results 2018." recovered from https://www.oecd.org/pisa/publications/pisa-2018-results.htm.

Oh, Dong-hyun (2010), "A global Malmquist-Luenberger productivity index." *Journal of productivity analysis*, 34, 183–197.

Pastor, J. T. and C. K. Lovell (2005), "A global Malmquist productivity index." *Economics Letters*, 88 (2), 266–271.

Portela, MCAS and AS Camanho (2010), "Analysis of complementary methodologies for the estimation of school value added." *Journal of the Operational Research Society*, 61 (7), 1122–1132.

Portela, MCAS, AS Camanho, and Abolfazl Keshvari (2013), "Assessing the evolution of school performance and value-added: trends over four years." *Journal of Productivity Analysis*, 39, 1–14.

Queiroz, M. V. A. B., R. M. B. Sampaio, and L. M. B. Sampaio (2020), "Dynamic efficiency of primary education in brazil: Socioeconomic and infrastructure influence on school performance." *Socio-Economic Planning Sciences*, 70, 100738.

Rattini, Veronica (2014), "The causal effect of scholarships targeted at low income students on performance: Evidence from italy." 968.

Rubin, Irene S. (2019), *The politics of public budgeting: Getting and spending, borrowing and balancing*. CQ Press.

Salazar Cuéllar, Andrés Felipe (2014), "The efficiency of education expenditure in latin america and lessons for colombia." *Revista Desarrollo y Sociedad*, 19–67.

Simar, Leopold and Paul W Wilson (2007), "Estimation and inference in two-stage, semi-parametric models of production processes." *Journal of econometrics*, 136, 31–64.

The World Bank (2020), "[base de datos]." recovered from https://data.worldbank.org/indicator/SE.XPD.TOTL.GD.ZS.

Villar, Leonardo, Natalia Salazar, Mary Simpson, Martha Delgado, Roberto Steiner, Juan G Zapata, Orlando Gracia, Carlos A Mesa, Gabriela González, Felipe Trujillo, et al. (2016), "Sistema general de participaciones y su vínculo con las políticas e incentivos dirigidos a la excelencia docente y el mejoramiento continuo de la calidad de la educación."

Woessmann, Ludger (2016), "The importance of school systems: Evidence from international differences in student achievement." *Journal of Economic Perspectives*, 30 (3), 3–32.

Wolszczak-Derlacz, Joanna (2018), "Assessment of TFP in european and american higher education institutions–application of malmquist indices." *Technological and Economic Development of Economy*, 24 (2), 467–488.

A. Appendix: Additional Tables

Table A.1 *Total Number of schools versus the number included in sample*

Year	Number of schools - original	Number of schools – final sample
2014	5,176	3,854
2015	5,016	3,854
2016	5,520	3,854
2017	5,749	3,854
2018	5,909	3,854
2019	6,415	3,854

Source: Own elaboration.

Table A.2 *Descriptive statistics of variables used in the empirical analysis*

Non-discretionary variables												
Year	Variable	sd_SES	sd_Saber11 _math	sd_Saber11 _language	dropouts	academic	normal	urban	size	tc_postgraduate	tc_female	tc_femaleStud
2014	Mean	5.16	7.94	8.08	52.11	0.82	0.02	0.65	1247.87	0.91	32.35	0.44
	Std. Dev.	1.2	1.27	1.17	66.75	0.39	0.12	0.48	899.99	0.12	22	0.13
2015	Mean	4.91	9.41	7.56	50.22	0.82	0.02	0.65	1248.65	0.92	32.68	0.45
	Std. Dev.	1.27	1.66	1.19	63.3	0.39	0.12	0.48	897.25	0.1	21.82	0.13
2016	Mean	4.96	9.44	8.01	48.7	0.82	0.02	0.65	1220.52	0.93	33.05	0.44
	Std. Dev.	1.16	1.33	1.08	62.76	0.39	0.12	0.48	876.79	0.1	21.72	0.12
2017	Mean	6.2	9.69	8.15	42.36	0.82	0.02	0.65	1207.87	0.93	32.8	0.44
	Std. Dev.	1.14	1.37	1.08	56.6	0.39	0.12	0.48	876.68	0.09	21.63	0.12
2018	Mean	6.79	9.51	8.34	39.97	0.82	0.02	0.65	1188.32	0.94	33.29	0.45
	Std. Dev.	1.25	1.38	1.13	52.94	0.39	0.12	0.48	862.74	0.09	22.84	0.12
2019	Mean	6.83	9.63	8.82	36.51	0.82	0.02	0.65	1199.22	0.94	33.78	0.45
	Std. Dev.	1.28	1.35	1.16	51.11	0.39	0.12	0.48	870.74	0.09	23.01	0.12
Total	Mean	5.81	9.27	8.16	44.98	0.82	0.02	0.65	1218.74	0.93	32.99	0.45
	Std. Dev.	1.48	1.52	1.2	59.45	0.39	0.12	0.48	881.01	0.1	22.18	0.12
	Min	0	1.73	0	0	0	0	0	4	0	0	0
	Max	12.9	20.85	17.23	1021	1	1	1	7297	1	231	1
	Obs	23124	23124	23124	23124	23124	23124	23124	23124	23124	23124	23124

Source: Own elaboration.

Table A.3 *Definition of variables used in the empirical analysis*

Variable	Name
sd_SES	standard deviation of SES (Socioeconomic Status)
sd_math11	standard deviation of Saber11$_{m}ath$
dropouts	number of dropouts in previous year
academic	dummy for academic schools
normal	dummy for normal schools
urban	dummy for schools located in urban area
size	number of students
female_stud	proportion of female students
tc_postgrade	proportion of teachers with postgraduate studies
tc_female	number of female teachers
state	a set of dummies for schools located in each state of Colombia
vt	time dummies

Source: Own elaboration.