Economic Analysis and Policy 77 (2023) 897–912

Contents lists available at ScienceDirect

Economic Analysis and Policy

journal homepage: www.elsevier.com/locate/eap

Analyses of topical policy issues

Does admission type matter? An analysis of the performance of federal high school students in Brazil

Lauana Rossetto Lazaretti *, Marco Túlio Aniceto França

Pontifical Catholic University of Rio Grande do Sul, Avenida Ipiranga 6681, Building 50, Partenon, Porto Alegre, RS, 90619-900, Brazil

ARTICLE INFO

Article history:
Received 30 June 2022
Received in revised form 4 December 2022
Accepted 22 December 2022
Available online 29 December 2022

JEL classification:
C31
D04
I21

Keywords:
Selection effect
High school
Performance on Saeb
State and federal aid

ABSTRACT

This study aims to verify whether a difference exists between the forms of entry (lottery draw vs. merit-based) on the performance of students entering the federal education schools. The data used were collected from students in their last year of high school in relation to mathematics and Portuguese. As each federal education unit has administrative independence in terms of the selection method, the identification strategy is based on assigning students to two groups: treatment (lottery draw) and control (merit-based admission process). Three matching methods were employed to estimate these effects. The empirical results indicate that an effect exists between 0.14 and 0.40 standard deviation smaller in mathematics for students who enter by lottery draw, with differences in gender and the distribution of grades. Among the female students, the effect was greater in Portuguese, and among the male students, the effect was greater in mathematics. Students in the lower quantiles of grade distribution showed the greatest difference.

© 2022 Economic Society of Australia, Queensland. Published by Elsevier B.V. All rights reserved.

1. Introduction

Following the publication of the Coleman report in 1965, the empirical literature on understanding the different educational outcomes of schools gained greater relevance. Studies under this theme have found that the selection process and the self-selection of students are important mechanisms to explain these differences in the peer effect and school segregation. Thus, great efforts have been made to understand what is actually due to the school effect and what is a result of the characteristics of the students at each school.

The literature on school sector effects sought to determine how much of the performance of certain schools is due to the so-called "school effect" and to the characteristics of the students (Angrist et al., 2006; Baude et al., 2020; Benevides and Soares, 2020; Chudgar and Quin, 2012; Hanushek et al., 2007; Howell et al., 2002; Hoxby and Rockoff, 2004; Lubienski and Lubienski, 2006; McEwan, 2004; Sapelli and Vial, 2002; Somers et al., 2004). Controlling for the endogenous characteristics of students demonstrated that the results were less than expected or insignificant. Part of the educational outcomes reflects the characteristics of the group of students, known as the "selection effect".

Thus, the selection process is crucial in explaining the differences in performance between schools, contributing to the discussion on the quality of the Federal Education (FE) in Brazil. In its first large-scale and comparable assessment of schools that offer secondary education, the 2017 Basic Education Assessment System (Saeb) pointed to a difference in performance between units of the educational system. The most recent data from 2019 indicate that the FE shows better

* Correspondence to: Ipiranga Avenue, 6681, Building 50, Paternon, Porto Alegre, Brazil.
E-mail addresses: lauana.lazaretti@gmail.com (L.R. Lazaretti), marco.franca@pucrs.br (M.T.A. França).

https://doi.org/10.1016/j.eap.2022.12.019
0313-5926/© 2022 Economic Society of Australia, Queensland. Published by Elsevier B.V. All rights reserved.

performance than other schools that comprise the public (state and municipal) and private school. The average FE score in Portuguese (LP) was 335, 321.85 for the private, and 271.08 for the remainder of the public schools.[1] The difference was even greater for mathematics (MT), with the FE reaching 356.13 points, whereas the private and public schools scored 334.45 and 267.76, respectively.

Therefore, the question arises: better schools or better students? Performance cannot be attributed entirely to the school effect, as the backgrounds of students who enter the FE differ from those of other students who remain in the public school. Thus, a part of the FE performance may be related to the student selection process.

In this context, the present study seeks to provide empirical evidence on the magnitude of the difference between the forms of entry (lottery draw vs. merit-based) of students entering federal education schools on their performance in the last year of high school. Therefore, this study investigates the extent to which the lottery-based process is superior or inferior to the merit-based process[2] from the point of view of student's performance. The data used refer to students in the last year of high school, available in the Basic Education Assessment System (Saeb), for Mathematics and Portuguese based on the autonomy of the teaching units in the selection process and admission of students. Given the existence of two forms of admission, lottery draw and merit-based, students were assigned to two groups: treatment and control, respectively. The assumption is that all students have completed their enrollment (choice of school) and that, conditional on observable characteristics, the difference in performance between the groups can be attributed to the selection process.

Three matching methods were used to estimate the effect: propensity score matching (PSM), coarsened exact matching (CEM), and entropy balancing. The empirical results indicate that an effect exists between 0.14 and 0.40 standard deviation smaller in mathematics for students who enter by lottery draw, with differences in gender and the distribution of grades. Among the female students, the effect was greater in Portuguese, and among the male students, the effect was greater in mathematics. Students in the lower quantiles of grade distribution showed the greatest difference.

This study fills an important gap in the argumentative discussion on the superior performance of the FE. Some of the mechanisms that explain the best outcomes are found in the literature on peer effects. Lazear (2001) pointed out two fundamental points of class organization and peer effect: (1) the segregation of students by academic skills (tracking) maximizes educational outcomes; contrastingly, it stratifies schools and increases inequality; and (2) without student tracking, integration tends to contribute to the peer effect; that is, students with a broader background can help others increase their performance.

In summary, the effect of the school sector is smaller than expected. Most of the higher average performance of certain schools is explained by the parents' background, which influences the choice of school, and the way students are admitted, which establishes the performance criteria (Lara et al., 2011). Thus, efforts have been made to determine what is actually due to each type of effect. In this study, the effect of school choice becomes less relevant, given that all students have already performed when they compete for a place in federal schools.

Additionally, to the best of our knowledge, no other studies have measured the selection effect in secondary and elementary education; investigating the existence of the school effect is also common. Therefore, this study is the first according to our knowledge to provide empirical evidence of secondary education in Brazil. It should be noted that in addition to free and quality education, two other characteristics are part of the utility function of parents' choice of school: (1) when attending public schools, students can later compete for quotas in higher education; and (2) peer effect gains—in more homogeneous classes (in the case of the merit-based admission process), performance gains are expected.

The remainder of the paper is structured as follows. Section 2 deals with the context of Brazil's federal secondary education schools. Section 3 describes the empirical strategy and data sample. Section 4 discusses the results. Finally, Section 5 provides the conclusions.

2. Characteristics of the federal secondary education system in Brazil

The Federal Basic Education in Brazil comprises 38 units of federal education institutes (IFs), the Federal University of Technology (UTFPR), two Federal Centers for Technological Education (Cefet[3]), and 23 technical schools linked to Federal Universities and Colégio Pedro II. The current federal education structure is derived mainly from the expansion plan of the Federal Network of Professional and Technological Education, with two main legal frameworks. The first comes from Law No. 11,195 of November 18, 2005, which proposed the expansion of professional education. Second, based on Law No. 11,892 of December 29, 2008, the Federal Institutes of Education, Science and Technology were created, and in addition to new educational institutions, other federal school were named as such. Altogether, 500 new teaching units have been created in all regions of the country. In 2006, 144 units were in operation, which increased to 643 units in 2018 (MEC, 2021).

By offering professional education integrated with students entering the federal secondary schools, the federal schools now share responsibility for the last stage of basic education with state and private education schools. In 2020, based on

[1] The municipal school serves less than 1% of the student population. When analyzed separately, the average for the state school is 271, and that for the municipal school is 285.29. Thus, for the sake of simplicity, we use the term "public" for state and municipal schools.

[2] What we defined as the selection process is the merit-based admission process. Therefore, the control group was formed by students who were admitted after doing an exam and the treatment group was formed by students who were admitted after a lottery.

[3] Federal Centers for Technological Education Celso Suckow da Fonseca - CEFET-RJ and Minas Gerais - CEFET-MG.

the Basic Education Census in Brazil, 3% of all enrollments in secondary education were in the federal education, 84% in the state, and 13% in the private. Thus, despite the increased number of places in the federal schools, enrollments are limited by the service capacity of teaching units, and no guarantee exists that all applicants will be admitted.

The federal schools differ in terms of the amount of investment per student and academic training for teachers, which provides them with higher remuneration. The annual expenditure per enrollment in the federal school was R$15,741.98 in 2019, whereas a full-time high school student in the rest of the public schools cost R$6,564.00 in 2020 (PNP, 2020; SIMCAQ, 2021). The faculty primarily comprised master's and doctorate degree (51.39% and 34.09%, respectively).

The student goes through the selection stage after being given the option of attending a school that is part of the federal education. The teaching unit can adopt different types of admissions under Law No. 11,892 passed on December 29, 2008, which provides for the autonomy of entities included in the federal education. By using school censuses from 2013 to 2018, the type of admission of students in the federal schools can be identified.[4] The most common, which represents 86% of enrollments in 2018, is the selection exam through knowledge tests or evaluation of the academic record from the previous stage (in the case of high school, grades in the final years of elementary school). To a lesser extent, students admitted through a lottery draw account for 1.5% of all integrated secondary school students—approximately 2700 students. The remainder came from the school transfers.

The units that used lotteries as a selection method belonged to the Federal Institutes of Brasília (IFB), Acre (IFAC), Santa Catarina (IFSC), Rio Grande do Sul (IFRS), and Paraná (IFPR). The selection approach adopted by educational institutions may be the same for all campuses or may differ among or within them. At the IFB, the only form of selection is through a lottery draw, whereas at the IFAC, lottery draws were the selection process used from 2010 to 2015. At the IFSC, IFRS, and IFPR, some campuses or programs that were created in 2016 adopted the lottery draw for selection. Finally, some teaching units adopted the lottery draw for the remaining spots.

3. Empirical strategy

Considering the FE independence scenario for the adoption of the selection method, the identification strategy was based on designating students into two groups. The treatment group comprised of students who were selected through lottery draws. The control group also registered but underwent a selection process (merit based) to attend high school in the federal government. For both groups, students must compete for a place at the FE.

This strategy has advantages and disadvantages. On the one hand, it minimizes the problem of self-selection or selection bias, which represents one of the main factors of school choice. Regardless of the selection type, the students registered, which signals that the federal school is within the range of school choice. In addition, by comparing students from only one school system, homogeneity of the teaching structure was guaranteed. To ensure the validity of the hypothesis and rule out possible problems with other variables linked to the characteristics of the schools affecting the results, the mean difference for a set of characteristics between the schools in the two groups under analysis (treatment and control) was tested. The result of the statistical tests demonstrated that they did not reject the hypothesis of equality of means.[5] As for the salaries of teachers in federal schools, they follow federal legislation and are equally followed by all schools.

On the other hand, the students analyzed were those in the last year of high school, and it must be considered that the sample represents the "winners". Although the high school dropout rate is only 1.7% in federal schools.[6] Among schools that adopt different forms of selection, the average remains statistically the same.[7]

[4] The identification variable is divided into nine admission categories: (1) No selection process, (2) Lottery draw, (3) Transfer, (4) Selection exam without place reservation, (5) Selection exam, place reserved for public school students, (6) Selection exam, place reserved for students from the public school, with low income and self-declared black, multiracial or indigenous, (7) Selection exam, place reserved for other affirmative action programs, (8) Other form of admission, and (9) Selection exam, place reserved for low-income students from the public school. As of 2019, this variable was discontinued; the last information on the selection process is for the year 2018.

[5]

Variable	Mean		Difference	*t-test*
	Treatment	Control		
Class size	28.77	31.34	2.56	1.27
Class time	7.52	7.43	−0.09	−0.11
Teaching regularity	3.29	3.24	−0.04	0.32
Percentage of teachers with higher education	99.0	98.8	−0.02	−0.13
Adequacy of Teacher Training				
Group 1	67.74	69.99	2.3	0.42
Group 2	17.50	12.99	−4.50	−1.35
Group 3	5.51	7.15	1.63	0.69
Group 4	8.25	8.79	0.54	0.13
Group 5	1.00	1.05	0.05	0.04

[6] The failure rate of federal secondary schools is 1.7%, higher than state schools (5.5%) and private schools (0.2%) in 2019.

[7] The average dropout rate for schools using performance-based selection is 1.44%, the average dropout rate for schools using the lottery is 1.91. Under the hypothesis of equal means, the test of equality of means presented a $t = -0.59$, which indicates that the null hypothesis was not rejected. That is, the means have no statistical difference.

The students in the two groups, treatment and control, tended to differ in terms of observable characteristics. In the absence of a random allocation between the two groups, the simple specification of a model via ordinary least squares (OLS) leads to an inconsistent estimation as there may be an error term (ε) correlated with the variable of interest (T) – in this study, admission through a lottery draw – that is, $cov\ (T, \varepsilon) \neq 0$.

To estimate a consistent parameter, finding a technique that addresses the problem listed in the OLS estimation is necessary. Among the possible techniques to resolve this problem, balancing observable characteristics is a widely used tool in the empirical literature. Given the existence of a selection process conditional on observable characteristics, the treatment was assumed to be random.

With such reservations, the following sections formalize the model in terms of potential results and present three estimators for balancing the observable characteristics between the treatment and control groups. The first, widely used in the literature and taken from the seminal text by Rosenbaum and Rubin (1983), is propensity score matching (PSM). PSM comprises of calculating a student's probability of receiving treatment conditional on his/her observable characteristics. The second method is coarsened exact matching (CEM), and the third is entropy balancing. Both methods seek to reduce matching imperfections by using propensity scores. The CEM advances the exact matching model and forms groups of similar strata. Entropy balance data pre-treatment reduces the imbalance of covariates between the treatment and control groups, increasing the independence of the treatment variable. The use of these three methods aims to present robust results regarding the balance of observable characteristics, and consequently, greater precision in determining causal effects.

3.1. Potential outcomes model

When taking the standardized performance test (Saeb) in the third year of high school, students from the federal school obtained grade Y. Thus, $y\,(1)$ is determined to be the potential outcome of the student who was admitted through the lottery draw—treated, and $y\,(0)$ is the potential outcome of the same student if he or she was admitted through a test—untreated. Thus, the student's observed outcome is $y_i = T_i y_i\,(1) + (1 - T_i)\,y_i\,(0)$, where $T_i = 1$ indicates that the student received treatment, and $T_i = 0$ indicates that he or she did not.

To identify the effect of the treatment for each student, the potential outcome must be subtracted the two possibilities: $\tau_i = y_i\,(1) - y_i\,(0)$. However, the student was only subject to one of the events; that is, they either received or did not receive the treatment. If $T_i = 1$, then $y\,(0)$ is unobserved; likewise, if $T_i = 0$, then $y\,(1)$ is unobserved. Consequently, a missing data problem exists. Generalizing among the groups, the Average Treatment Effect on the Treated (ATT) is $ATT = \mathrm{E}\,[y(1)|T = 1] - \mathrm{E}\,[y(0)|T = 1]$. Finding a group that represents the counterfactual of the students treated in the case that they have not received treatment. The closer this control group was to the students' counterfactual, the smaller the measurement bias in identifying the causal effect.

To represent the counterfactual, matching estimators are used to bring students closer in terms of observable characteristics (X) to identify the effect of treatment on the treated.

3.2. Propensity score matching (PSM)

Given that $y_i(0)$ is observed only among individuals who are not treated and assignment to treatment is not random, Rosenbaum and Rubin (1983) proposed two assumptions to ensure that ATT is identified.

The first hypothesis was the selection of observables or conditional independence. Conditional on the variables contained in the vector of observable characteristics X, the treatment is random between treatment and control groups. In other words, the potential results are independent of the treatment (Eq. (1)).

$$y_i(0), y_i(1) \perp T|X_i. \tag{1}$$

In addition, students in the treatment group must have a peer in the control group who represents the case in which they did not receive treatment. This implies the second hypothesis of the model, the condition of overlap or common support (Eq. (2)).

$$0 < Pr\,[T = 1|X] < 1. \tag{2}$$

This ensured that students with the same characteristics had a positive probability of being treated. Rosenbaum and Rubin (1983) also advanced the dimensional difficulty of exact matching. The more characteristics included in vector X, the lower the chance of finding a match that is compatible with the characteristics of the treated individuals. To solve this, the hypothesis of selection of the observables is also valid when conditional on the probability of receiving the treatment (Eq. (3)).

$$y_i\,(0)\ \perp T\,|X_i \Rightarrow y_i\,(0)\ \perp T|\,p\,(X_i)\,, \tag{3}$$

where $p(X_i)$ is the propensity score, which represents the probability of receiving treatment from the set of features of the vector X. In practice, $p(X_i)$ is estimated using linear probability models, either probit or logit models.

Thus, for Caliendo and Kopeinig (2008), the ATT estimator in PSM is the difference in means between the treated and controls located within the common support and weighted by the propensity score distribution (Eq. (4)).

$$ATT_{PSM} = \mathrm{E}_{p(X)|T=1}\left\{\mathrm{E}\left[Y(1)|T=1, p(X)\right] - \mathrm{E}\left[Y(0)|T=0, p(X)\right]\right\} \tag{4}$$

To define which observations are part of the control group, one can choose a measure of the proximity of the propensity score, such as n nearest neighbors, caliper and radius, or the kernel matrix. For this study, various metrics were used to estimate the results and doubly robust method with inverse probability treatment weighting (IPWRA) proposed by Emsley et al. (2008) and Lunceford and Davidian (2004). The double-robust technique allows for estimation with robust and clustered standard errors, which is not possible in PSM.

Although consolidated in the literature, PSM is susceptible to criticism regarding the magnitude of unobservable characteristics that can bias results. The following can be mentioned in the case of the study: low self-confidence or insufficient information about parents regarding their children's knowledge. To test whether the model is affected by such a problem, the next section presents a sensitivity analysis technique.

3.3. Sensitivity analysis

With the assumption of conditional independence in PSM, the idea is that the variables that determine the condition of the treated are included in the vector of observable characteristics X. According to Rosenbaum (2002), formally speaking, the probability of receiving treatment (π_j) is a function $\lambda\left(x_j\right)$ of this vector of observable characteristics of each j individual. Thus, if two individuals (j and k) have the same covariates ($x_j = x_k$) but the probability of receiving treatment is different between them ($\pi_j \neq \pi_k$), an indication of bias exists; it may be that unobservable characteristics affect the treatment.

Although testing the conditional independence hypothesis was not possible, Rosenbaum (2002) proposed a sensitivity analysis of the results. The idea is to attempt to understand the size of this difference to change the conclusions of the study.

Given that the chances of receiving treatment are $\pi_j/(1-\pi_j)$ and it is also valid for individual k, the odds ratio between j and k is a number $\Gamma \geq 1$ – Eq. (5).

$$\frac{1}{\Gamma} \leq \frac{\pi_j(1-\pi_k)}{\pi_k(1-\pi_j)} \leq \Gamma \text{ For every } j \text{ and } k \text{ with } x_j = x_k \tag{5}$$

When the chances of receiving treatment are equal, $\Gamma = 1$, and no bias exists. If $\Gamma > 1$, for example, $\Gamma = 2$, although individuals have similar observable characteristics, one of them is twice likely to receive treatment. In the latter case, bias was observed. Thus, Γ represents the degree of departure from the unbiased outcome. For several values of Γ, Rosenbaum (2002) pointed out that if the outcomes are sensitive to values very close to $\Gamma = 1$, the validity of the idea that no other unobservable characteristics (u) affect the treatment can be questioned.

3.4. Coarsened exact matching - CEM

As in PSM, coarsened exact matching (CEM) presupposes the independence of conditional treatment in observables: $P\left\{T|X, y\left(0\right), y\left(1\right)\right\} = P\left(T|X\right)$. In this case, the treatment designation is independent of potency outcomes when conditioned on X.

Unlike PSM, which calculates the probability of each unit of analysis being treated, CEM inserts an interval (bin) between the values of the X vector variables to determine substantially significant groups (Blackwell et al., 2009; Iacus et al., 2011; Iacus and King, 2008). In addition, CEM makes it possible to choose variables with exact matching, for example, categorical variables, such as gender or color/race. The CEM algorithm is well known for its simplicity and applicability to large data samples. In summary, the algorithm procedures are as follows (1) make a copy (X*) of the variables contained in vector X, (2) perform the groupings with the cuts defined or automatically generated by the algorithm, (3) create a stratum for each X* observation and allocate each observation to a stratum, and (4) assign strata to the original data (X). Strata that did not have at least one individual in the treatment group and one in the control group were excluded. The ATT can be generated automatically by the difference in means between the groups or by linear regression, and matching by CEM can be used together with another a posteriori method.

Given that the bin is defined ex ante to the matching, one advantage of CEM is that it reduces the imbalance bias a priori, whereas in most matching methods, the tests to verify the adjustment are performed ex post. This attribute is called monotonic imbalance bounding (MIB), where: the maximum imbalance of some empirical characteristics is limited before matching is performed.

3.5. Entropy balancing

With the idea of improving balance, in the case of entropy, the objective is to make the treatment variable more independent (orthogonal) from individual characteristics. To do this, pre-treatment was performed between the variables such that the characteristics of the units of analysis were re-weighted in the weight function to ensure a better adjustment of the two groups (treatment and control).

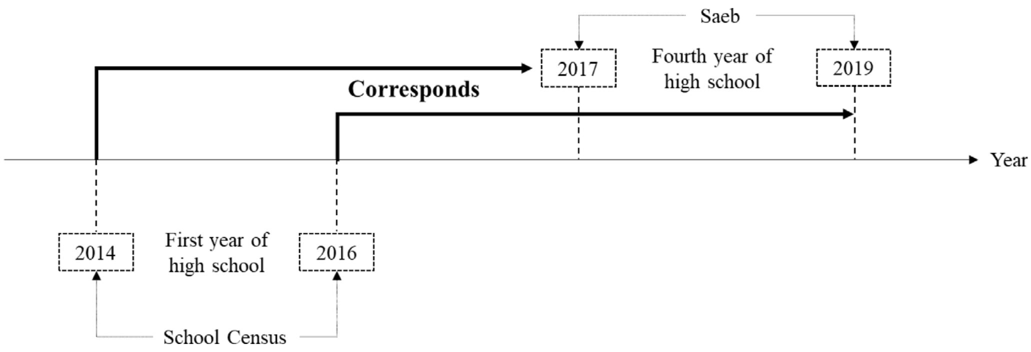


Fig. 1. Identification of admission type.
Source: Prepared by the authors.

Based on the same model of potential outcomes formally written by Hainmueller (2012), entropy balancing reweighs the control group (Eq. (6)).

$$E[\widehat{y(0)|T=1}] = \frac{\sum_{\{i|T=0\}} y_i \omega_i}{\sum_{\{i|T=0\}} \omega_i} \tag{6}$$

where ω_i denotes the weight of the control unit. The choice of weight for reweighting is $\min_{\omega_i} H(\omega) = \sum_{\{i|T=0\}} h(\omega_i)$, where *h(.)* is a distance metric conditional on the shortest distance between the observations.

In summary, this means that several moments of the sample distribution are created from the set of characteristics of the treatment group, and with the reweighting of the control group, the X densities of the two groups becomes very close. Therefore, unlike in the case of PSM, which first estimates unit weights on the basis of a probabilistic distribution and then verifies the adjustment, the weights are estimated directly, considering the moments of the sample.

3.6. Sample and database

The performance data of Brazilian students were extracted from the Basic Education Assessment System (Saeb). This survey was a large-scale assessment carried out every two years and students are evaluated in Portuguese and Mathematics. In addition to the knowledge test, the students answered a questionnaire containing information about their education, parents' education, socioeconomic characteristics, and demographic profile. Additionally, to identify the selection method for each teaching unit, the School Census of Basic Education, which is carried out annually and on a census basis, is used.

As students in the federal schools attended secondary education integrated with a technical course, the students in this analysis took four years to completed the last stage of basic education, whereas students in the introductory courses complete it in three years. Therefore, to identify the method of student admission, it must be verified which selection process each teaching unit adopted in the four-year period prior to the analysis (Fig. 1).

The cohort of students who took the Saeb test in 2017 and 2019 entered their first year of high school in 2014 and 2016, respectively. Thus, through the teaching unit code, the type of selection can be identified, and subsequently, the cohort of students who entered through the lottery draw or a selection exam (merit-based selection process). Therefore, we used datasets from 2017 and 2019. In addition, the period was determined by whether high schools were surveyed in 2017 for the first time.

Of the 25,017 students who took Saeb in their fourth year of high school in the federal school, 5638 did not have school identification and were only included in the calculation of state and national averages.[8] In addition, to maintain a comparison group without other students entering school through the lottery draw, the schools that used this selection method for the remaining spots were excluded from the sample. The final sample included 348 students who entered through the lottery draw and 18,637 newcomers through the selection exam.

The characteristics of the treated and control students before and after matching are shown in Table 1. Differences with negative values indicated that students who entered through the lottery draw lots had higher values than students in the selection process, and the inverse was valid in the case of positive values.

Most of the characteristics were statistically different between the two groups before matching, except for parents having completed high school, students living with both parents, and socioeconomic status. After matching, the averages of the characteristics showed no statistically significant differences.

8 Through Law No. 13,005, of June 2014, the results of students with school identifications are disclosed to schools with an attendance rate of 80% of enrolled students. As of Inep Ordinance No. 366, of April 29, 2019, the cumulative enrollment of at least 10 students at the time of application of the assessment instrument is still included.

Table 1
Descriptive statistics.
Source: Prepared by the authors.

Variable	Matching	Mean		Difference	*t-test*
		Treatment	Control		
Caucasian	Before	0.23	0.34	0.11***	−3.97
	After	0.25	0.26	0.01	−0.28
Black	Before	0.12	0.09	−0.03*	1.79
	After	0.25	0.26	0.01	1.15
Multiracial	Before	0.49	0.37	−0.12***	4.67
	After	0.53	0.55	0.02	−0.65
Asian	Before	0.04	0.02	−0.02***	2.62
	After	0.04	0.04	0.00	0.21
Mother completed high school	Before	0.27	0.31	0.04	−1.50
	After	0.30	0.31	0.01	−0.27
Mother completed higher education	Before	0.31	0.26	−0.05**	2.18
	After	0.33	0.30	−0.03	0.79
Father completed high school	Before	0.24	0.28	0.04	−1.60
	After	0.26	0.26	0.00	−0.19
Father completed higher education	Before	0.18	0.13	−0.05***	2.62
	After	0.20	0.20	0.00	−0.10
Works outside the home	Before	0.15	0.30	0.15***	−6.16
	After	0.08	0.08	0.00	−0.15
Went to daycare or preschool	Before	0.52	0.70	0.18***	−6.71
	After	0.57	0.57	0.00	−0.16
Public education	Before	0.77	0.68	−0.09***	3.46
	After	0.83	0.84	0.01	−0.44
EJA	Before	0.32	0.55	0.23***	−8.08
	After	0.36	0.36	0.00	−0.08
Lives with both parents	Before	0.54	0.56	0.02	−0.80
	After	0.59	0.64	0.05	−1.43
Failure	Before	0.23	0.27	0.04**	−1.97
	After	0.16	0.19	0.03	−0.86
Dropout	Before	0.12	0.17	0.05**	−2.15
	After	0.06	0.06	0.00	−0.00
Parents are encouraging	Before	0.82	0.70	−0.12***	4.93
	After	0.89	0.90	0.01	−0.40
Computer	Before	0.70	0.80	0.10***	−4.60
	After	0.69	0.68	−0.01	0.35
Socioeconomic level	Before	0.04	−0.04	−0.08	1.46
	After	0.04	0.01	−0.03	0.46

Note: 1. *, **, and *** represent 90%, 95%, and 99% confidence levels, respectively.

In terms of color/race, the treatment group had 12 percentage points (p.p.) more multiracial individuals and fewer Caucasian individuals than did the control group. On average, after matching, 53% of the sample comprised of self-declared multiracial students; Caucasian and black students represented 25% each, and 4% were Asian. The proportion of mothers who had completed high school and higher education was greater than that the fathers. Among mothers, roughly 30% had completed both levels of education, and among fathers; 26% had completed secondary education, and 20% had completed higher education.

Students in the control group worked more outside the home than the treatment group did (30% and 15%, respectively). The proportion of the control group was also higher among those who went to day care or preschool (18 p.p.) and those who studied in classes for youth and adult education (EJA) (13 p.p.). A greater proportion of the students in the treatment group studied in public schools (5 p.p.).

Students who lived with their parents (fathers and mothers) represented more than half of the sample, and there was no statistically significant difference between the two groups. Regarding student performance, those who entered by grade had a higher proportion of failure (4 p.p.) and dropout (5 p.p.) students. Students in the treatment group were encouraged by their parents to study more often than those in the control group (82% and 70%, respectively).

Regarding socioeconomic characteristics, although the control group had more students with computers (10 p.p.), the variable that indicated the socioeconomic level of the students did not show a statistical difference between the two groups of students. This variable is constructed through principal component analysis, with the extraction of the first component, which includes the characteristics of the student's residence, such as having a refrigerator,

Table 2
Differences in means for Portuguese and Mathematics grades.
Source: Prepared by the authors.

Subject	Mean		Difference	*t-test*
	Treatment	Control		
Portuguese	297.73	315.35	17.61***	2.55
Mathematics	297.12	326.52	29.39***	2.88

Note: 1. *, **, and *** represent 90%, 95%, and 99% confidence levels, respectively.

Table 3
Results of the treatment effect on the treated using PSM.
Source: Prepared by the authors.

	OLS (a)	NN (1) (b)	NN(1)WR (c)	NN (3) (d)	NN (5) (e)	Kernel (f)	Radius (g)	IPWRA (h)
Portuguese	−0.09	−0.04	−0.09	−0.04	−0.07	−0.09*	−0.20***	0.06
	(0.100)	(0.036)	(0.068)	(0.059)	(0.056)	(0.050)	(0.048)	(0.044)
Mathematics	−0.24**	−0.27***	−0.28***	−0.24***	−0.24***	−0.26***	−0.40***	−0.14***
	(0.113)	(0.080)	(0.076)	(0.066)	(0.063)	(0.057)	(0.054)	(0.054)

Notes: 1. *, ** and *** represent 90%, 95%, and 99% confidence levels, respectively. 2. OLS = ordinary least squares; NN (1) = nearest neighbor – first nearest neighbor; NN (1) WR = first nearest neighbor without replacement; NN (3) = first three nearest neighbors; NN (5) = first five nearest neighbors; radius = 0.10; IPWRA = inverse probability-weighted regression adjustment.

number of bedrooms, televisions, bathrooms, cars, and washing machines. The higher the value, the higher the students' socioeconomic level. Appendix A presents the adequacy tests used in the analysis. The procedure was carried out given that the Saeb questionnaire did not include information on students' family income.

Regarding the performance of students in Portuguese and Mathematics, the difference in means test confirmed that disparities existed between the treatment and control groups (Table 2).

Students who entered school through merit-based selection exhibited superior performance in both subjects. Students in the treatment group had a mean score on the Saeb scale of 297.73 in Portuguese and 297.12 in Mathematics, whereas those in the control group achieved 315.35 points in Portuguese and 326.52 in Mathematics. The difference between the groups was 17.61 points for Portuguese and 29.39 points in mathematics. As the analyzed students were from two editions of the Saeb, to minimize the effect of the cohort on the grades, the outcomes are presented as a standardized score with a mean of zero and a standard deviation (SD) equal to one.

4. Results and discussion

To isolate the selection effect of students entering the federal education schools, three matching methods were used, and combinations were presented. The first is PSM, for which, as presented in the previous section, after its use, the difference in the means of the observable characteristics was not statistically significant, which represents an indication of adjustment. Appendices B and C present the bias size test and the probability score distribution for the treatment, respectively. The reduction in the values of the mean bias, median bias, and pseudo R2 test as well as the overlapping of the propensity scores of the treatment and control groups reinforces the hypothesis of adjustment and balancing of the sample.

The first step in identifying the effect of treatment on those treated via PSM is to calculate the probability of receiving treatment, that is, having entered school through the lottery draw. The variables used for the estimation are presented in Table 1 in Section 3.6. The model indicates that the probability of receiving treatment is associated with the region of residence and decreases if the student works outside the home, goes to daycare or preschool, or studies in a public school (Appendix D). As expected, students living in the North and Midwest regions are more likely to enter the federal school through the lottery draw, as these regions most frequently adopt this type of selection.[9] The other characteristics do not influence how students are admitted, generally confirming that the choice of the selection process is intrinsic to the school (each institution has administrative autonomy, Law No. 11,892, passed on December 29, 2008) and exogenous in relation to students. The second step, shown in Table 3, is the estimation of ATT.

FE students who were admitted through different types of selection processes, generally showed no difference in performance in Portuguese, with the exception of the PSM estimation with kernel and radius weight matrices, which demonstrates that students who enter school through lottery draws have lower performance in Portuguese. Contrastingly, the results for the mathematics subject were robust to all specifications. With the estimation of the effect between 0.14sd and 0.40sd less in grades, this is equivalent to saying that students who enter school through lottery draws have lower math grades than those who enter through merit-based admission. Initially, the OLS estimation shows that students who

9 The Federal Institute of Acre is located in the North region, and the Federal Institute of Brasília is located in the Midwest region.

Table 4
Results of the treatment effect on treated using CEM and entropy balancing.
Source: Prepared by the authors.

	CEM (a)	Entropy (b)	CEM/Entropy (c)	CEM/PSM (d)
Portuguese	−0.12* (0.068)	−0.07 (0.043)	−0.13*** (0.047)	−0.18*** (0.077)
Mathematics	−0.27*** (0.071)	−0.22*** (0.046)	−0.27*** (0.050)	−0.36*** (0.086)

Note: 1. *, **, and *** represent 90%, 95%, and 99% confidence levels, respectively.

Table 5
Results for the 2017 Saeb with inclusion of the gender characteristic and effects separated by gender.
Source: Prepared by the authors.

	Inclusion of gender			Girls			Boys		
	IPWRA	CEM	CEM/Entropy	IPWRA	CEM	CEM/Entropy	IPWRA	CEM	CEM/Entropy
	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)
LP	−0.10 (0.112)	−0.11 (0.087)	−0.10 (0.067)	−0.14 (0.105)	−0.17* (0.098)	−0.20** (0.087)	−0.12 (0.127)	−0.02 (0.137)	−0.05 (0.094)
MT	−0.34*** (0.100)	−0.32*** (0.077)	−0.27*** (0.076)	−0.23*** (0.073)	−0.24*** (0.077)	−0.27*** (0.093)	−0.49*** (0.176)	−0.38*** (0.112)	−0.34*** (0.119)

Note: 1. *, **, and *** represent 90%, 95%, and 99% confidence levels, respectively. 2. LP = Portuguese; MT = mathematics.

enter school through lottery draws have 0.24 fewer standard deviations in the mathematics subject in the Saeb exam than students who enter through the selection process based on grades. The effect is maintained or higher when PSM estimation is used (Columns (b) to (g) of Table 3). Only in Column (h), which considers a doubly robust approach and advances when considering robust and clustered standard errors per class, a process that is not possible in PSM, does the magnitude of the effect decrease, and the ATT becomes −0.14sd.

The other matching methods, CEM and entropy balancing, are presented in Columns (a) and (b) of Table 4. In addition to the effects found with the PSM specifications, the effect in mathematics remains statistically significant and of a similar magnitude (−0.22 and −0.27 standard deviations), whereas in Portuguese, it does not present statistical significance with entropy balancing. Furthermore, according to Iacus et al. (2009), CEM can be used in conjunction with other methods to improve the reliability of results. In Columns (c) and (d) of Table 4, the coefficients generated by CEM+Entropy and CEM+PSM reinforce the results. As a complementary contribution, a statistically significant exists inference for the Portuguese subject (−0.13 and −0.18 standard deviations).

However, although with a smaller magnitude in the IPWRA estimation, the mathematical outcomes are robust to different specifications and similar in terms of the effect direction. The parameters corroborate previous literature on the selection effect (Duflo et al., 2011; Hanushek and Wözmann, 2006; Lara et al., 2011) with regard to the existence of some form of tracking or choice by skill; to an extent, the superior performance of certain schools is a result of the students' own selection.

According to Lubienski and Lubienski (2006), the fact that there is an effect in mathematics and less robustness in Portuguese is often related to how subjects are taught. Languages and reading may, in part, reflect the family context, which, in this study, is supposed to be balanced between groups, and the effects on mathematics are the result of teaching within the school. In addition, specific literature differentiates language and mathematics skills according to gender (Contini et al., 2017; Gevrek et al., 2020). Girls are more skilled in Portuguese and boys are more skilled in mathematics. To ensure that there was no gender confusion in the results, as the 2019 Saeb does not distinguish students in terms of gender and the characteristics were thus not included in the study, Table 5 presents the results for 2017.

The validity of the results after controlling for gender characteristics is confirmed in Columns (a), (b), and (c) of Table 5. The effect remained robust for mathematics and there was no treatment effect on the treated Portuguese. When analyzing the separate effect for girls and boys (columns (d) to (h)), as expected, in 2017, the self-declared female students showed a statistical difference in LP (−0.17 and −0.20 standard deviations), and there was no effect among the self-declared male students. Complementarily, in mathematics, the ATT was higher among boys (−0.34 to −0.49 standard deviations) when compared to the result for girls (−0.23 to −0.27 standard deviations).

4.1. Sensitivity analysis and entropy balancing adjustment

As PSM hypothesizes the independence of the treatment conditional on observable characteristics, a sensitivity measure was used to verify whether the estimation is reliable. Although testing this hypothesis is not possible, the use of Rosenbaum bounds makes it possible to verify the magnitude of the unobservable effects to generate a bias in the matching estimates. The higher the critical value of Γ, the less sensitive the outcomes are to the unobservable characteristics affecting the treatment, and the closer it is to unity, the greater the chance.

Table 6
Results by quantiles of grade distribution.
Source: Prepared by the authors.

	Q10	Q25	Q50	Q75	Q90
IPWRA					
Portuguese	−0.01 (0.15)	0.14*** (0.054)	0.04 (0.046)	−0.02 (0.040)	0.06 (0.490)
Mathematics	−0.25** (0.133)	−0.31*** (0.114)	−0.11** (0.059)	−0.13*** (0.519)	0.21*** (0.065)
CEM					
Portuguese	−0.13 (0.128)	−0.03 (0.080)	0.02 (0.049)	−0.13*** (0.043)	0.02 (0.040)
Mathematics	−0.32*** (0.119)	−0.32*** (0.072)	−0.27*** (0.083)	−0.21*** (0.044)	−0.15*** (0.056)

Note: 1. *, ** and ***, represent 90%, 95% and, 99% confidence levels, respectively. 2. Robust standard errors per class.

The sensitivity analysis calculated the difference in the confidence intervals of the means between the treatment and control groups. Appendix E presents the critical test values for the estimates that showed statistical significance. For Portuguese, with the CEM+PSM specification, the critical value of Γ was 1.3. That is, the unobservable factors must increase the chances of treatment by 1.3 times so that the treatment effect is not significant. For mathematics, the critical values are even higher: 1.5 and 1.7 via PSM and CEM + PSM, respectively. The further away from the unit, the smaller the chance that unobserved effects will affect the results. This estimate corroborates with the robustness of the selection effect found in this study.

Specifically, for entropy balancing, which always reweights the data sample, the adjustment must be verified in the variance and asymmetry of the distribution in addition to the mean. Appendix F demonstrates that after entropy, the sample becomes precisely balanced.

4.2. Mechanisms

To understand the factors that contribute to the reduction in the Saeb score among students who enter school through lottery draws, this section discusses and tests the possible paths that led to this result. First, empirical literature emphasizes the possibility of different effects between grade quantiles (Cummins, 2017; Epple et al., 2002). The quantile regression specification (Table 6) allows the use of a doubly robust approach (IPWRA) in conjunction with CEM. In both specifications, mathematics has a greater effect on the lower distribution of grades (quantiles 10 and 25), and it loses magnitude in higher quantiles. Thus, the students with lower performance showed the greatest difference between the two selection modes.

In the existence of different effects between grades and greater differences between students in the lower quantiles, the second possibility to be tested is the existence of a peer effect. If high-performing students contribute to the learning of lower-performing students, it is convenient for students from classes that enter school randomly (given the lottery draws) to have fewer high-ability peers, and consequently, to lose from low-ability peers. That is, classes with students selected on the basis of grades tend to have a higher average ability and are more homogeneous, and classes comprising students selected through lottery draws tend to be more heterogeneous. In this case, a positive relationship likely exists between the average performance of a class and students' grades. An alternative specification[10] is tested to check for peer effects (Table 7).

The estimated peer effects parameter confirmed the influence of peers on students' grades. In terms of the magnitude of the association, the other students in the class contribute approximately 0.70sd in the grades of the two subjects. The ATT coefficients of both subjects also lost magnitude and, in some cases, significance, which reinforces the strong influence of peers. The inclusion of alternative peer effect control corroborates one of the main mechanisms of grade differences for the assignment of students by performance. According to Carrell et al. (2011), Cummins (2017), and Epple et al. (2002), in the existence of different grades between classes with tracking, some results are explained by the peer effect.

Another explanation for the differences in grades is school choice (student self-selection for schools) (Allen, 2007; Ivaniushina et al., 2018; Kim, 2018; Piopiunik, 2014). However, in the design of this study's empirical strategy, all students have already made their choice of school; one of the main sources of bias for omitted variables when calculating the school effect concerns the unobservable characteristics of the cohort in the region (Angrist et al., 2006; Baude et al., 2020; Benevides and Soares, 2020; Chudgar and Quin, 2012; Hanushek et al., 2007; Howell et al., 2002; Hoxby and Rockoff, 2004; Lubienski and Lubienski, 2006; McEwan, 2004; Sapelli and Vial, 2002; Somers et al., 2004). To investigate whether there

[10] The most traditional way to capture the peers effect is through the linear-in-means model. Formally, it follows the specification $y_i = \alpha + \beta_1 \bar{y}_{-i} + \gamma X_i + \rho \overline{X}_{-1} + \varepsilon_i$. This means that the outcome variable y is a function of the observable characteristics of student i, the average grade of the class without the specific grade of the student, $\bar{y}_{-i}$, and the average characteristics of the class (Sacerdote, 2011).

Table 7
Average treatment effect on the treated and peer effects.
Source: Prepared by the authors.

	IPWRA	CEM	CEM/Entropy
Portuguese	0.05*	−0.02	−0.05*
	(0.021)	(0.023)	(0.029)
Peer effects	0.70***	0.77***	0.64***
	(0.048)	(0.022)	(0.051)
Mathematics	−0.04	−0.05**	−0.09***
	(0.032)	(0.023)	(0.028)
Peer effects	0.62***	0.77***	0.62***
	(0.068)	(0.020)	(0.058)

Note: 1. *, ** and ***, represent 90%, 95% and, 99% confidence levels, respectively. 2. Robust standard errors per class.

Table 8
Average treatment effects controlling for the average grade of ninth-grade students in the municipality in the years 2013 and 2015.
Source: Prepared by the authors.

	IPWRA	CEM	CEM/Entropy
Portuguese	0.05	−0.11	−0.11*
	(0.048)	(0.075)	(0.064)
Average 9th grade	0.0004	0.01***	0.01*
	(0.002)	(0.001)	(0.004)
Mathematics	−0.10*	−0.25***	−0.24***
	(0.058)	(0.077)	(0.069)
Average 9th grade	0.01**	0.01***	0.005
	(0.002)	(0.001)	(0.004)

Note: 1. *, ** and ***, represent 90%, 95% and, 99% confidence levels, respectively. 2. Robust standard errors per class.

Table 9
Average treatment effect for the treated on failure and dropout variables.
Source: Prepared by the authors.

	IPWRA	CEM	CEM/Entropy
Failure	−0.02	−0.01	−0.01
	(0.021)	(0.03)	(0.02)
Dropout	0.01	0.02*	0.02*
	(0.009)	(0.01)	(0.01)

Note: 1. * 90% confidence level. 2. Robust standard errors per class.

exists any type of effect intrinsic to students in the region or in relation to the average quality of the cohort of students, Table 8 adds the average grade of ninth-grade students in the municipality corresponding to each student for 2013 and 2015 as a control. These groups were cohorts of high school students in 2017 and 2019.

Although it had a statistically significant relationship with students' grades, the average grade of the ninth-grade cohort had a magnitude of 0.01sd and did not affect the ATT of interest. The effect on math remains robust; students who enter through lottery draws have lower performance.

In terms of the unobservable variable, whether treatment influences the incidence of failure and dropout among students can be verified. Both characteristics are controlled for in the estimates; however, the direct effect (Table 9) reinforces that the two groups are extremely close in terms of unobservable characteristics.

Finally, we return to the discussion regarding the analysis centered on finishers, then called winners. The idea is that the Saeb is a low-risk low-stakes exam; thus, the use of a high-risk high-stakes exam, such as the National High School Exam (ENEM), can change the conduct of students during the test, as it is used in part or in full for admission to higher education in several teaching units. In this case, to the extent that student effort can vary, if the difference in results is maintained in ENEM (Table 10), there exist indications of the effects of the selection type in stages after high school, that is, in entering higher education and, consequently, in the labor market as well as in wages.

The ENEM grades were divided into five blocks and assigned independently to each block. As a result, the effect on the mathematics subject remains robust to different specifications, and a difference also exists in candidates' essay grades.

Table 10
Average treatment effect for the treated on ENEM scores.
Source: Prepared by the authors.

	IPWRA	CEM	CEM/Entropy
Natural sciences	−9.66*	−6.46	−8.45
	(4.990)	(4.654)	(5.266)
Human sciences	−19.99***	−16.35***	−7.27
	(4.786)	(3.781)	(5.152)
Language and codes	−11.05***	−10.77***	−5.39
	(3.626)	(3.053)	(3.863)
Mathematics	−27.46***	−22.39***	−35.50***
	(5.662)	(6.426)	(6.995)
Essay	−47.66***	−48.86***	−33.61***
	(8.327)	(7.847)	(10.934)

Note: 1. *, **, and *** represent 90%, 95%, and 99% confidence levels, respectively. 2. Robust standard errors per class. 3. The control variables were gender, color/race, parents' education, income, whether the student lived alone, whether the student had a computer, socioeconomic level, region of residence, and information available in the ENEM questionnaire.

5. Final considerations

The different grades achieved on standardized exams by educational institutions are normally seen as a measure of effectiveness. Students entering school are rarely viewed directly as part of the set of grades provided by the school. This study sought to provide empirical evidence on how the selection process affects the grades of Brazilian students in an innovative way, and is based on an empirical strategy that explores how students enter the federal education schools.

The results indicate that a difference exists in math grades from 0.14sd to 0.40sd when students enter FE through lottery draws. In terms of the magnitude of the effect, the total standard deviation of the scores in the sample is 0.94, which represents a reduction in dispersion between 14% and 42%. Additionally, the average grades of students from 2017 to 2019 increased by 0.20sd, which indicates a considerable magnitude, as the size of the impact is equivalent to one to four years of grade increase.

The effects were distinct between student genders and across the distribution of grades. The magnitude of the impact coefficient on mathematics was greater among self-declared male than among self-declared female students. In addition, among the girls who entered through lottery draws, there were also indications of differences in their performance in Portuguese. The selection effect was greater among students who fell into the lower end of the grade distribution.

One possible explanation for this selection effect is the gain in terms of peers. As the average performance of the class increased, there were indications of a positive association between students' grades in Portuguese and mathematics. In addition, the selection effect can be reflected in the grade achieved for admission to higher education, such as ENEM, and can consequently impact on students' lives.

The results were robust to a set of specifications, and the tests showed a low probability of unobserved effects, generating bias in the estimates.

This work has some limitations, as the results refer to a specific profile of students – those in the last year of high school in the federal education schools – which results in low external validity. Nonetheless, the evidence emphasizes the importance of how students are admitted, which should be taken into account in public educational policies that consider how students are assigned to schools and classes. Ultimately, the work contributes to the Brazilian educational context and inspires future research on the subject.

Funding

This work was supported by the Higher Education Personnel Improvement Coordination - Brazil (CAPES) Code 001. Lauana Rossetto Lazaretti was granted by Grant 88887.178904/2018-00. Marco Tulio Aniceto França was granted by National Council for Scientific and Technological Development – CNPq under Grant Process Number 312144-2019-9.

Appendix A. KMO and Bartlett tests

Kaiser-Meyer-Olkin measure of sampling adequacy	0.7929	
Bartlett's test of sphericity	Chi-square	$1.8e^{+4}$
	Sig.	0.0000

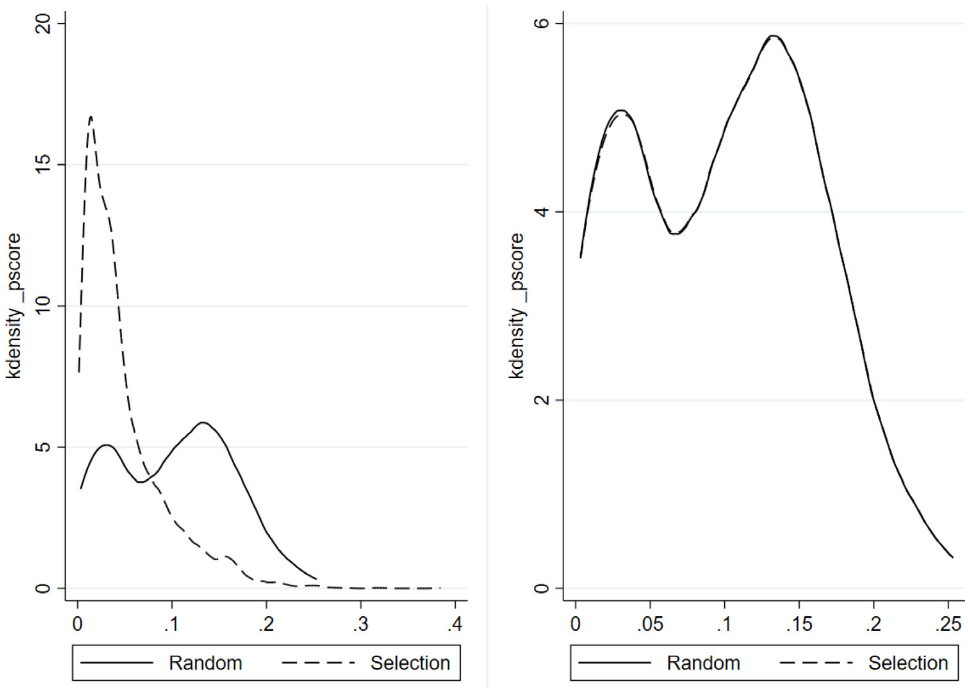
Source: Prepared by the authors.

Appendix B. Pseudo R2 test, maximum likelihood, medium bias, and median bias

Sample	Pseudo R2	p value χ^2	Medium bias	Median bias
Unmatched	0.095	0.000	27.4	19.2
Matched	0.010	0.995	2.9	2.2

Source: Prepared by the authors.

Appendix C. Distribution of propensity score for treated and untreated with a closest neighbor



Appendix D. Estimated results for the probability of receiving treatment

Covariate	Estimated parameter
Northeast region (base = North)	–
Southeast region	–
South region	−0.25***
	(0.091)
Midwest region	−0.05
	(0.071)
Afternoon shift (base = morning)	0.05
	(0.098)
Caucasian	−0.17
	(0.149)
Black	0.19
	(0.166)
Multiracial	0.13
	(0.146)

Asian	0.32
	(0.218)
Mother completed high school	−0.12
	(0.076)
Mother completed higher education	−0.07
	(0.081)
Father completed high school	−0.09
	(0.073)
Father completed higher education	0.03
	0.087
Works outside the home	−0.29***
	0.095
Went to daycare or preschool	−0.24***
	(0.063)
Public education	−0.39***
	(0.086)
EJA	0.21
	(0.132)
Lives with both parents	−0.01
	(0.063)
Failure	−0.085
	(0.085)
Dropout	0.22
	(0.149)
Parents are encouraging	−0.01
	(0.093)
Computer	−0.06
	(0.071)
Socioeconomic level	0.03
	(0.036)
R^2	0.1021

Source: Prepared by the authors.
Note: 1. *, ** and *** represent 90%, 95% and 99% confidence, respectively.

Appendix E. Sensitivity analysis critical values (Rosenbaum limits)

		Γ	p-critical
Portuguese	CEM + PSM	1	0.001557
		1.1	0.011663
		1.2	0.050047
		1.3	0.141531
Mathematics	PSM	1	0.000054
		1.1	0.000784
		1.2	0.00589
		1.3	0.026721
		1.4	0.082211
		1.5	0.187152
	CEM + PSM	1	<0.00001
		1.1	0.00001
		1.2	0.000133
		1.3	0.001022
		1.4	0.005117
		1.5	0.018278
		1.6	0.04982
		1.7	0.109157

Source: Prepared by the authors.

Appendix F. Adjustment of means, variance, and asymmetry for before and after entropy balancing

Variable	Before: weightless					
	Treatment			Control		
	Average	Variance	Asymmetry	Average	Variance	Asymmetry
Color/Race						
Caucasian	0.2567	0.1914	1.114	0.393	0.2386	0.4383
Black	0.13	0.1135	2.2	0.1115	0.09909	2.468
Multiracial	0.53	0.2499	−0.1202	0.4274	0.2448	0.2933
Asian	0.04	0.03853	4.695	0.0197	0.01931	6.913
Indigenous	0.0067	0.0066	12.12	0.0078	0.0078	11.15
Mother's education						
High school	0.3	0.2107	0.8729	0.366	0.2321	0.5562
Higher education	0.3333	0.223	0.7071	0.3014	0.2106	0.8659
Father's education						
High school	0.26	0.193	1.094	0.3242	0.2191	0.751
Higher education	0.2	0.1605	1.5	0.1595	0.1341	1.86
Works outside the home	0.08	0.07385	3.096	0.1906	0.1543	1.576
Went to daycare or preschool	0.5667	0.2464	−0.2691	0.8127	0.1522	−1.603
Public education	0.83	0.1416	−1.757	0.7889	0.1665	−1.416
EJA	0.3567	0.2302	0.5985	0.6338	0.2321	−0.5553
Lives with both parents	0.5867	0.2433	−0.352	0.6502	0.2275	−0.6299
Parents are encouraging	0.89	0.09823	−2.493	0.8154	0.1505	−1.626
Computer	0.69	0.2146	−0.8216	0.769	0.1776	−1.277
Socioeconomic level	0.0452	0.9755	0.4132	0.6486	0.2279	−0.6226
Variable	After: with weight					
	Treatment			Control		
	Average	Variance	Variable	Average	Variance	Variable
Color/Race						
Caucasian	0.2567	0.1914	1.114	0.2565	0.1907	1.115
Black	0.13	0.1135	2.2	0.1301	0.1132	2.199
Multiracial	0.53	0.2499	−12.02	0.53	0.2491	−0.1202
Asian	0.04	0.03853	4.695	0.04	0.0384	4.695
Indigenous	0.0067	0.0066	12.12	0.0067	0.0066	12.13
Mother's education						
High school	0.3	0.2107	0.8729	0.3	0.21	0.8727
Higher education	0.3333	0.223	0.7071	0.3332	0.2222	0.7078
Father's education						
High school	0.26	0.193	1.094	0.26	0.1924	1.094
Higher education	0.2	0.1605	1.5	0.1999	0.1599	1.501
Works outside the home	0.08	0.07385	3.096	0.08001	0.07361	3.096
Went to daycare or preschool	0.5667	0.2464	−0.2691	0.5672	0.2455	−0.2714
Public education	0.83	0.1416	−1.757	0.8298	0.1413	−1.755
EJA	0.3567	0.2302	0.5985	0.3567	0.2295	0.5985
Lives with both parents	0.5867	0.2433	−0.352	0.5867	0.2425	−0.3522
Parents are encouraging	0.89	0.09823	−2.493	0.89	0.09792	−2.493
Computer	0.69	0.2146	−0.8216	0.6899	0.214	−0.821
Socioeconomic level	0.0452	0.9755	0.4132	0.445	0.9422	0.2942

Source: Prepared by the authors.

References

Allen, R., 2007. Allocating pupils to their nearest secondary school: The consequences for social and ability stratification. Urban Stud. 44, 751–770. http://dx.doi.org/10.1080/00420980601184737.

Angrist, J., Bettinger, E., Kremer, M., 2006. Long-term educational consequences of secondary school vouchers: Evidence from administrative records in Colombia. Am. Econ. Rev. 96, 847–862. http://dx.doi.org/10.1257/aer.96.3.847.

Baude, P.L., Casey, M., Hanushek, E.A., Phelan, G.R., Rivkin, S.G., 2020. The evolution of charter school quality. Economica 87, 158–189. http://dx.doi.org/10.1111/ecca.12299.

Benevides, A. de A, Soares, R.B., 2020. Diferencial de desempenho de alunos das escolas militares: o caso das escolas públicas do Ceará. Nova Economia 30, 317–343.

Blackwell, M., Iacus, S., King, G., Porro, G., 2009. Cem : Coarsened exact matching in Stata. Stata J. 9, 524–546.

Caliendo, M., Kopeinig, S., 2008. Some practical guidance for the implementation of propensity score matching. J. Econom. Surveys 22 (1), 31–72.

Carrell, S.E., Hoekstra, M., West, J.E., 2011. Does drinking impair college performance? Evidence from a regression discontinuity approach. J. Public Econ. 95, 54–62. http://dx.doi.org/10.1016/j.jpubeco.2010.08.008.

Chudgar, A., Quin, E., 2012. Relationship between private schooling and achievement: Results from rural and urban India. Econ. Educ. Rev. 31, 376–390. http://dx.doi.org/10.1016/j.econedurev.2011.12.003.

Contini, D., Tommaso, M.L.Di, Mendolia, S., 2017. The gender gap in mathematics achievement: Evidence from Italian data. Econ. Educ. Rev. 58, 32–42. http://dx.doi.org/10.1016/j.econedurev.2017.03.001.

Cummins, J.R., 2017. Heterogeneous treatment effects in the low track: Revisiting the Kenyan primary school experiment. Econ. Educ. Rev. 56, 40–51. http://dx.doi.org/10.1016/j.econedurev.2016.11.006.

Duflo, E., Dupas, P., Kremer, M., 2011. Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in Kenya. Am. Econ. Rev. 101, 1739–1774. http://dx.doi.org/10.1257/aer.101.5.1739.

Emsley, R., Lunt, M., Pickles, A., 2008. Implementing double-robust estimators of causal effects. pp. 334–353.

Epple, D., Newlon, E., Romano, R., 2002. Ability tracking, school competition, and the distribution of educational benefits. J. Public Econ. 83, 1–48. http://dx.doi.org/10.1016/S0047-2727(00)00175-4.

Gevrek, Z.E., Gevrek, D., Neumeier, C., 2020. Explaining the gender gaps in mathematics achievement and attitudes: The role of societal gender equality. Econ. Educ. Rev. 76, 101978. http://dx.doi.org/10.1016/j.econedurev.2020.101978.

Hainmueller, J., 2012. Entropy balancing for causal effects : A multivariate reweighting method to produce balanced samples in observational studies. Polit. Anal. 20, 25–46. http://dx.doi.org/10.1093/pan/mpr025.

Hanushek, E.A., Kain, J.F., Rivkin, S.G., Branch, G.F., 2007. Charter school quality and parental decision making with school choice. J. Public Econ. 91, 823–848. http://dx.doi.org/10.1016/j.jpubeco.2006.09.014.

Hanushek, E.A., Wözmann, L., 2006. Does educational tracking affect performance and inequality? differences-in-differences evidence across countries. Econ. J. 116, 63–76. http://dx.doi.org/10.1111/j.1468-0297.2006.01076.x.

Howell, W.G., Wolf, P.J., Campbell, D.E., Peterson, P.E., 2002. School vouchers and academic performance: Results from three randomized field trials. J. Policy Anal. Manag. 21, 191–217. http://dx.doi.org/10.1002/pam.10023.

Hoxby, C.M., Rockoff, J.E., 2004. The impact of charter schools on student achievement. http://dx.doi.org/10.4324/9781351213318.

Iacus, S.M., King, G., 2008. Matching for causal inference without balance checking.

Iacus, S.M., King, G., Porro, G., 2009. Cem : Software for coarsened exact matching. J. Stat. Softw. 30.

Iacus, S.M., King, G., Porro, G., 2011. Causal inference without balance checking : Coarsened exact matching. Polit. Anal. August, http://dx.doi.org/10.1093/pan/mpr013.

Ivaniushina, V., Makles, A.M., Schneider, K., Alexandrov, D., 2018. School segregation in St. Petersburg–the role of socioeconomic status. Educ. Econ. 27, 1–21. http://dx.doi.org/10.1080/09645292.2018.1538408.

Kim, Y., 2018. The effects of school choice on achievement gaps between private and public high schools: Evidence from the seoul high school choice program. Int. J. Educ. Dev. 60, 25–32. http://dx.doi.org/10.1016/j.ijedudev.2017.10.019.

Lara, B., Mizala, A., Repetto, A., 2011. The effectiveness of private voucher education: Evidence from structural school switches. Educ. Eval. Policy Anal. 33, 119–137. http://dx.doi.org/10.3102/0162373711402990.

Lazear, E.P., 2001. Educational production. Q. J. Econ. 116, 777–803.

Lubienski, S.T., Lubienski, C., 2006. School sector and academic achievement: A multilevel analysis of NAEP mathematics data. Am. Educ. Res. J. 43, 651–698.

Lunceford, J.K., Davidian, M., 2004. Stratiÿcation and weighting via the propensity score in estimation of causal treatment e ects : a comparative study. Stat. Med. 23, 2937–2960. http://dx.doi.org/10.1002/sim.1903.

McEwan, P.J., 2004. The potential impact of vouchers. Peabody J. Educ. 79, 57–80. http://dx.doi.org/10.1207/s15327930pje7903_4.

MEC, Ministério da Educação. 2021. [www document] URL http://portal.mec.gov.br/setec-programas-e-acoes/expansao-da-rede-federal.

Piopiunik, M., 2014. The effects of early tracking on student performance: Evidence from a school reform in Bavaria. Econ. Educ. Rev. 42, 12–33. http://dx.doi.org/10.1016/j.econedurev.2014.06.002.

PNP, 2020. Plataforma Nilo Peçanha [www document]. URL http://plataformanilopecanha.mec.gov.br/.

Rosenbaum, P.R., 2002. Observational Studies, second ed. In: Springer Series in Statistics.

Rosenbaum, B.Y.P.R., Rubin, D.B., 1983. The central role of the propensity score in observational studies for causal effects. Biometrika 70, 41–55.

Sacerdote, B., 2011. Peer effects in education: How might they work, how big are they and how much do we know thus far? In: HandBook of the Economics of Education. Elsevier B.V., pp. 249–277. http://dx.doi.org/10.1016/B978-0-444-53429-3.00004-1.

Sapelli, C., Vial, B., 2002. The performance of private and public schools in the chilean voucher system. Cuad. Econ. 42, 3–454.

SIMCAQ, 2021. Simulador de custo aluno qualidade [www document]. URL https://simcaq.c3sl.ufpr.br/simulator/results/caqreport.

Somers, M.A., McEwan, P.J., Willms, J.D., 2004. How effective are private schools in Latin America? Comp. Educ. Rev. 48, 48–69. http://dx.doi.org/10.2307/3542000.