The Oscars of Education: The effects of a Tournament Between Public Schools

Casemiro Campos*

Goethe University Frankfurt

July 27, 2023

Abstract

Despite discussions and implementations of performance-related incentive pay for teachers, limited research exists on the medium-term effects of combining financial and non-pecuniary group incentives within the public school system. This study focuses on a yearly rank-order tournament among schools in Brazil and presents quasi-experimental evidence. By exploring the timing and eligibility criteria of the tournament, a differences-in-discontinuities design is implemented and reveals substantial improvements in student performance by their school being eligible for the prize. These effects primarily stem from schools at both ends of the performance spectrum, indicating the varying levels of incentives generated by the tournament. Furthermore, by exploring an exogenous cutoff that defines the prize winners, I document significant ex-post changes in the beliefs and behavior of teachers and students upon receiving the prize. Awarded teachers tend to implement innovative teaching methods, and students report higher learning efforts and performance.

1 Introduction

Incentive designs are widely adopted by firms as a mechanism to induce effort and behavior in employees [Lazear (2018)]. Over the last few years, there has been an increasing effort in rationalizing these schemes for the public sector, including the public education system. Particularly relevant for low-income and developing countries, individual incentive mechanisms, such as teacher performance pay, have been used as a tool to enhance the productivity of public schools, as a substitution for input-based policies. However, empirical analyses of such

*PhD Candidate at Goethe University Frankfurt. Contact: casemirocampos@em.uni-frankfurt.de

programs have produced mixed evidence in terms of concrete results (see Santibañez (2020) for a comprehensive review).

On the other hand, more complex mechanisms that involve both financial and non-pecuniary team incentives are still overlooked by the literature. Recent advances in the Personnel Economics literature highlight the positive reinforcement of the use of incentive schemes at the team level. That is, implementing an incentives mechanism targeted to a group of workers can achieve further productivity gains as compared to individual incentives, once that peer effects can interact with incentives and yield further efficiency gains, whether through monitoring, coaching, or motivation [Friebel et al. (2017); Lavy (2009); Chevalier (2022)]. On top of that, non-financial and implicit incentives also have the potential to enhance group identity, especially considering the intrinsic characteristics of the job as an educator [Dixit (2002); Akerlof and Kranton (2005); Dang and King (2016)].

This paper aims to contribute to the empirical literature on incentive designs in the education sector, by providing quasi-experimental evidence on the introduction of a tournament between public schools that combines both financial and non-pecuniary group incentives. Informally denominated the "The Oscars of Education", the School Score A Award (*Prêmio Escola Nota Dez* in Portuguese, hence PEN10), yearly awards the top performing public schools from the Brazilian federal state of Ceará. On top of that, the program offers yearly financial help to the worst-performing schools.

This study is divided in two main set of results. First, by exploring exogenous variations on the prize eligibility criteria, this paper documents large gains in student's academic performance from the schools competing for the prize, denominated here the tournament effect. Furthermore, by again exploring an exogenous variation, now on the prize winning criteria, this paper also documents significant ex-post effects within the winners of the PEN10, here called the prize effect. Being a prize winner changes teacher's and principal's perception on their job, and increases the evaluated and reported effort by the students.

To estimate the causal tournament effects of the award, my research design exploit two dimensions of the prize: the introduction timing and the eligibility criteria. More specifically, I will combine the before/after dimension of my data set with the discontinuity of the prize eligibility criteria (only schools with at least 20 students enrolled at the evaluated grade are eligible) to implement a diff-in-disc design. The introduction of the policy had a local increase in the scores of 0.25 and 0.35 standard deviations for Language and Mathematics, respectively.

Furthermore, the estimated effects are related to the school's relative performance: for Mathematics, schools at the top and bottom of the prior scores distribution observed the highest positive impact on their performance; for Language, on the other hand, the prize only had an impact on schools at the bottom half of the distribution.

To estimate the prize effect, I will explore the rank of the worst-performing winner as a cutoff for a regression discontinuity design. That is, using a biannual public school survey from the Brazilian federal government (SAEB), I will refrain my analysis to the schools located just above and below the winning threshold on a given year and estimate the difference in my set of outcomes between the worst performing winners and the best-performing losers. Overall, I observed significant changes in the reported behavior and performance of school actors between one and two years after receiving the prize. Teachers become more confident regarding their performance and implement more innovative teaching methods. Furthermore, students report an increase in academic effort, followed by a higher performance in the SAEB evaluation exam.

Overall, this study highlights how policymakers can explore mechanisms beyond teacher performance pay and individual compensations. The combination of financial compensations and non-pecuniary incentives in a setting that fosters cooperation between schools yields an increase in learning by the students through two different channels, by competing for the prize and by winning it.

This study contributes to the growing literature on incentives programs in the education sector. Primarily focused on low-income and developing countries, previous evidence has documented the relevance of teacher performance pay mechanisms to induce effort and enhance the productivity of teachers in developing countries [Muralidharan and Sundararaman (2011); Duflo et al. (2015); Barrera-Osorio and Raju (2017)]. Although these studies produce mixed evidence, with the most significant results coming from small-scale interventions, they still show that targeted incentives are more cost-effective than traditional salary increases.

On the other hand, little is known about the effects of targeted group incentives. Some of the previous studies mixing individual and group incentives in small-scale interventions have shown more promising results [Lavy (2002); Glewwe et al. (2003)]. Further, little is known about the effects of non-pecuniary incentives on teachers' efforts, apart from a few correlational studies [Dang and King (2016)]. In this sense, this study provides novel evidence on a large-scale permanent tournament program between public schools that combines financial and non-pecuniary group incentives in a developing country setting. Finally, this paper also documents

the presence of ex-post performance effects of awards in the public education sector, in line with the literature on ex-post effects of awards in the firms setting [Neckermann et al. (2014); Ashraf et al. (2014)].

This paper is organized as follows. In section 2 the institutional setting and the role of the incentives of the prize are described. In section 3, I proceed with the empirical analysis of the tournament effects. Further, in section 4 I describe the analysis and results for the prize effects. Section 5 concludes the paper.

2 School Score A Award

2.1 Institutional Setting

The School Score A Award (*Premio Escola Nota Dez*, in Portuguese) is a state-level policy aimed at public schools in the state of Ceará, Brazil. The main objective of the prize is to award the best-performing schools on a yearly basis. The top 150 schools in the 2nd, 5th, and 9th grades receive a financial prize. Further, the prize also gives an extra budget to 150 worst-performing schools (supported), conditional on the development of a cooperation program with a top-performing school (awarded).

First implemented in 2009, the prize was initially designed to reward the schools cohorts at the first grade, as part of an effort to increase the literacy rates within the public school system. The Right Age Literacy Program (PAIC in Portuguese), was a consequence of an internal study conducted by the state parliament, pointing out the alarming literacy rate of the state, concluding that only 16% of the children read, write and understand a text in an age-appropriate way [Ceará (2006)]. Later on, the prize was implemented in the 5th grade, in 2011, and in the 9th grade in 2015.

The prize is structured as a tournament in which, every year, the eligible public schools within the state are ranked based on a school independent evaluation conducted by a third-party entity [1]. To be eligible for the tournament, the schools needs to fulfill the following criteria:

1. To have, at the moment of the evaluation, at least 20 students enrolled at the evaluated grade.

[1] The entire evaluation process used to construct the ranking for the prize, including the analysis of the eligibility criteria, are all conducted by the Center for Public Policy and Evaluation of Education of the Federal University of Juiz de Fora (CAEd/UFJF).

2. An average between 7.5 and 10 on the School Performance Index (SPI).

3. To have at least 90% of the enrolled students present at the conducted evaluation.

In the Brazilian education system, the provision of primary and secondary education is mostly the responsibility of the state and municipal governments. In practical terms, usually, the municipalities are the ones responsible to manage primary education, or *Ensino Fundamental*, which is equivalent to elementary and middle school. On the other hand, the state government is responsible to provide education at the high school level. In this sense, the award is a state-level policy that focuses on the grades that are funded and managed by the municipalities.

2.2 The School Performance Index (SPI)

The schools are ranked based on their performance on a yearly evaluation with all schools part of the public system of the state. As part of the System for Permanent Evaluation of Education from the State of Ceará (SPAECE in Portuguese), all students in the 2nd, 5th, and 9th grades are evaluated in Mathematics and Portuguese. From these evaluations, the School Performance Index (SPI) for each grade g in school i is constructed as follows:

$$SPI_{gi} = \frac{SPI_{gi}^{Math} + SPI_{gi}^{Port}}{2} \tag{1}$$

Where for each subject s:

$$SPI_{gi}^{s} = Score_{gi}^{s} * PR_{gi}^{s} * LA_{gi}^{s} \tag{2}$$

$Score_{gi}^{s}$ is the school score at grade g and subject s between 0 and 10, the latter being the highest score possible. PR_{gi}^{s} is the Participation Rate, the fraction of enrolled students in grade g that took part in the evaluation exam. Finally, LA_{gi}^{s} is the Level Adjustment variable, which is a weighted average of the distributions of the students on each proficiency level[2].

To summarize, the SPI takes into account not only the overall performance at the exam, but also the share of students that were actually evaluated, and it prevents that schools only select the top students to be part of the exam. Further, it also considers the performance distribution within the school, by the level adjustment variable. The higher the proportion of students in

[2]This variable takes a value between 0 and 1 depending on the share of students on each level of performance. Students are assigned to 4 categories: Very Critical, Critical, Intermediate, and Adequate. For example, School A has 10% of the students at the very critical level, 5% at the critical level, 15% at the intermediate level, and 70% at the adequate level. The Level Adjustment Variable will be equal to: 0.25*10% + 0.5*5% + 0.75*15% + 1*70% = 86.25%

the lowest performance levels, the lower the value of the variable. This mechanism encourages schools to not leave the worst-performing students behind.

2.3 The Financial Prize

The 150 awarded schools are entitled to a financial prize of around 386 dollars per student, while the bottom 150 schools are entitled to an extra budget of around 193 dollars per student. To put that into perspective, according to third-party estimates, the annual average expenditure per student at a primary school in Brazil was around 722 dollars per student, in 2015 [Educação (2021)]. Hence, the prize for the awarded schools is more than half of the average expenditure per student of the Brazilian public schools.

This discretionary budget comes from the State Fund for the Fight Against Poverty (*Fundo Estadual de Combate à Pobreza (FECOP)*, in Portuguese) and is managed by the Department of Education of the State of Ceará. These transfers are divided into two terms. By the end of the year of the announcement (one year after the schools were evaluated), the awarded schools receive 75% of the prize and the supported schools receive 50% of the prize. In the following year, the remaining prize is distributed to the school's condition on a cooperation program between a rewarded and a supported school, where the rewarded school needs to maintain or raise the results from the past year, and the supported school needs to obtain a minimum grade of 5 in the SPI.

The cooperation program between rewarded and supported schools is a central part of the prize. Each awarded school is matched with a supported school by the department of education, based on the SPI ranking and the geographical proximity between them. Once the match is done, each municipality has a representative which will coordinate the partnerships of the schools and will report it to the department of education. This partnership is centered around bilateral meetings between the schools, in which technical and pedagogical elements are discussed with the final objective of enhancing student's performance.

As pointed out by the literature on the program, this set of conditions is not easily fulfilled. A recent analysis, by Costa and Vidal (2021), shows that between 2008 and 2015, from all the schools that received the prize for the literacy grade, only around 38% of the awarded and 58% of the supported schools achieved the conditions to receive the second part of the prize. Beyond these conditions, the prize establishes certain criteria to be followed in the execution of the received budget.

For the awarded schools, the first part of the financial prize needs to be pledged as follows: 20% of the total value is addressed to bonuses for teachers and staff; 70% needs to be invested in the acquisition of pedagogical materials, extension courses for teachers, tutoring, or infrastructure improvements of the schools; and the remaining 10% goes to finance the cooperation meetings between the schools. For the second part of the prize, up to 100% of the prize can go to investments within the schools, as stated by the second item, and up to 30% can go to bonuses for the school staff.

For the supported schools, the first part of the prize cannot be directed to bonuses for teachers and staff. 90% of the budget needs to be invested back into the schools, and 10% goes to finance the partnership program. On the other hand, if the supported schools managed to achieve an SPI of at least 5 in the following year, up to 30% of the remaining prize can be destined to bonuses for the staff, and up to 100% can be invested within the schools.

2.4 Ceremony: The Oscars of Education

Each year, all awarded schools are invited by the state administration to the prize ceremony at the state capital. The ceremony is usually conducted by the state governor with the presence of important political figures. Given the size of the prize, the ceremony is conducted at the biggest conference hall in the state. Beyond these ceremonies conducted by the state, some municipalities also conduct ceremonies within their jurisdiction.

The event is highly appraised and expected, as it gives an important platform for the best-performing schools to be recognized for their effort. Informally named by the community "The Oscars of Education", the ceremony is the most important event of the year for the education system and is a major motivational drive for the teachers, principals and students. Previous qualitative studies highlights how the event is expected by all entities within the schools, as it is considered to be a major spotlight for not only the schools, but the municipalities [Conceição (2022)].

It is crucial to recognize the non-monetary incentives that arise from the policy. Since its establishment, the state department of education has employed the award as a motivational tool, fostering collaboration among key stakeholders in the education system. The ceremony extends beyond school staff and involves mayors from municipalities with high-performing schools. These politicians utilize the documentation from the ceremony as a promotional instrument, often sharing it on their social media platforms.

2.5 The role of Incentives

The award triggers distinct incentive mechanisms. On the one hand, there is the monetary aspect, where the top and bottom schools receive additional funding through a competition. Since school performance is assessed as a whole, without individual distinctions among teachers, the prize is designed as a group incentive. Further, unlike traditional teacher performance pay incentives, a significant portion of the budget cannot be allocated to individual staff bonuses. As previously mentioned, there are restrictions of 20% and 30% on bonuses for the first and second parts, respectively.

As summarized in Garibaldi (2006), the tournament theory yields essentials predictions over the characteristics of such relative compensation designs which are crucial to the Grade A prize setting: first, the effort exerted by the participants depends directly on the size of the premium that comes with the performance; further, the effort will also be positively linked to the level on which marginal effort increases the chances of winning and negatively related to its marginal cost. Since there are monetary compensations for both high and low-performance schools, there is a potential heterogeneity in the levels, and the direction, of the incentives, depending on how each school perceives its relative performance.

In other words, each school will potentially have the incentive to behave accordingly to its expected position on the ranking, which could yield the second-best effort. If a certain school has been performing poorly over the last years, the perceived probability of achieving a high ranking is lower than the probability of being located at the bottom of the rank. Hence, the tournament can have adverse effects, as it could lead to a lower level of effort than the scenario without a tournament, for the school to receive the financial compensation for being in the bottom 150. Also, schools that perceive themselves as average performance schools will tend to have lower expectations in terms of receiving a financial award from the tournament.

Usually, these adverse features of a simple tournament are fixed by introducing different mechanisms. To deal with the natural advantage that some players may have, one can introduce a handicapping system by giving the player at disadvantage a head start or offer strategies within the tournament with different risk profiles [Garibaldi (2006)]. Further, it has been shown that rank-order tournaments with prizes that differs among players can also be used to solve the inefficiency caused by the heterogeneity of the players [Gürtler and Kräkel (2010)]. Since the PEN10 does not have any of these mechanisms, in a heterogeneity analysis I will test whether the effects of the prize are different depending on the previous performance levels of the schools.

3 The Tournament Effect

3.1 Data

This empirical analysis will be based on a panel data set, at the school level, provided by the system for permanent evaluation of education from the state of Ceará (SPAECE). First implemented in 1992 and administered by the department of education of the state, since 2007 the SPAECE conducts an yearly evaluation of all primary and secondary students enrolled in the public system. The evaluation is used as a guide for different educational policies of the state, including the Grade A Prize. Beyond the average scores of the schools in Language and Mathematics, the main data set also contains information on the number of enrolled and evaluated students of each school.

For this study, I will use data from 2008, the first available year, until 2019. However, there are no data available for 2009 and 2013 [3]. Given that the time dimension will be crucial in my identification strategy, I will focus only on the effects of the prize for the 9th grade. As mentioned, the prize for the 2nd and 5th grade was implemented in 2009 and 2011, respectively, hence does not give me enough pre-treatment units. On the other hand, the prize for the 9th grade was implemented in 2015, giving me a balanced data-set in terms of before and after treatment units.

As documented in table 1, the sample begins with 2808 evaluated schools but decreases constantly throughout the years. This movement, is related to an official guideline of the Department of Education in the direction of centralization of the municipal school systems, with the clustering of schools, usually located in the rural area, with more structured and bigger schools [Ceará (2012); Costa and Vidal (2021)]. The observed share of eligible schools, that is schools with 20 or more students enrolled in the 9th grade, remains fairly constant throughout the years.

Further, to complement the main data set, I collected different data sources from the Brazilian federal government. First, to add further school characteristics, yearly school-level data from the Brazilian School Census is added. Through the Census I can observe the infrastructure characteristics of the schools and so construct what I called the Infrastructure Index [4]. Furthermore, to complement my main outcomes variables, I added a yearly data set that

[3] According to the department of education, in 2009 the SPAECE exam was not applied, and in 2013 the exam was applied for only a sample of schools.

[4] The Infrastructure Index is an average of the following dummy variables: whether the school has filtered water, if it has a main auditorium, a library, a reading room, a science laboratory, a computer laboratory, a

Table 1: Number of Evaluated and Eligible Schools by Year

Year	Evaluated Schools	Eligible Schools	Eligible Schools (%)
2008	2808	1793	63.85
2010	2815	1779	63.19
2011	2709	1794	66.22
2012	2778	1781	64.11
2014	2566	1577	61.45
2015	2522	1590	63.04
2016	2516	1626	64.62
2017	2466	1565	63.46
2018	2324	1503	64.67
2019	2258	1458	64.57

reports the yearly failing and dropout rate for each school [5].

Table 2 presents the descriptive statistics of the presented variables for the sample used in the study. First, the main outcomes are standardized to have mean zero and standard deviation one, to simplify the interpretation of the results. Further, the average of students enrolled in the 9th grade is around 43 students What is also notable is the average attendance rate of the SPAECE evaluation, around 95% [6]. Finally, the average failing and dropout rate are around 3%.

Table 2: Descriptive Statistics

	Mean	SD	Min	Max	N
Score Port	0.00	1.00	-4.64	4.58	25726
Score Math	0.00	1.00	-3.63	5.68	25726
Students Enrolled at 9th Grade	43.03	45.35	0.00	574.00	25735
Students Attended Exam	40.25	41.22	0.00	491.00	25762
Exam Attendance Rate (%)	94.21	10.23	0.00	100.00	25762
Failing Rate 8th Grade (%)	2.95	5.82	0.00	83.30	25680
Dropout Rate 9th Grade (%)	3.23	6.01	0.00	100.00	25680
Rural Area	0.52	0.50	0.00	1.00	25751
Infrastructure Index	0.51	0.20	0.00	1.00	25751

playground, a sports court, internet, and at least one computer.

[5] Failing rate is defined here as the percentage of the students that did not achieve all the requirements to move on to the next grade, and so they must repeat the grade at the next school year. The dropout rate is the percentage of students that drop out of school compared to the previous year.

[6] For some schools, the number of evaluated students where higher than the number of enrolled students, but the attendance rate variable was capped at 100%.

3.2 Empirical Strategy

Here I want to estimate what the literature calls the tournament effect. In my set-up, it is the effect on the student's performance of the introduction of the award on the eligible schools. The design of the policy allows us to explore the eligibility criteria as part of the identification strategy. As discussed, there are three criteria: first, the school must have at least 20 students enrolled at the evaluated grade; it needs to have an average between 7.5 and 10 at the SPI; and it needs to have a presence rate in the evaluation of 90%. The last two criteria are endogenous to the actions of the school in the respective year of the evaluation. On the other hand, the number of enrolled students is, in theory, exogenous in the short term. I can explore the cutoff created by the first criteria as a proxy for the eligibility status.

Given that the school status sharply changes at the 20 enrolled students threshold, I could estimate the local average treatment effect of having the number of students just above the cutoff, compared to having a number just below. Hence, run a cross-sectional Regression Discontinuity Design (RDD) at the cutoff. In order to produce a causal interpretation of my estimates it is necessary that the continuity assumption holds. That is, in the absence of the tournament, there are no discontinuities at the running variable cutoff.

On a typical RDD, this discontinuities are the biggest threat to the identification strategy, once researchers cannot take into account all the unobservables that are potentially discontinuous at the cutoff. To avoid these threats, I will incorporate a time dimension in my design. Since I have access to a panel data structure, I will also include the timing of the introduction of the prize for each grade as a further treatment status dimension. The idea is to use the information on the period prior to the award to remove possible selection bias. That is, compare the outcomes before and after treatment as in a Differences-in-Differences (DiD) design.

Hence, I will use the estimated discontinuities of the cross-sectional RDD on each time t as the outcome variable in a DiD design that uses the timing of the introduction of the prize. With that, I can relax the continuity assumption, but, now, it is required that observations around the cutoff are on a local parallel trend in the absence of treatment. This assumption is easier to be tested, using the same criteria as in the conventional DiD setup. Denominated the Differences-in-Discontinuities (Diff-in-Disc) estimator, it was first proposed by Grembi et al. (2016). Since then, the framework has been extended to cases with imperfect compliance, and it has been explored as a local average treatment effect (LATE) estimation method in different scenarios [Galindo Silva et al. (2018); De Benedetto and De Paola (2019); Millán-Quijano (2020); Butts

(2023)].

In my case, I have the following assignment mechanism: first, there is the time dimension, the tournament started for a given grade at time t_0; further, we have the eligibility criteria, only schools with 20 or more enrolled students are eligible for the prize. To produce causal estimates of the tournament effect I will then explore both of these variations.

Formally, the estimations procedure consists of fitting a local linear regression to the observations within a distance h, which is the bandwidth, on each side of the cutoff S_c (20 students) before and after treatment t_0. The boundaries point, h, is an average of the before and after treatment optimally chosen bandwidth based on the mean square error (MSE) optimal bandwidth selector algorithm [Calonico et al. (2020)]. Hence, restricting the sample to schools within the interval $S_{it} \in [S_c - h, S_c + h]$ the following equation is estimated:

$$Y_{it} = \delta_0 + \delta_1 S^*_{it} + T_{it}(\gamma_0 + \gamma_1 S^*_{it}) + Post_t[\alpha_0 + \alpha_1 S^*_{it} + T_{it}(\beta_0 + \beta_1 S^*_{it})] + \delta_2 X_{it} + \theta_m + \epsilon_{it} \quad (3)$$

Where S^*_{it} is the normalized cohort size ($S^*_{it} = S_{it} - 20$) for a given school i at the year t. Further, T_{it} is the treatment status dummy, indicating if the school i has 20 or more students enrolled at time t, and $Post_t$ is a dummy indicating if $t > t_0$. My parameter of interest is then β_0, which gives us the estimated effect of the interaction between the post-period and the treatment dummy. X_{it} is a vector with school characteristics. Further, θ_m is the municipality fixed effects, which account for time-invariant municipalities characteristics that are potentially correlated to my assignment criteria. Standard errors are clustered at the school level.

3.2.1 Validity of the Differences-in-Discontinuities Strategy

In order to validate the empirical strategy, it is important to check for the continuity of the forcing variable, in my case the cohort size. If there are significant discontinuities at the eligibility cut-off point, it raises the concern that the schools are artificially raising the number of students per cohort in order to be eligible for the prize. To test for that, I conduct a data-driven manipulation test following Cattaneo et al. (2018, 2020).

This test uses local polynomial estimators to compute the density of the running variable on both sides of the cutoff. Figure 1 plots the estimated third-order polynomial on both sides of the cutoff for 2015, the year of the introduction of the prize. The test does not point out a statistically significant discontinuity at the cutoff. Figures 10 and 11 in the appendix, plots the

same test but for the years before and after the introduction of the prize, with no discontinuities at the cutoff found.

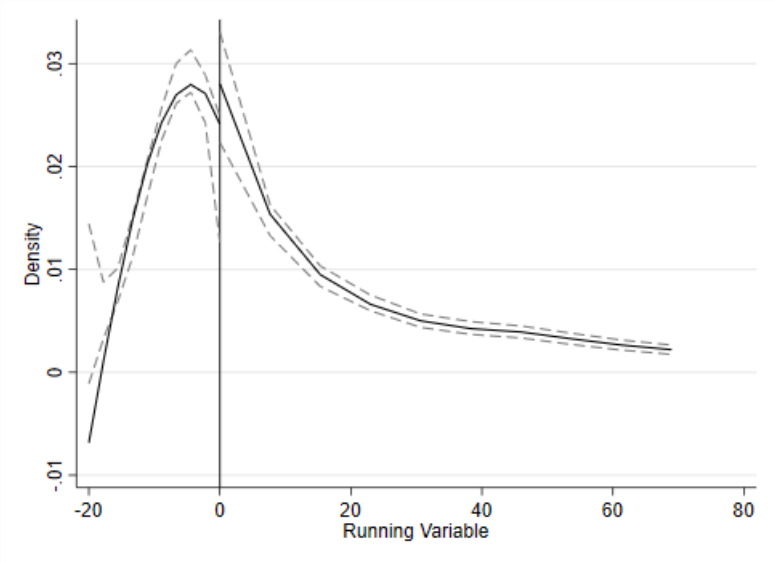
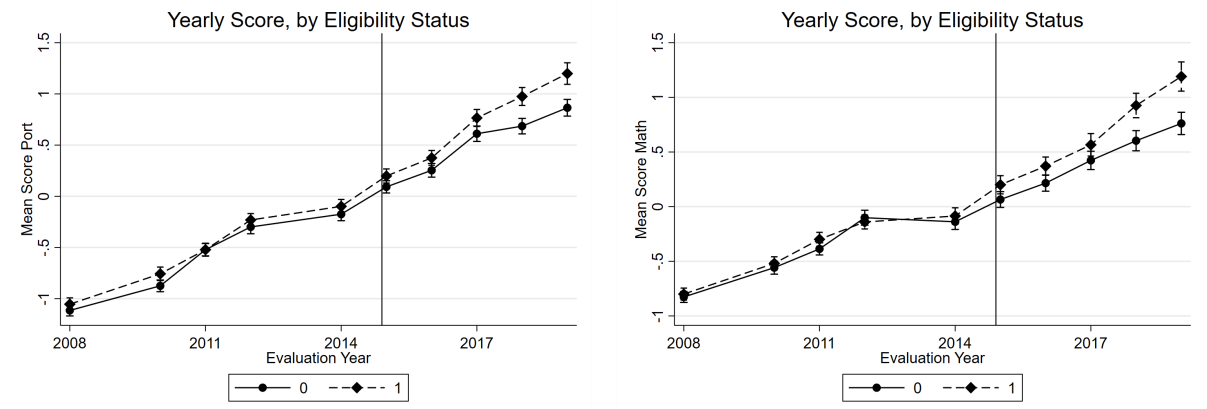


Figure 1: Schools Distribution by Cohort (9th Grade) Size - 2015

Further, I also tested for plausible threats on the parallel trends assumption, the main underlying assumption in the diff-in-disc estimator. For that, I can check whether the performance levels of the schools just above and just below the cut-off had similar trends before the announcement of the prize for the 9th grade. Figure 2 plots the average levels of performance in Portuguese and Mathematics, respectively, for schools just above and below the cutoff. It reassures us that both groups were following similar trends before 2015. In fact, both groups had quite similar performance levels. On the other hand, after 2015, the gap between both groups grows constantly.



(a) Language (b) Mathematics

Figure 2: Average Trends in Perfomance by Eligibility Status

3.3 Results

Table 3 displays the main Differences-in-Discontinuities estimates. The main outcomes are student's related to performance measures: the yearly scores in Mathematics and Language at the SPAECE evaluation. For each of these outcomes, a bandwidth is optimally chosen, and the coefficients of Eq. 3 are estimated. Column 5 and 6 displays the results from the full specification, including controls at the school level and municipalities fixed effects.

Table 3: Tournament Effects: School Grade A Award, Diff-in-Disc Estimates (9th Grade)

	(1)	(2)	(3)
Panel A: Port			
Diff-in-Disc Est	0.313***	0.311***	0.247***
	(0.070)	(0.070)	(0.062)
Bandwidth	8.8	8.8	8.8
N	10456	10455	10453
Mun FE	N	N	Y
Controls	N	Y	Y
Panel B: Math			
Diff-in-Disc Est	0.479***	0.475***	0.356***
	(0.076)	(0.076)	(0.065)
Bandwidth	9.2	9.2	9.2
N	11428	11427	11425
Mun FE	N	N	Y
Controls	N	Y	Y

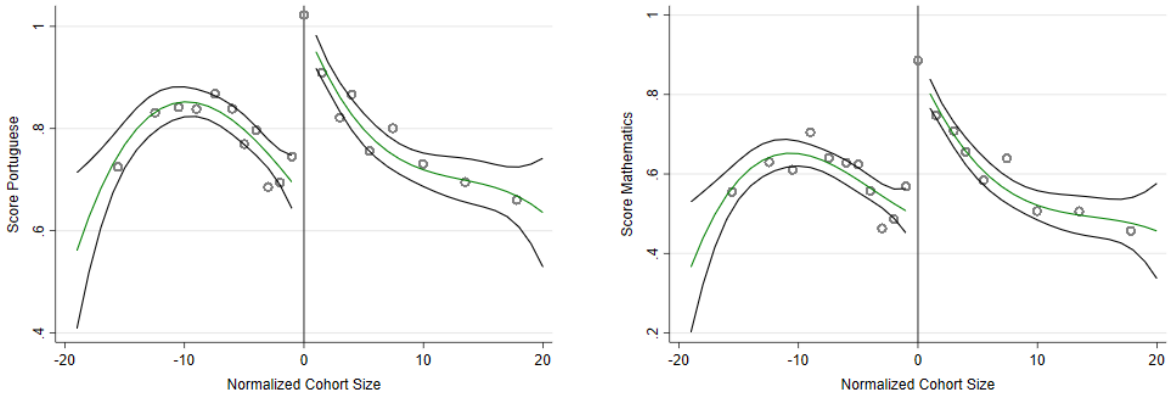
Note: Table reports Diff-in-Disc estimates. Scores are normalized. Controls includes rural area dummy and a infrastructure index variable. Parentheses contain standard errors, clustered at the school level. Significance levels are indicated by $*<.1$, $**<.05$, $***<.01$.

The introduction of the School Grade A Award had a positive and significant impact on the school's performance at the SPAECE evaluation exam, on all three specifications. To be precise, the award had a local increase in the scores of 0.25 Standard Deviation (SD) for Language and 0.35 SD for Mathematics in the full specification. Adding school characteristics do not alter these results. On the other hand, the inclusion of municipality fixed effects lowers the magnitude of the effects substantially. One possible explanation is that a school being just above the eligibility cut-off after the treatment is correlated with time-invariant characteristics of their respective municipality.

Looking into further outcomes, Table 11 in the appendix display the estimations using the dropout and failing rates, respectively, as outcomes. As one can observe, we do not find any effects of the introduction of the award on those outcomes.

These estimates are consistent with the evidence displayed in the descriptive graphs in

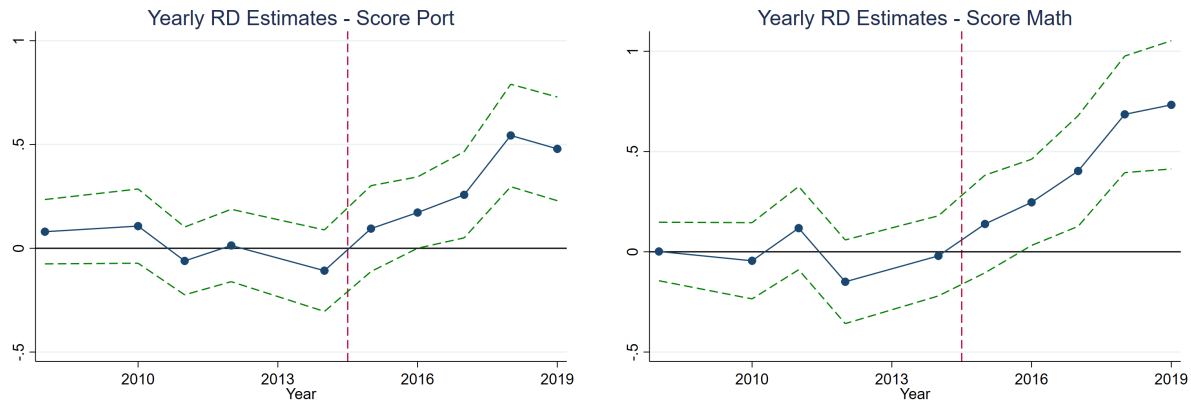
Figure 3. Following previous empirical studies [Grembi et al. (2016); De Benedetto and De Paola (2019)], the figure plots the estimated difference in scores between post and pre-treatment values by 20 quantiles of the normalized cohort size. The central line is a third-order polynomial fit, and the lateral lines are the 95% confidence intervals. Each point in the graph marks the average difference between post and pre-treatment for a given cohort size quantile. In this sense, there is a positive, and statistically significant, discontinuity in the neighborhood of the prize eligibility cutoff. Overall, schools with different cohort sizes increased their average scores comparing before and after 2015, but around the 20 students' cutoff, there is a discontinuity. Schools with a cohort size just above the cutoff had a discontinuous increase in their performance, compared to schools just below the cutoff.



(a) Language (b) Mathematics

Figure 3: Differences-in-discontinuities plots

To illustrate further these results, figure 4 plots the yearly cross-section regression discontinuity designs estimates. This plot is particularly useful to get insights on the time trends of the discontinuities around the cutoff, and also to give an additional support to the parallel trends assumption. As we can infer from the graphs, before the introduction of the award for the 9th Grade, in 2015, there were no statistically significant discontinuities in the scores at the 20 students' cutoff. That is, both groups of schools just below and above the cutoff were on parallel trends before the award. After 2015, there is an increase in the size of the discontinuities for both Mathematics and Language. These estimates contribute as supportive evidence that there are no potential unobservables that could trigger a discontinuity in the outcomes and also suggest that the award was not anticipated in the previous years.



(a) Language (b) Mathematics

Figure 4: Yearly Regression Discontinuity Estimates

3.3.1 The Effects of the Prize by Prior Performance Level

As pointed out, the productivity level of each school before the introduction of the tournament could be linked to the realized effort induced by the tournament. To test whether the effects of the prize vary by prior performance levels, I estimated the main specification of Eq. 3 separately by performance quartiles. That is, I calculated a combined average of the performance in the years before the introduction of the prize and assigned each school to overall performance quartiles, combining the scores for both Language and Mathematics.

As displayed in table 4, the LATE in the performance in Mathematics by quartile (panel B) follows a u-shaped curve. That is, schools located in both the lowest (1st) and the highest (4th) quartiles before the introduction of the prize, had the highest local effects. On the other hand, the estimates for Portuguese, in panel A, do not follow the same pattern, schools at the top of the distribution did not presented a significant change in their scores, only the schools at the bottom. In terms of the dropout and failing rate I find a drop, due to the introduction of the prize, for the first and second quartely, respectively.

Although the patterns are not consistent between outcomes, these results do not support the hypothesis that the prize would undermine any efforts by the schools with the worst prior performance, so they could receive the money distributed to the schools at the bottom of the ranking. On the other hand, schools in the 3rd quartile show, on average, the smallest effects on both scores outcomes. Being around the median in terms of prior performance means that the chances of both receiving the prize or the support for bottom-ranking schools are lower. These results support the hypothesis that the PEN10 triggers heterogeneous incentives depending on

Table 4: Outcome: Score in Mathematics, by Quartile, Diff-in-Disc Estimates

	(1) 1st Quartile	(2) 2nd Quartile	(3) 3rd Quartile	(4) 4th Quartile
Panel A: Port				
Diff-in-Disc Est	0.321***	0.362***	0.067	0.204
	(0.121)	(0.104)	(0.116)	(0.145)
Bandwidth	8.8	8.8	8.8	8.8
N	2538	2926	2451	2371
Panel B: Math				
Diff-in-Disc Est	0.476***	0.292***	0.131	0.481***
	(0.120)	(0.108)	(0.120)	(0.160)
Bandwidth	9.2	9.2	9.2	9.2
N	2807	3187	2656	2592
Panel C: Drop				
Diff-in-Disc Est	-1.886**	0.302	0.304	0.003
	(0.928)	(0.753)	(0.801)	(0.598)
Bandwidth	8.9	8.9	8.9	8.9
N	2542	2923	2451	2371
Panel D: Fail				
Diff-in-Disc Est	1.294	-1.674**	0.264	0.881
	(0.924)	(0.764)	(0.887)	(0.764)
Bandwidth	7.2	7.2	7.2	7.2
N	2263	2633	2189	2136

Note: Table reports Diff-in-Disc estimates by performance level. Municipality fixed effects and controls are included. Significance levels are indicated by $*<.1$, $**<.05$, $***<.01$. Parentheses contains standard errors clustered at the school level.

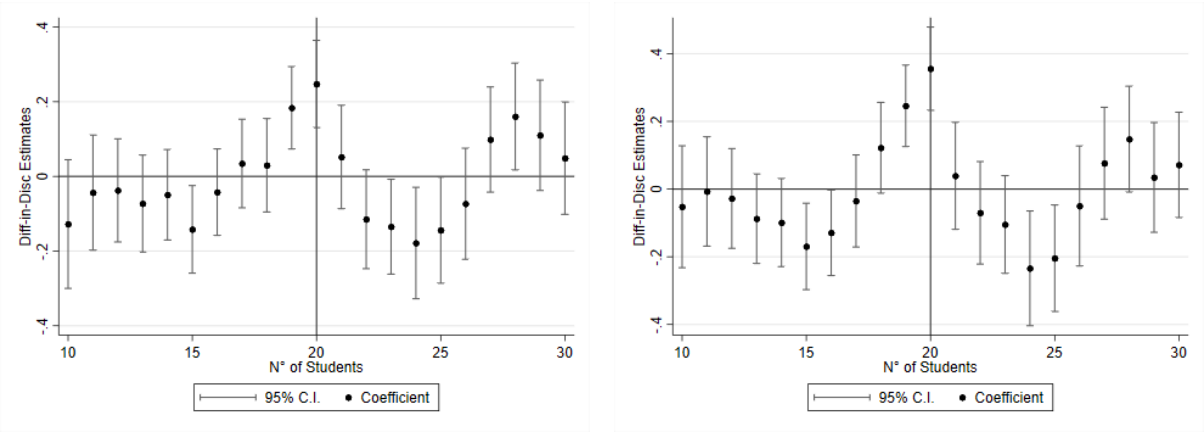
the relative performance of the schools. Schools located in the 3rd quartile do not have any gains in terms of the student's performance triggered by the award.

3.4 Robustness Checks

3.4.1 Placebo Cutoffs

As in any quasi-experimental setting, it is important to provide supportive evidence on the credibility of the identification strategy. By incorporating a time dimension in the setting, I control for time-invariant endogenous factors that could cause a discontinuity at the cutoff. To take into consideration other potential threats I conducted another series of tests.

First, if the design is credible, there should not be any significant effect of the introduction of the tournament by choosing placebos eligibility criteria. Figure 5 plots the main estimations for a range of placebo cutoffs. These graphs mostly support our design, the diff-in-disc estimates become positive and significant for estimations using the actual of 20 students. For most of my estimations using placebo cutoffs the estimations do not have any positive and significant effects.



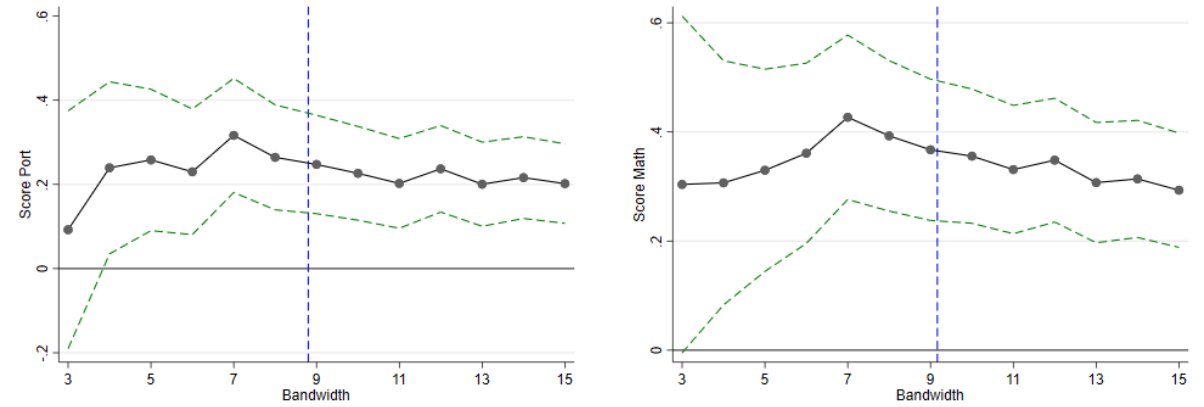
(a) Language (b) Mathematics

Figure 5: Placebo Cutoffs - Diff-in-Disc Estimates

3.4.2 Estimates by Bandwitdth

One critical choice in the empirical design is the optimal bandwidth. RDD results are sensitive to the estimation window of choice. In our case, the bandwidth selection followed the MSE-Optimal Bandwidth procedure, standard in the current literature. Figure 6 plots the diff-in-disc estimates for different bandwidth choices. The blue vertical line indicates the bandwidth used in the main specification.

This evidence suggests that my estimations are robust to different bandwidth sizes, for both of the outcomes. As expected, with smaller intervals the confidence intervals become larger, due to the decrease in the number of observations. On the other hand, for bandwidths slightly higher than the optimal ones, the estimates become more biased due to the use of observations further away from the running variable cutoff.



(a) Language (b) Mathematics

Figure 6: Diff-in-Disc Estimaes by Different Bandwitdh Choices

3.4.3 Dynamic Eligibility Status

One concern regarding the use of the diff-in-disc estimator for this setting is the lack of perfect compliance to the treatment. Differently from previous empirical studies using this estimator, the eligibility status are not fixed throughout time [Grembi et al. (2016); De Benedetto and De Paola (2019)]. Schools do not have the same amount of students enrolled every year, so the eligibility for the prize is yearly defined.

This uncertainty regarding eligibility could bias the results. That is, a school with more than 20 students in one year can become ineligible in the following year if it observes a drop in the cohort size, or vice-versa. Figure 7 displays the number of schools that switch treatment status by year. Previous to the introduction of the treatment there was an upward trend of schools switching from the treatment to the control group. However, since 2015 the number of switchers from both groups has remain slightly constant.

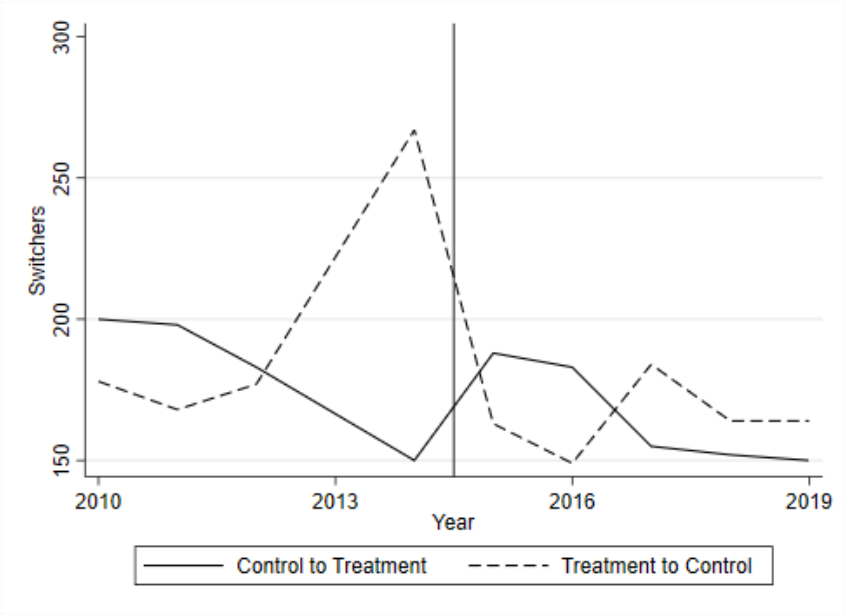


Figure 7: Eligibility Status Dynamic)

Given the significant presence of switchers in all years, it is important to investigate whether these dynamics in the eligibility status have consequences for my estimations. In order to check for that I re-sampled my data to only include schools that never changed their treatment status throughout all the observed years. Table 12, in the appendix, displays the yearly numbers of total and eligible schools with this new sample, where now we have roughly 1000 fewer schools per year. Table 5 displays the estimates for this sample. As we can observe the results remain fairly similar to the mains specification in table 3.

Hence, the fact that the eligibility status is not constant throughout the years does not seem to be introducing significant bias in the estimates. In particular, it slightly increases the size of the effects in the full specified estimates. But restricting our sample to the schools that do not switch their eligibility status reduces the amount of units around the cutoffs, almost doubling the optimal bandwidth and reducing the precision of the diff-in-disc estimator.

Table 5: School Grade A Award, Diff-in-Disc Estimates (9th Grade): Non-Switchers

	(1)	(2)	(3)
Panel A: Port			
Diff-in-Disc Est	0.250**	0.280***	0.273***
	(0.107)	(0.107)	(0.096)
Bandwidth	15.3	15.3	15.3
N	6571	6564	6562
Mun FE	N	N	Y
Controls	N	Y	Y
Panel B: Math			
Diff-in-Disc Est	0.408***	0.432***	0.407***
	(0.126)	(0.126)	(0.107)
Bandwidth	14.4	14.4	14.4
N	6195	6189	6187
Mun FE	N	N	Y
Controls	N	Y	Y

Note: Table reports Diff-in-Disc estimates using only non-switchers schools. Significance levels are indicated by $*<.1$, $**<.05$, $***<.01$. Parentheses contains standard errors clustered at the school level.

4 The Prize Effect

In the last section, I documented significant medium-term learning gains of being eligible for the School Grade A Prize. These gains reflect only the average improvement throughout the years of the schools participating in the yearly tournament. The heterogeneity analyses show that they were spread throughout the school's performance distribution. However, not only participating in the competition, but winning the prize can have important consequences for the schools. The aim of this section is to understand and quantify behavioral and academic changes at the schools after winning the prize.

Previous empirical evidence has shown that awards at work can affect workers' performance not only ex-ante but also ex-post. Further, it also yields positive spillovers to other job dimensions [Neckermann et al. (2014); Ashraf et al. (2014)]. Taking advantage of the setting established by the prize in Brazil, I can check whether these conclusions also hold for the public

education setting. More specifically, I will check whether I can observe changes in the behavior and productivity of teachers, principals, and students after being awarded the School Score A Prize.

4.1 Data

To conduct this empirical analysis I will focus on the 5th grade cohorts, due to data limitations of the two main data sets that I will explore in this section. First, the list of the yearly awarded and supported schools by the tournament was provided by the Secretary of Education of the state of Ceará. Further, to use behavioral outcomes in the analysis, I will rely on micro-data from a biannual evaluation conducted by the Brazilian Federal Government (SAEB). The SAEB is a series of evaluations conducted at public schools, and in a sample of private schools, that serves as an important guide for the Brazilian Educational policy.

Importantly for my analysis, the SAEB not only tests students on Mathematics and Language, but also surveys teachers, principals, and students on a broad range of issues regarding the school's environment, curriculum, and structure. Due to changes in the survey I will only use data from the survey years 2013, 2015, and 2017. Table 6 summarizes the statistics of the main outcomes from the SAEB used in this section. The survey contains a separate questionnaire for the principal, teachers, and students. Hence, I divided my main outcomes into three sections. All variables, excluding dummies, are normalized.

For the surveys done with the teachers and principals, I will be focusing on variables related to job performance and beliefs about the students. As displayed in table 6, I have information on how often the principal communicates with the parents and whether he promoted extracurricular activities in the school throughout the year. Further, I also observed the principal's reported belief in community support and on the issue of teacher absenteeism.

Regarding the teachers, I have measures on the reported level of communication between them and their relationship with the principal. Furthermore, I also have information on the teacher's beliefs about the students, once they are asked the fraction of the students in the cohort (5th grade) who will graduate high school. I also have information on the reported teaching practices, and whether the teachers implemented innovative teaching methods throughout the year. Finally, the teachers report the main issues observed in that school on that year, and I group the possible answers into whether these issues are related to the school, the teachers, the parents, or the students.

In panel C statistics for the variables from the survey completed by the students are displayed. Here, I can observe the reported effort of the students throughout the year, on how often they used the library and completed the homework. Further, I also have information on how the students perceived their parents ' presence on school issues. More specifically the students are asked whether the parents talk with them about what is happening at school. To conclude, I also have information on the their performance on the evaluation conducted by the SAEB in Portuguese and Mathematics. Section A.3 in the appendix provides further details on all these variables.

Table 6: Descriptive Statistics - SAEB

	Mean	SD	Min	Max	N
Panel A: Principal					
How offer does the school communicate with the parents	0.00	1.00	-4.38	1.35	4541
How often the school runs extracurricular projects	0.00	1.00	-2.59	2.35	4549
Teacher absenteeism is an issue	0.00	1.00	-0.52	3.74	4477
The community supports your mandate (dummy)	0.95	0.22	0.00	1.00	4629
Observations	4683				
Panel B: Teachers					
How often did your exchange Ideas with your peers	0.00	1.00	-4.15	2.11	4584
Teacher feel trusted by the director	0.00	1.00	-3.86	1.52	4589
How many of your students will graduate high school	0.00	1.00	-3.95	0.76	3030
Implement Innovative teaching practices (Portuguese)	0.00	1.00	-5.93	1.55	4557
Implement Innovative teaching practices (Mathematics)	0.00	1.00	-10.16	1.19	4532
Issue School (dummy)	0.20	0.21	0.00	1.00	4543
Issue Teacher (dummy)	0.20	0.28	0.00	1.00	4529
Issue Parents (dummy)	0.91	0.20	0.00	1.00	4552
Issue Student (dummy)	0.69	0.25	0.00	1.00	4570
Observations	4598				
Panel C: Students					
Students goes to the Library	0.00	1.00	-4.21	2.96	4633
Parental presence	0.90	0.05	0.42	1.00	4635
Student does the homework	0.00	1.00	-6.89	2.10	4634
Professor Corrects the Homework	0.00	1.00	-7.93	1.43	4635
Score SAEB Port	0.00	1.00	-3.04	4.11	4636
Score SAEB Math	0.00	1.00	-3.04	4.56	4636
Observations	4636				

4.2 Empirical Strategy

As done on the estimations of the tournament effect, I will again rely on exogenous variations in the treatment to exert causal evidence on the effects of receiving the prize. Since now I am interested in the effects of not competing, but winning the prize I will take advantage of the rank of the worst performing winner as a cutoff in a regression discontinuity design. That is, given that the rank is solely based on the school's performance, I will estimate the local difference in

my outcomes between the schools above and below the threshold.

As mentioned in the previous setting, the continuity assumption must hold. That is, in the absence of the treatment, my expected outcomes would remain a smooth function of the running variable, here the performance rank, across the cutoff. Since the performance rank does not give space to a rank manipulation, the continuity assumption here becomes less of an issue. In other words, the rank is solely defined by the school's performance in the SPAECE evaluation, given that they met all the program criteria. This assessment is done by a third-party evaluator, hence less subject to manipulations.

Here, the timeline of the School Grade A prize and the SAEB evaluation becomes important, as illustrated in Figure 8. The SAEB evaluation for a given year t (i.e., 2013, 2015, or 2017) is usually conducted between October and November of the respective year. However, the School Score A results for a given year, are announced by the end of June of the following year. Hence, by the time the survey for a year t was conducted, schools were aware whether they won the prize in $t-1$ and before, but not for t. In other words, using a survey from a given year t, I can compare schools around the cutoff in $t-1$ to estimate the effects of being awarded the prize in $t-1$ on my set of outcomes in t. Further, since there is a gap of two years between surveys, I can also compare schools around the cutoff in $t-2$ to estimate the effects of being awarded the prize in $t-2$ on my set of outcomes in t.

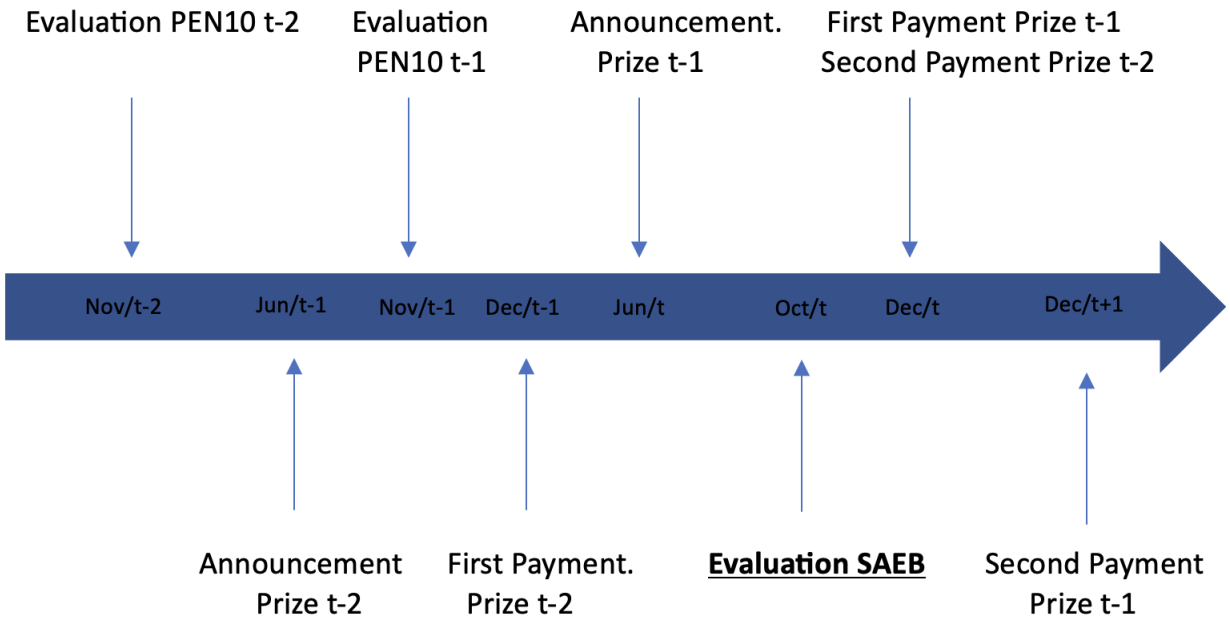


Figure 8: Timeline of the prize for years t-2 and t-1

Using the ranking in t-1, I will estimate the ex-post short-term effects of only being awarded, once that the monetary prize is usually transferred by the end of the year t, that is after the survey was collected. Further, using the ranking in t-2 I can estimate the effects almost two years after receiving the prize combined with the monetary effect of the prize transfers.

In the data, I do not observe full compliance with treatment. That is, some schools met all the eligibility criteria and scored above the worst performing winner, but still did not receive the prize. Although there is only a small number of cases (6% of all theoretically winners), to not introduce bias to the estimates I will implement a fuzzy RDD. Formally I will use an instrumental variable approach, by using the theoretical assignment to treatment T_{it} as an instrument for the actual treatment assignment variable P_{it}. Hence, within a optimally chosen bandwidth h (estimated following Calonico et al. (2020), I estimate the following model:

$$P_{it-1} = \alpha + \beta T_{it-1} + G(R_{it-1}) + \epsilon_{it} \tag{4}$$

$$Y_{it} = \mu + \delta \hat{P}_{it-1} + G(R_{it-1}) + X_{it} + \alpha_t + \alpha_m + \epsilon_{it} \tag{5}$$

Equation 4 represents the first stage, where β represents the effect of being above the cutoff on actually receiving the prize, assuming a flexible functional form on the ranking in t-1, $G(R_{it-1})$. Equation 5 estimates my main parameter of interest, δ, which represents the local effect, on average, of being above the winning cutoff on my set of outcomes Y_{it}. Further, X_{it} is a vector of individual characteristics of the respondent (gender, race, and age). Finally, α_t and α_m are survey year and municipality fixed effects, respectively. Error terms ϵ_{it} are clustered a the school level.

4.3 Results

Table 7 displays the regression discontinuity estimates for the set of outcomes reported by the principals. Receiving the prize in $t-1$ leads on average to a local increase in the reported communication of the principals with the parents of around 0.8 standard deviations. Further, it also leads to an increase of around 12% in the support of the community to the Principal's mandate, as perceived by the latter. On the other hand, these effects fade away when we look into the effects of receiving the prize in $t-2$, as displayed in panel B.

Moving into the teacher's questionnaire, table 8 combines a series of outcomes in which

Table 7: Prize Effects - Principal

	(1) Parents Communication	(2) Extracurricular Activities	(3) Teach. Abseenteism	(4) Community Support
Panel A: T-1				
RD Est	0.823***	0.279	0.047	0.119**
	(0.212)	(0.232)	(0.125)	(0.056)
Bandwidth	118.9	130.0	122.5	78.7
N	3794	3783	3753	3790
Panel B: T-2				
RD Est	0.504	0.252	0.131	-0.041
	(0.354)	(0.307)	(0.164)	(0.057)
Bandwidth	70.7	113.1	98.4	98.1
N	3789	3778	3747	3786

Note: Table reports RDD estimates on outcomes reported by the principals. The ranking of t-1 and t-2 are used as the running variable. Parentheses contain standard errors, clustered at the school level. Significance levels are indicated by $*<.1$, $**<.05$, $***<.01$.

teachers are asked whether the main issues for the learning development of the students are due to problems related to the school, the teachers, the parents or the students themselves. Interestingly, receiving the prize in $t-1$ leads to an average local increase in the perception of the teachers that the learning problems that occur in the school are due to the students. This points out the direction that the prize gives a confidence boost to the teachers, and if learning problems exist among some students, it is not due to the teachers or the school.

Table 8: Prize Effects - Teachers

	(1) Issue School	(2) Issue Teachers	(3) Issue Parents	(4) Issue Students
Panel A: T-1				
RD Est	-0.019	0.024	0.081	0.136**
	(0.044)	(0.056)	(0.051)	(0.065)
Bandwidth	129.7	134.3	127.5	98.2
N	3933	3936	3947	3965
Panel B: T-2				
RD Est	-0.079	0.050	0.033	-0.057
	(0.052)	(0.064)	(0.062)	(0.068)
Bandwidth	99.2	116.0	94.7	112.4
N	3937	3937	3948	3967

Note: Table reports RDD estimates on outcomes reported by the teachers. The ranking of t-1 and t-2 are used as the running variable. Parentheses contain standard errors, clustered at the school level. Significance levels are indicated by $*<.1$, $**<.05$, $***<.01$.

Further, column 3 in table 9 corroborates this hypothesis: receiving the prize in $t-1$ leads to a decrease in the reported belief of the teachers on the share of the students, among the 5th grade cohort, that will eventually graduate high school. That is, they believe that the students with current learning issues will not manage to catch up later on and some will eventually drop

out of school.

Table 9 also displays the RD estimates for outcomes related to teaching methods. Outcomes in columns 4 and 5 displays the estimates on the reported frequency in which teachers implement innovative pedagogical practices, as described in section A.3 in the appendix. The estimates indicate a causal increase in the reported use of innovative teaching methods two years after being awarded the School Grade A Prize, with an average of 0.5 SD combining Portuguese and Mathematics.

This prize effects indicates that being awarded changes the beliefs and behaviors of the school. Teachers becomes more open to innovative teaching methods and students reports an increase in their academic effort. Since these effects appears only two years after receiving the prize, it is likely due to the combination of the monetary and non-pecuniary prizes.

Table 9: Prize Effects - Teachers

	(1) Teachers Exchange	(2) Relation w.Principal	(3) Graduate High Sch.	(4) Port. Innovative Teaching	(5) Math. Innovative Teaching
Panel A: T-1					
RD Est	-0.099	-0.208	-0.396*	-0.146	-0.170
	(0.188)	(0.203)	(0.227)	(0.225)	(0.212)
Bandwidth	118.5	121.9	132.2	116.8	126.1
N	3974	3975	2561	3941	3918
Panel B: T-2					
RD Est	0.226	0.185	-0.389	0.495**	0.696***
	(0.218)	(0.227)	(0.322)	(0.246)	(0.255)
Bandwidth	113.8	101.9	95.6	114.6	125.1
N	3977	3978	2584	3945	3922

Note: Table reports RDD estimates on outcomes reported by the teachers. The ranking of t-1 and t-2 are used as the running variable. Parentheses contain standard errors, clustered at the school level. Significance levels are indicated by $*<.1$, $**<.05$, $***<.01$.

These changes in reported teaching practices are followed by significant changes in students' behavior and academic performance. As we can observe in panel B of table 10, two years after the schools received the prize, students report that they visit the library more often, that their parents give extra support on school matters, that they increased on average homework completion and, consequently, a local increase of almost 1 standard deviation on their performance on the exam conducted by the SAEB evaluation on both Portuguese and Mathematics.

This last result is illustrated here by Figure 9, in which the average SAEB scores in Mathematics in t are plotted against the School Grade A ranking in $t-2$. As expected, the school's performance in t is positively correlated to their performance on $t-2$. However, around the winner's ranking cutoff, there is a discontinuous jump in the scores in t. In other words, there is a significant gain on barely winning the prize two years before, compared to the students in those schools that performed slightly below the cutoff and, hence, were not awarded.

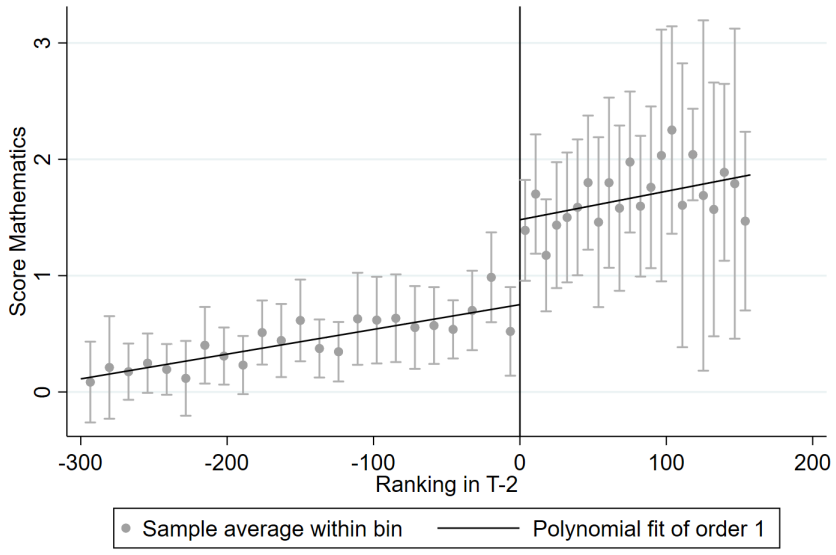
Table 10: Prize Effects - Students

	(1) Library	(2) Parents Support	(3) Homework	(4) Homework Correction	(5) Score Port	(6) Score Math
Panel A: T-1						
RD Est	0.037	-0.010	-0.206	-0.254	-0.172	-0.331
	(0.243)	(0.012)	(0.201)	(0.214)	(0.210)	(0.241)
Bandwidth	122.4	110.2	115.7	115.9	135.0	131.9
N	4171	4173	4173	4173	4173	4173
Panel B: T-2						
RD Est	0.453*	0.031**	0.765**	0.276	0.978***	0.926***
	(0.255)	(0.012)	(0.299)	(0.251)	(0.293)	(0.232)
Bandwidth	116.0	114.5	71.2	106.3	80.8	132.9
N	4183	4185	4185	4185	4185	4185

Note: Table reports RDD estimates on outcomes reported by the students. The ranking of t-1 and t-2 are used as the running variable. Parentheses contain standard errors, clustered at the school level. Significance levels are indicated by $*$ $< .1$, ** $< .05$, *** $< .01$.

Overall, these estimations indicate a relevant change in the schools' environment after being awarded as a School Grade A winner. On the one hand, right after receiving the prize school's principals indicates on average a higher support by the community on their mandates, and followed by that an increase in their communication with the parents. At the same time, the results point out temporary damage to the teacher's beliefs in the low-performing students, and they tend to attribute to them the potential issues within the schools. Further, the prize seems to also trigger a change in teaching methods and the student's effort and performance in the medium term.

Figure 9: Score in Mathematics across the T-2 cutoff



5 Conclusion

The use of incentive mechanisms to induce effort in teachers has been at the center of educational policy discussion. Previous research, focused on individual performance incentives, has produced mixed evidence when looking into the effects on students' academic scores. In this paper, I have relied on a quasi-experimental design to show that there is a gap to be filled by incentives mechanisms that do not solely rely on teacher performance pay schemes but incorporate group dynamics and non-pecuniary incentives.

The School Score A Award is a rank-order tournament that yearly rewards the top-performing schools in the state of Ceará in Brazil, and also supports the worst-performing schools through a cooperation program and an extra budget. The design of the award has important singularities, the financial prize needs to be mostly allocated to investment within the schools, limiting the distribution of bonuses. Further, the prize is highly advertised throughout the state and has a yearly ceremony which is documented by previous qualitative studies to be a source of motivation and pride for the participant's schools.

Through a differences-in-discontinuities design, I find positive and significant effects of a school being eligible for the award on its student's performance. Although I do not find effects on the approval and failing rates, the effects on the student's cores are robust to different specifications and tests. Overall, schools located just above the eligibility cutoff had a gain of 0.25 SD in Language and 0.35 SD in Mathematics. When estimating the effects by performance quartiles, I observe that the effects for Mathematics are mainly driven by schools at the top and bottom of the distribution, consistent with the incentives mechanisms of the program, in which these groups have higher chances to be granted an extra financial budget and be part of the cooperation program set out by the tournament.

Furthermore, on top of the ex-ante effect of competing for the prize, there are significant ex-post effects linked to the schools winning the prize. Through a RDD, I estimated the effects of winning the PEN10 on outcomes one and two years later. To summarize, this paper documents a series of beliefs and performance changes after receiving the prize. These findings highlights that implementing school awards with both monetary and non-pecuniary prizes not only yields in gains due to the competition effect, but also enhance good practices and students efforts among the prize winners.

These findings highlight the importance of going beyond bonus programs for teachers when

aiming to increase learning outcomes in developing countries. First, setting group incentives for schools can help achieve further productivity gains through peer effects. Further, non-pecuniary incentives can foster cooperation and motivation within public schools, and ultimately trigger intrinsic social values related to the work as an educator.

Given the structure of the policy, I do not disentangle the financial and non-pecuniary effects of the prize. Although I showed that these incentives combined are responsible for enhancing students' scores, a necessary further step is to estimate the size of the non-pecuniary effects on its own. Furthermore, it is important to understand the mechanisms behind these enhancements in the students performance and the role of the extra financial budget and the cooperation program in it.

References

Akerlof, G. A. and R. E. Kranton (2005). Identity and the economics of organizations. *Journal of Economic perspectives 19*(1), 9–32.

Ashraf, N., O. Bandiera, and S. S. Lee (2014). Awards unbundled: Evidence from a natural field experiment. *Journal of Economic Behavior & Organization 100*, 44–63.

Aucejo, E., T. Romano, and E. S. Taylor (2022). Does evaluation change teacher effort and performance? quasi-experimental evidence from a policy of retesting students. *Review of Economics and Statistics 104*(3), 417–430.

Barrera-Osorio, F., J. Cilliers, M.-H. Cloutier, and D. Filmer (2022). Heterogenous teacher effects of two incentive schemes: Evidence from a low-income country. *Journal of Development Economics 156*, 102820.

Barrera-Osorio, F. and D. Raju (2017). Teacher performance pay: Experimental evidence from pakistan. *Journal of Public Economics 148*, 75–91.

Behrman, J. R., S. W. Parker, P. E. Todd, and K. I. Wolpin (2015). Aligning learning incentives of students and teachers: Results from a social experiment in mexican high schools. *Journal of Political Economy 123*(2), 325–364.

Besley, T. and M. Ghatak (2017). The evolution of motivation. *Work. Pap., London School Econ.*

Butts, K. (2023). Geographic difference-in-discontinuities. *Applied Economics Letters 30*(5), 615–619.

Calonico, S., M. D. Cattaneo, and M. H. Farrell (2020). Optimal bandwidth choice for robust bias-corrected inference in regression discontinuity designs. *The Econometrics Journal 23*(2), 192–210.

Cattaneo, M. D., N. Idrobo, and R. Titiunik (2019). *A practical introduction to regression discontinuity designs: Foundations.* Cambridge University Press.

Cattaneo, M. D., M. Jansson, and X. Ma (2018). Manipulation testing based on density discontinuity. *The Stata Journal 18*(1), 234–261.

Cattaneo, M. D., M. Jansson, and X. Ma (2020). Simple local polynomial density estimators. *Journal of the American Statistical Association 115*(531), 1449–1455.

Ceará (2012). *Regime de Colaboração para a garantia do direito à aprendizagem: o Programa Alfabetização na Idade Certa (Paic) no Ceara*. Secretaria de Educação (SEDUC).

Ceará, A. L. d. E. (2006). *Educação de Qualidade começando pelo Começo*. Relatório Final do Comitê Cearense para a Eliminação do Analfabetismo Escolar.

Chevalier, A. (2022). Does employee ownership improve performance? *IZA World of Labor*, 311–311.

Conceição, A. d. S. (2022). Educação como forma de crescimento social: o caso do prêmio escola nota dez no ceará.

Contreras, D. and T. Rau (2012). Tournament incentives for teachers: Evidence from a scaled-up intervention in chile. *Economic development and cultural change 61*(1), 219–246.

Costa, A. G. and E. M. Vidal (2021). Prêmio escola nota dez no estado do ceará: concessão, ajustes e responsabilização. *Revista Brasileira de Estudos Pedagógicos 102*, 415–436.

Cotofan, M. (2021). Learning from praise: Evidence from a field experiment with teachers. *Journal of Public Economics 204*, 104540.

Dang, H.-A. H. and E. M. King (2016). Incentives and teacher effort: Further evidence from a developing country. *Economics of Transition 24*(4), 621–660.

De Benedetto, M. A. and M. De Paola (2019). Term limit extension and electoral participation. evidence from a diff-in-discontinuities design at the local level in italy. *European Journal of Political Economy 59*, 196–211.

Delfgaauw, J., R. Dur, O. Onemu, and J. Sol (2022). Team incentives, social cohesion, and performance: A natural field experiment. *Management Science 68*(1), 230–256.

Dixit, A. (2002). Incentives and organizations in the public sector: An interpretative review. *Journal of human resources*, 696–727.

Duflo, E., P. Dupas, and M. Kremer (2011). Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in kenya. *American economic review 101*(5), 1739–1774.

Duflo, E., P. Dupas, and M. Kremer (2015). School governance, teacher incentives, and pupil–teacher ratios: Experimental evidence from kenyan primary schools. *Journal of public Economics 123*, 92–110.

Duflo, E., R. Hanna, and S. P. Ryan (2012). Incentives work: Getting teachers to come to school. *American economic review 102*(4), 1241–1278.

Educação, T. P. (2021). *Anuário Brasileiro da Educação Básica, 2021*. Moderna.

Eggers, A. C., R. Freier, V. Grembi, and T. Nannicini (2018). Regression discontinuity designs based on population thresholds: Pitfalls and solutions. *American Journal of Political Science 62*(1), 210–229.

Friebel, G., M. Heinz, M. Krueger, and N. Zubanov (2017). Team incentives and performance: Evidence from a retail chain. *American Economic Review 107*(8), 2168–2203.

Fryer, R. G. (2013). Teacher incentives and student achievement: Evidence from new york city public schools. *Journal of Labor Economics 31*(2), 373–407.

Galindo Silva, H., N. H. Somé, and G. Tchuente (2018). Does obamacare care? a fuzzy difference-in-discontinuities approach. *Available at SSRN*.

Ganimian, A. J. and R. J. Murnane (2016). Improving education in developing countries: Lessons from rigorous impact evaluations. *Review of Educational Research 86*(3), 719–755.

Garibaldi, P. (2006). *Personnel economics in imperfect labour markets*. OUP Oxford.

Glewwe, P., N. Ilias, and M. Kremer (2003, May). Teacher incentives. Working Paper 9671, National Bureau of Economic Research.

Green, C., F. Lozano, and R. Simmons (2015). Rank-order tournaments, probability of winning and investing in talent: evidence from champions' league qualifying rules. *National Institute Economic Review 232*, R30–R40.

Grembi, V., T. Nannicini, and U. Troiano (2016). Do fiscal rules matter? *American Economic Journal: Applied Economics*, 1–30.

Gürtler, O. and M. Kräkel (2010). Optimal tournament contracts for heterogeneous workers. *Journal of Economic Behavior & Organization 75*(2), 180–191.

Lavy, V. (2002). Evaluating the effect of teachers' group performance incentives on pupil achievement. *Journal of political Economy 110*(6), 1286–1317.

Lavy, V. (2009). Performance pay and teachers' effort, productivity, and grading ethics. *American Economic Review 99*(5), 1979–2011.

Lazear, E. P. (2018). Compensation and incentives in the workplace. *Journal of Economic Perspectives 32*(3), 195–214.

Leonardi, M. and G. Pica (2013). Who pays for it? the heterogeneous wage effects of employment protection legislation. *The Economic Journal 123*(573), 1236–1278.

Loureiro, A., L. Cruz, I. Lautharte, and D. K. Evans (2020). The state of ceara in brazil is a role model for reducing learning poverty.

Loyalka, P., S. Sylvia, C. Liu, J. Chu, and Y. Shi (2019). Pay by design: Teacher performance pay design and the distribution of student achievement. *Journal of Labor Economics 37*(3), 621–662.

Mbiti, I., K. Muralidharan, M. Romero, Y. Schipper, C. Manda, and R. Rajani (2019). Inputs, incentives, and complementarities in education: Experimental evidence from tanzania. *The Quarterly Journal of Economics 134*(3), 1627–1673.

McEwan, P. J. (2015). Improving learning in primary schools of developing countries: A meta-analysis of randomized experiments. *Review of educational research 85*(3), 353–394.

Millán-Quijano, J. (2020). Fuzzy difference in discontinuities. *Applied Economics Letters 27*(19), 1552–1555.

Muralidharan, K. and V. Sundararaman (2011). Teacher performance pay: Experimental evidence from india. *Journal of political Economy 119*(1), 39–77.

Neal, D. (2011). The design of performance pay in education. In *Handbook of the Economics of Education*, Volume 4, pp. 495–550. Elsevier.

Neckermann, S., R. Cueni, and B. S. Frey (2014). Awards at work. *Labour Economics 31*, 205–217.

Santibañez, L. (2020). Teacher incentives. In *The Economics of Education*, pp. 431–441. Elsevier.

Winters, M., J. Greene, G. Ritter, and R. Marsh (2008). The effect of performance pay in little rock, arkansas on student achievement. research brief. *National Center on Performance Incentives*.

A Appendix

A.1 tables

Table 11: Tournament Effects: School Grade A Award, Diff-in-Disc Estimates (9th Grade)

	(1)	(2)	(3)
Panel A: Drop			
Diff-in-Disc Est	-0.614	-0.629	-0.442
	(0.401)	(0.402)	(0.390)
Bandwidth	8.9	8.9	8.9
N	10455	10455	10455
Mun FE	N	N	Y
Controls	N	Y	Y
Panel B: Fail			
Diff-in-Disc Est	-0.038	-0.066	0.170
	(0.418)	(0.417)	(0.410)
Bandwidth	7.2	7.2	7.2
N	9372	9372	9372
Mun FE	N	N	Y
Controls	N	Y	Y

Note: Table reports Diff-in-Disc estimates by performance level. Scores are normalized. Controls includes rural area dummy and a infrastructure index variable. Parentheses contain standard errors, clustered at the school level. Significance levels are indicated by $*<.1$, $**<.05$, $***<.01$.

Table 12: Number of Evaluated and Eligible Schools by Year - No Switchers Sample

Year	Evaluated Schools	Eligible Schools	Eligible Schools (%)
2008	1962	1340	68.29765
2010	1844	1237	67.08243
2011	1712	1217	71.08645
2012	1763	1186	67.2717
2014	1533	1085	70.77625
2015	1437	1048	72.92972
2016	1474	1072	72.72727
2017	1439	1043	72.48089
2018	1347	1016	75.42687
2019	1319	1008	76.42153

A.2 Figures

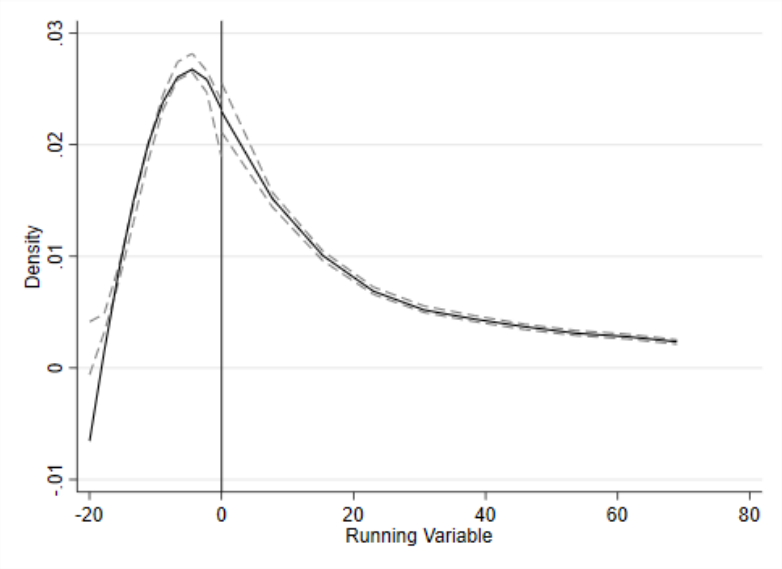


Figure 10: Schools Distribution by Cohort (9th Grade) Size - Before Treatment

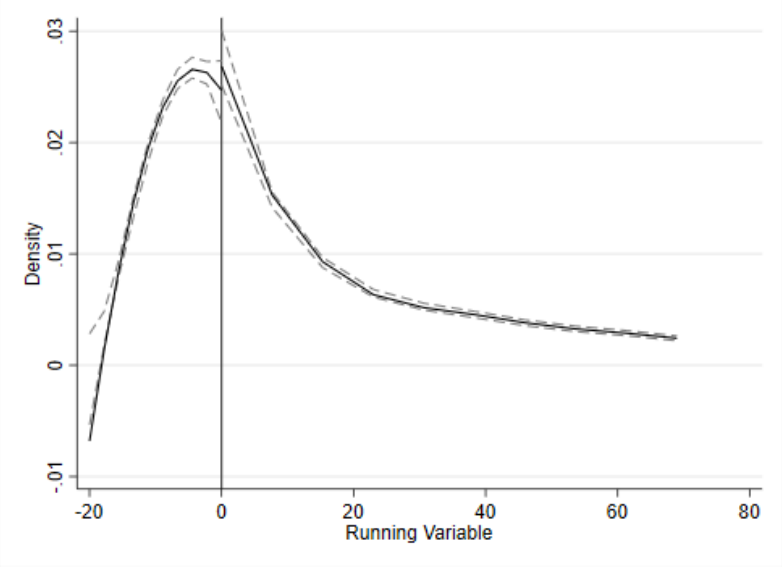


Figure 11: Schools Distribution by Cohort (9th Grade) Size - After Treatment

A.3 SAEB - Variables

A.3.1 Principal

- **Parents Comun**: Please indicate how often the following activities are carried out to minimize student absences this year and at this school (0- Never, 1- Sometimes, 2- Frequently, 3 - Always or almost always):

  - Parents/guardians are notified through school communication.
  - Parents/guardians are called to the school for a meeting with other parents.
  - Parents/guardians are called to the school individually to discuss the matter.

- **Extracurricular Act**: Please indicate how often you carried out the following activities this year at this school:(0- Never, 1- Sometimes, 2- Frequently, 3 - Always or almost always):
  - Developed extracurricular activities in sports.
  - Developed extracurricular activities in arts.
  - Developed thematic projects (e.g., bullying, environment, social inequalities, etc.).

- **Teach. Absenteeism**: Was the operation of the school hindered by any of the following problems? :(0- No, 1- Yes, to a small extent., 2- Yes, moderately. 3 - Yes, significantly.):
  - High absenteeism rate among teachers.

- **Commun. Support**: Consider the existing conditions for the position of principal in this school: :(0- No, 1- Yes):
  - Is there support from the community for your management?

A.3.2 Teachers

- **Teachers Exchange**: In this school, how often did you do the following::(0- Never, 1- Once a Year, 2- 3 to 4 times a year, 3 - Monthly, 4- Weekly):
  - Participated in the planning of the school curriculum or part of it.
  - Exchanged teaching materials with your colleagues.
  - Participated in meetings with colleagues who teach the same grade level as you.
  - Involved yourself in joint activities with different teachers (e.g., interdisciplinary projects).
  -

- **Relation w.Principal** : In this school and this year, indicate the frequency with which :(0- Never, 1- Sometimes, 2- Frequently, 3 - Always or almost always):

- The principal discusses educational goals with teachers during meetings.
- The principal informs teachers about professional development opportunities.
- The principal pays special attention to aspects related to student learning.
- The principal pays special attention to aspects related to administrative norms.
- The principal pays special attention to aspects related to school maintenance.
- The principal encourages and motivates me in my work
- The principal stimulates innovative activities.
- I feel respected by the principal.
- I have confidence in the principal as a professional.

- **Graduate High Sch.** : How many of the students in this class do you think will complete high school? (0- Few Students, 1- Slightly less than half of the students , 2- Slightly more than half of the students , 3 - Almost all of the students):

- **Port. Didactic**: Indicate the frequency with which you implement the following pedagogical practices in this class: (0- Never, 1- Once a Year, 2- 3 to 4 times a year, 3 - Monthly, 4- Weekly, 5- Daily)
  - Propose grammatical activities related to newspaper or magazine texts.
  - Promote the reading and discussion of short stories, chronicles, poems, or novels.
  - Use short stories, chronicles, poems, or novels to practice grammar aspects.
  - Utilize comic books as a learning tool.
  - Reinforce the names of grammatical and linguistic concepts.

- **Math. Didactic**: Indicate the frequency with which you implement the following pedagogical practices in this class: (0- Never, 1- Once a Year, 2- 3 to 4 times a year, 3 - Monthly, 4- Weekly, 5- Daily)
  - Do exercises to reinforce procedures and rules.
  - Discuss whether the numerical results obtained in solving a problem are suitable for the given situation.
  - Discuss different methods to solve problems and calculations.

  - Deal with topics that appear in newspapers and/or magazines, discussing their relation to mathematics.
  - Provide schemes/rules that enable obtaining the correct answers to calculations and problems.
  - Experiment with different actions (collecting information, cutting, exploring, manipulating, etc.) to solve problems.

- **Issue School**:In your perception, do possible learning problems in the evaluated grade(s) or year(s) occur in this school due to : (0- No, 1- Yes)
  - Lack of physical infrastructure?
  - Insufficient or ineffective supervision, coordination, and pedagogical guidance?
  - Inadequate curriculum content for the students' needs?

- **Issue Teachers**:In your perception, do possible learning problems in the evaluated grade(s) or year(s) occur in this school due to : (0- No, 1- Yes)
  - Excessive workload on teachers, making it difficult to plan and prepare lessons?
  - Dissatisfaction and demotivation of teachers with their teaching career?

- **Issue Parents**:In your perception, do possible learning problems in the evaluated grade(s) or year(s) occur in this school due to : (0- No, 1- Yes)
  - Cultural level of the students' parents?
  - Lack of assistance and parental involvement in the student's school life?

- **Issue Students**:In your perception, do possible learning problems in the evaluated grade(s) or year(s) occur in this school due to : (0- No, 1- Yes)
  - Low self-esteem of the students?
  - Lack of interest and effort from the students?
  - Student discipline issues in the classroom?
  - High rate of student absenteeism?

A.3.3 Students

- **Library**:Do you use the library or reading room of your school? : (0- Never or almost never, 1- Occasionally, 2- Always or almost always)
- **Parents Behvr**: Your parents or guardians : (0- No, 1- Yes)
  - Do they encourage you to study?
  - Do they encourage you to do homework and/or school assignments?
  - Do they encourage you to attend school and/or not miss classes?
  - Do they talk to you about what happens at school?
- **Homework**:Do you do your homework? : (0- Never or almost never, 1- Occasionally, 2- Always or almost always)
  - Portuguese.
  - Mathematics.
- **Homework Correction**: Does the teacher correct the Portuguese homework? : (0- Never or almost never, 1- Occasionally, 2- Always or almost always)
  - Portuguese.
  - Mathematics.