Economics of Education Review 80 (2021) 102067

Contents lists available at ScienceDirect

Economics of Education Review

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

Teacher Turnover and Financial Incentives in Underprivileged Schools: Evidence from a Compensation Policy in a Developing Country

Rafael Camelo[a], Vladimir Ponczek[1,*,b]

[a] Plano CDE Institute and Sao Paulo School of Economics - FGV. Brazil
[b] Sao Paulo School of Economics - FGV and LEARN Rua Itapeva 474, Sao Paulo, SP 01332-000, Brazil

ARTICLE INFO

JEL classification:
I20
J38
J45

Keywords:
teacher labor markets
teacher turnover

ABSTRACT

This paper evaluates the impacts of a policy that introduced a sizeable wage premium (24% to 36%) to teachers at disadvantaged schools on teacher turnover in the public school system in São Paulo, Brazil. We explore a discontinuity at the eligibility rule to identify the policy effect. We find that the wage compensation reduced turnover by 5 p.p. (10.4% over the pre-treatment average). We also show that this policy positively impacted the achievement of low-performing students, but had no effects on average test scores. We rule out alternative explanations, such as reallocation of teachers or direct effects of the wage increase. These results suggest a disruptive effect of teacher turnover on learning, especially on students at the bottom of the test score distribution.

1. Introduction

Attraction and retention of effective teachers influence educational results. Teacher turnover can be costly for students' academic performance (BOYD, GROSSMAN, LANKFORD, LOEB, & WYCKOFF, 2008; RONFELDT, LOEB, & WYCKOFF, 2013), as turnover breaks the routine of classes. Substitute teachers are, in general, less experienced (CLOTFELTER, GLENNIE, LADD, & VIGDOR, 2008; HANUSHEK, KAIN, O'BRIEN, & RIVKIN, 2005) and less effective (BOYD et al., 2008; HANUSHEK, KAIN, & RIVKIN, 2004).

When choosing a school in which to teach, teachers base their decisions on the profile of students (BOYD et al. (2008); HANUSHEK et al. (2004); KASMIRSKI (2012); SCAFIDI, SJOQUIST, and STINEBRICKNER (2007); SIMON and JOHNSON, (2015); CARVER-THOMAS & DARLING-HAMMOND, 2017) and the location of the school (BOYD et al., 2008; CARVER-THOMAS & DARLING-HAMMOND, 2017). In general, these preferences disadvantage the poorest and the most remote schools (BOYD et al., 2008; CARVER-THOMAS & DARLING-HAMMOND, 2017; LOEB, BETEILLE, & KALOGRIDES, 2012). Schools with underperforming students are the least likely to attract and retain teachers (ALLENSWORTH, PONISCIAK, & MAZZEO, 2009; HANUSHEK et al., 2005; HANUSHEK et al., 2004; HEMPHILL & NAUER, 2009; INGERSOLL, 2001; JOHNSON, BERG, & DONALDSON, 2005; MARINELL & COCA, 2013; RONFELDT et al., 2013).

Human resource management practices that foster teachers' professional development and rationalize the teacher-school matching can reduce turnover (LOEB et al., 2012). Another possible way to reduce teacher turnover is through financial compensation. Differential compensation has been studied in detail. In general, additional income can compensate for non-wage characteristics of a job (ROSEN, 1986). The idea behind creating a compensation differential for teachers working at non-ideal schools is that teachers, despite their preferences in schools' attributes, also respond to monetary stimulus. However, education research lacks a consensus on prescribing monetary compensatory policies to reduce teacher turnover. In fact, (HANUSHEK et al., 2004) show that the wage gap between schools in Texas plays only a minor role in a teacher's decision to switch schools. On the other hand, (CLOTFELTER et al., 2008) use hazard models to evaluate the impact of a wage premium for teachers in underperforming schools in North Carolina. They conclude that the reduction in average turnover was 17%, and the impact was higher for the most experienced teachers.

This paper evaluates the *ALE* program (*Adicional por Local de Exercício*), a wage compensation policy in the largest public school system in Brazil (São Paulo state). This school system serves over three million students with 150,000 teachers. Beginning in 2008, teachers working in disadvantaged schools started to receive a premium, equivalent to 24%

* Corresponding author.
E-mail address: vladimir.ponczek@fgv.br (V. Ponczek).

1 Ponczek gratefully acknowledges financial support from CNPq.

https://doi.org/10.1016/j.econedurev.2020.102067
Received 23 September 2019; Received in revised form 20 November 2020; Accepted 21 November 2020
Available online 22 December 2020

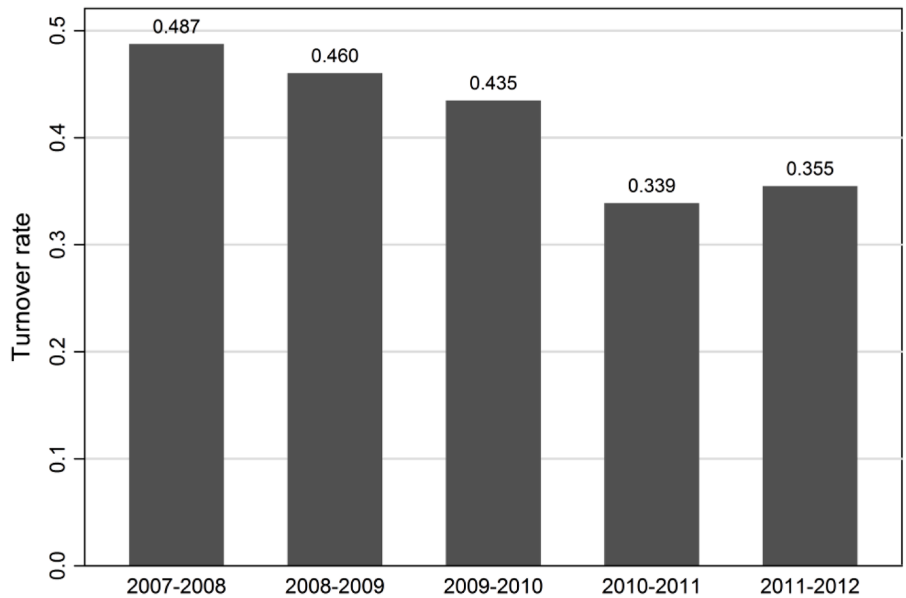

Fig. 1. Average school turnover rate

to 36% of their wage. Treated schools were chosen by an arbitrary rule, based on a socioeconomic index, outside of the school's control. This provides a regression-discontinuity design that allows us to estimate the causal impact of this policy. We analyze the impact of the compensation policy up to 2012; our data covers the period from 2007 to 2012.

Some studies have investigated the effects of specific monetary compensations for poor school on several outcomes. CABRERA and WEBBINK (2019) find that an Uruguayan program that increases teacher salaries and school inputs for target schools changed the teacher composition by increasing the average years of experience and tenure. However, they do not find any robust effect on student outcomes. CASTRO and ESPOSITO (2018) show that a bonus program to teachers in Peru decreased the number of teacher vacancies, but no effects on learning outcomes. This paper also dialogues with the literature focused on evaluating program for disadvantaged schools. For instance, DEE and WICKOFF (2015) and COWAN and GOLDHABER (2018) find that accountability measures increased teacher retention and performance at schools in thee U.S.

Different from previous studies that focused on policies that target specific teachers (particular grades or teachers with specific credentials), we focus on the effect of a general monetary stimulus unrelated to performance on turnover and its direct consequences on achievement. Recent literature indicates that turnover may directly harm achievement. RONFELDT et al. (2013) show that teacher turnover harms the performance of fourth and fifth graders, even after controlling for teacher quality, and also for students of the stayers (teachers that do not leave). The authors argue that part of this adverse effect comes from a disruption of the school's organization, which is also the claim of BOYD et al. (2008), which analyses the turnover of high-achieving teachers. When teachers leave a school, there may be a loss of institutional memory related to those teachers' knowledge about the school's routines, the principal's and coordinators' methods of work, relationship within the faculty, students' difficulties, and other aspects that may affect teaching.

Our benchmark models show that the extra payment reduces teacher turnover by 5 percentage points (p.p.), which means a drop of 10.4% over the pre-treatment average. We do not find any significant effects of the policy on average test scores. However, the wage premium appears to impact the achievement of low-performing students. The program reduced the proportion of low-achievers by 6.8 p.p. in Math and 5.4 p.p. in Reading. These effects are equivalent to reductions of 11.3% and 17.4% of low-performing students in Reading and Math, respectively.

It is well-known that students' and teachers' profiles in high-poverty schools in many developing countries (such as Brazil) are quite different from high-poverty schools in developed countries. Turnover rates are quite different between developing and developed countries. For instance, (CLOTFELTER et al., 2008) report a turnover rate of around 30% in North Carolina in the U.S.. In São Paulo, Brazil, this figure goes up to 50%. Teachers are more educated and have more experience in the sample from North Carolina. Additionally in developing countries, violence, severe learning difficulties, behavior problems, extreme poverty, lack of basic infrastructure (sanitation, electricity, etc.) are all very common in high-poverty schools. Moreover, teachers in developing countries are usually at the bottom of the country's overall wage distribution. Therefore, a teacher's reaction to monetary stimulus in the developing world is probably different than that of a teacher in a developed country, even if both work in underprivileged schools. For instance, teachers' wages in Brazil are lower compared to the U.S., which makes this kind compensation policies more feasible. In developed countries where wages are higher, financial compensation programs are likely to be much more expensive in order to have the same impact.

We further discuss alternative explanations by which this incentive may impact achievement of low-performing students: i) changes in the composition of teachers; and ii) changes in teachers' effort due to the wage increase.

First, *ALE* may have impacted achievement by changing the composition of teaching staff. In general, incoming teachers are less experienced and less effective than veteran teachers (BOYD et al., 2008; CLOTFELTER et al., 2008; HANUSHEK et al., 2005; HANUSHEK et al., 2004; HANUSHEK, RIVKIN, & SCHIMAN, 2016). These new teachers tend to be allocated to the most disadvantaged schools (BOYD et al., 2008; LOEB et al., 2012). PUGATCH and SCHROEDER (2014), using data from Gambia, find that a hardship allowance program attracted more qualified teachers to treated schools, although the program did not impact average achievement. Therefore, we check if the program in São Paulo changed the composition of teachers. We estimate whether *ALE* attracted teachers with different observed (experience, age, gender, education, working status and tenure) and unobserved characteristics, measured by the teacher's valued-added prior to the program.

Since *ALE*'s incentive is exclusively financial, we further investigate whether the increase in teacher earnings caused the improvements in achievement via teachers' effort. However, the link between teacher earnings and student performance is not well established. REE, MURALIDHARAN, PRADHAN, and ROGERS (2018) use a large-scale randomized experiment and find that unconditional payment increases are ineffective at improving student learning outcomes in Indonesia. A common explanation for such results is that teachers' pay is, generally, unrelated to their effort and productivity. However, even with pay-for-performance schemes, the link to achievement is unclear (FIGLIO and KENNY (2007), among others). *ALE* gives teachers with different workloads in treated schools different levels of compensation. We use this fact to try to uncover a possible effect of wage increase on students' proficiency through teachers' effort.

We present evidence that suggests that *ALE* has impacted students' performance solely through reducing teacher turnover. We find no evidence that *ALE* has impacted teachers' quality, nor that the increases in teachers' remuneration directly affects low-performing students' achievement.

The paper proceeds as follows. The next section describes the main institutional rules for hiring and allocating teachers in São Paulo's state schools. Section 3 describes the compensation wage policy (*ALE*). Section 4 discusses the data and identification strategy. Section 5 presents the results and discusses alternative explanations. Section 6 shows robustness tests. Finally, the last section summarizes our conclusions and discussion on the issue.

2. Institutional background

The state of São Paulo has the largest public education system in Brazil, with over 3.6 million students and about 150,000 teachers in 5,600 schools, including elementary, middle and high school. As in any public school system in Brazil, a teacher's career is governed by a wide set of rules defined, mostly, by the central administrative body, the State Department of Education (*Secretaria Estadual de Educação* - SEE).

The school year in São Paulo goes from February to December. The allocation of teachers to schools and classes takes place annually before the start of the new school year. It usually happens in January. Public school teachers, like most civil servants in Brazil, are hired by a public tender and have a permanent contract with job stability guaranteed by law.[2] Priority for choosing schools is granted using a scoring system based on teachers' tenure. This process is fully centralized by SEE and depends solely on a teacher preferences and attributes; principals have no power to interfere.

Graph 1 shows turnover rates in the analyzed period. The turnover rate is the proportion of teachers in a school who are replaced between two years. That is, on average in São Paulo state's schools, 48.7% of teachers changed schools between 2007 and 2008. This rate has been dropping since 2007. In 2011-2012, it reached 35.5%.

In São Paulo state's public schools, a teacher's hourly wage depends, generally, on two variables: tenure and certification. Both of them define what we call the "base wage". Beyond this remuneration, some teachers receive grants, other compensation and bonuses.[3,4]

The tenure determines a teacher's salary by the so-called "quinquennial rule", a state law that establishes a 5% increase on the base wage for all public servants for every five years of work, independent of any other characteristic, such as position, qualification or productivity.

Additionally, a teacher's salary can increase through career progression. The State Department of Education establishes that a teacher's career can progress through five levels. At each level, the base wage also increases by 5%. This progression depends on the teacher's certification. In São Paulo's public schools, every teacher must have at least an undergraduate degree in teaching or in a specific area (Portuguese language, English language, arts, math, physics, history, etc.). As a teacher obtains additional certificates (in post-graduate courses, for example) he/she can advance to higher career levels. For instance, the 2008 hourly nominal wage, for elementary school teachers, ranged from US$4.46 (for the least experienced and qualified level I) to US$6.92 (for the most-experienced and qualified level V). Secondary and high school teachers' base salary ranged from US$5.17 to US$8.01 per hour.[5]

On average, teachers work 30 hours per week, so a typical elementary school teacher earned between US$535.20 and US$830.40 per month in 2008. An average secondary/high school teacher earned between US$620.40 and US$961.20 dollars.[6]

Table 1
Summary statistics of pre-treatment characteristics

	Control	Treatment	p-value
Turnover	0.453	0.528	0.000
Standard. Reading Score	0.103	-0.198	0.000
Standard. Math Score	0.071	-0.213	0.000
% Low Perform. Stud. in Reading	0.229	0.310	0.000
% Low Perform. Stud. in Math	0.506	0.599	0.000
Stud. Gender	0.514	0.516	0.227
School Delay	0.113	0.137	0.000
Mother's Schooling	0.478	0.328	0.000
Father's Schooling	0.499	0.346	0.000
Teacher's Age	43.355	41.092	0.000
Teacher's Experience	13.127	12.174	0.000
Temporary Teachers	0.410	0.504	0.000
Teacher's Workload	20.479	20.676	0.099
Teacher's Gender (Male)	0.170	0.176	0.641
Teacher's Tenure	4.361	3.387	0.000

3. *ALE* - The compensation policy

Adicional por Local de Exercício (*ALE*) is a policy aimed at reducing turnover of staff (including teachers, coordinators, principals, and support staff) in schools in poor urban areas.[7] This target staff receives an additional monthly compensation equivalent to 24% to 34% of their gross base wage, depending on their career position. The Department of Education announced on January 30th of 2008[8] that *ALE* would start in February of the same year.

[2] Job stability is a common characteristic of civil servants' contracts in Brazil.

[3] Some of the additional remuneration is determined by Brazilian labor law, such as the compensations for night shifts (20% over the basic wage).

[4] It is important to note that these wage policies have been altered since 2010. In this study, we focus on the current rules, between 2007 and 2012, a period in which the rules did not change as much.

[5] Table A1 in the appendix shows wage per hour values by tenure and career.

[6] For comparison, the Brazilian minimum wage was US$224 per month in 2008.

[7] All schools in rural areas are also part of the policy, but since they do not have a suitable control group, they will not be analyzed.

[8] After the end of the teachers' allocation period for 2008.

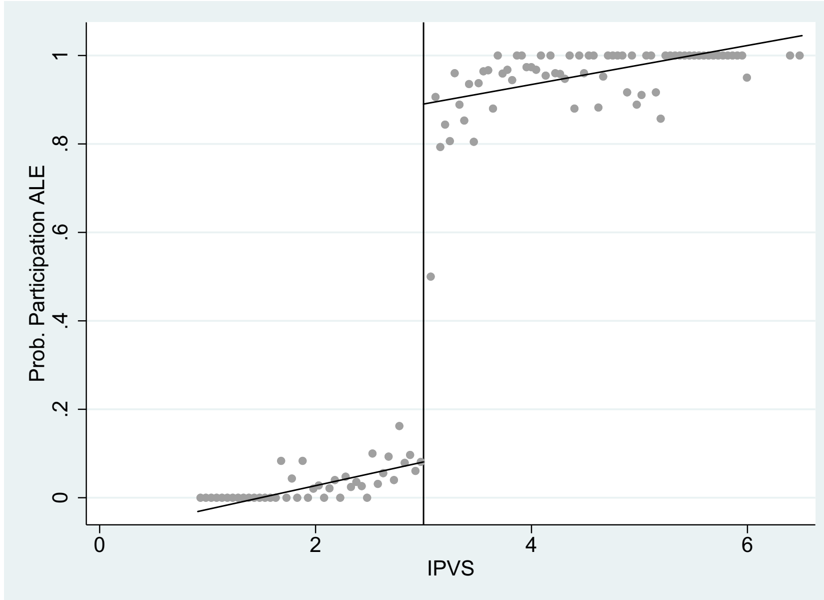

Fig. 2. Treatment assignment by *IPVS*

The value of this incentive is fixed independent of a teacher's career level or tenure. In 2008 dollars, teachers earned an additional US$1.53 per hour worked in *ALE* schools. Therefore, the total additional value received depends on the teacher's workload in *ALE* schools. Additionally, the weight of this incentive on teachers' earnings varies across career level and experience. More precisely, *ALE*'s relative impact on earnings decreases over experience and career position. For a beginning teacher (level-I), *ALE*'s incentive raises hourly earnings by 34% (elementary school) or 29% (secondary/high school). For the most experienced teachers (level-V), *ALE* increases hourly earnings by 22% (elementary school) or 19% (secondary/high school).

A school's location is the only eligibility criterion. The policy applies to schools in the 39 cities of the metropolitan area of São Paulo and in 14 other large municipalities (those with more than 300,000 inhabitants). In these municipalities, the policy targets schools in the poorest areas, defined by a socioeconomic index called *IPVS* (*Índice Paulista de Vulnerabilidade Social*). In 2008, there were 2,746 schools in eligible municipalities. Out of these 1,422 received the compensation and 1,324 did not.

IPVS was calculated by the state statistical bureau (*Fundação Seade*). Each census tract in the state of São Paulo is classified according to a degree of social vulnerability. São Paulo state has over 50,000 census tracts. The index is composed of socioeconomic variables, such as household income, characteristics of household head, and family composition. The indicator has a discrete scale ranging from one (no vulnerability) to six (very high vulnerability) (Table A2 in appendix). *IPVS* was calculated in 2002, six years before the implementation of *ALE*, based on the 2000 Brazilian population census. After the release of the new 2010 population census database in 2012, the state government decided to change *ALE*'s assignment rule. For that reason, we focus our analysis on the period before 2012.

Using *IPVS*, each state public school also received its own vulnerability index, calculated as the average of the discrete *IPVS* ratings of the census tracts within 300 meters around the schools. Therefore, the bureau created a continuous index for all public schools in the state. The map in Figure A1 in the appendix illustrates how the areas around schools are defined. Using this criterion, the State Department of Education (*SEE*) decided which schools would take part in *ALE*. Figure 2 shows the probability of treatment according to the continuous values of the *IPVS* for eligible schools. It shows a significant jump (64 p.p.) in this probability at IPVS=3. The SEE included most, but not all, schools with IPVS$\geq$ 3. Some schools below this threshold were also included.[9] This decision generated a fuzzy assignment for schools with *IPVS* around 3.

Table 1 outlines a general profile of schools in eligible regions during the pre-treatment period, both those that did and those that did not receive *ALE*. It shows that the compensation policy is focused on lower socioeconomic schools. Treated schools actually had higher turnover rates before the program (almost 53%). They also had more students with worse profiles: lower average scores on proficiency exams, a higher proportion of poor performers, and less-educated parents. In addition, although differences are marginal, vulnerable schools had less-experienced teachers and fewer permanent contract ones. All these differences indicate that treated schools are indeed those with more difficulty retaining teachers.

4. Data and Identification Strategy

The data come from three different sources: (i) the Brazilian Education Census, an annual survey that allows us to identify the match between teachers and schools, as well as some teachers' characteristics such as age, experience, tenure, contract type (temporary or permanent)[10] and workload;[11] (ii) State School Performance Assessment System (*SARESP*)[12] from which we extract data on students' performance and their attributes such as gender, grade and parental level of education (it is available only for students in 3rd, 5th, 7th, 9th, and 11th grades);

[9] The reasons for these inclusions vary: pressure from schools, principals, and local politics. We have no control over these reasons.

[10] Public school teachers can be divided into two broad functional categories, according to the type of employment contract: about 54% of teachers have permanent contracts and the remainder have fixed-term contracts (they are usually called temporary teachers).

[11] For experience, teaching load and contract type, we have information only until 2009.

[12] *SARESP* is a state-wide standardized assessment based on multiple-choice questions. Scores are computed by the Department of Education using IRT approach and range on a scale from 0 to 500. Teachers do not participate in this process. We have data on Reading and Math scores of students from the 3rd, 5th, 7th, 9th and 11th grades.

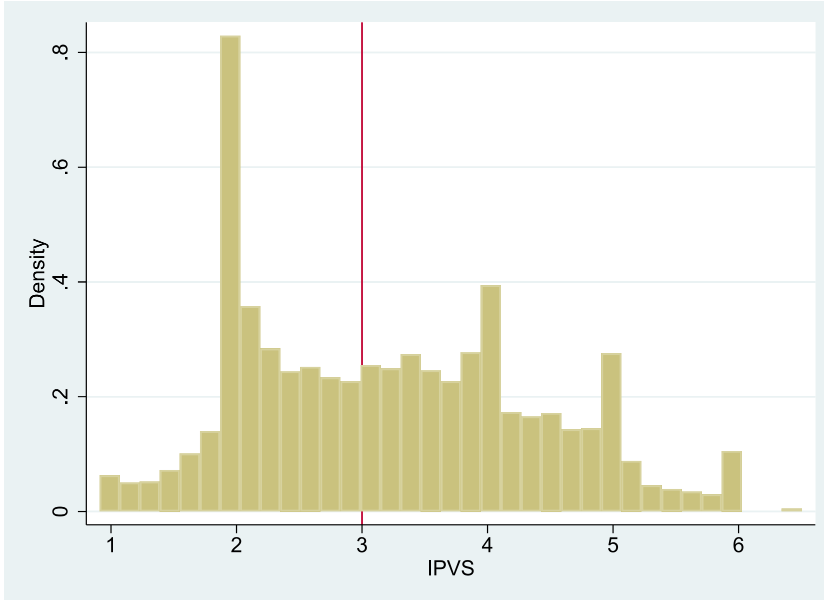

Fig. 3. School Density by *IPVS*

Table 2
Estimates of differences in pre-treatment characteristics

	(1) Turnover	(2) Reading Ave. Score	(3) Math Ave. Score	(4) Reading % low perform.	(5) Math % low perform.	
RD_Estimate	0.005	-0.077	-0.089	0.013	0.003	
	(0.028)	(0.082)	(0.091)	(0.029)	(0.043)	
Effective n. of schools	1249	683	727	859	858	
BW	0.884	0.514	0.546	0.652	0.649	
	Student's Gender	School Delay	Mother's Schooling	Father's Schooling		
RD_Estimate	-0.003	-0.018	-0.015	-0.023		
	(0.011)	(0.016)	(0.024)	(0.026)		
Effective n. of schools	867	848	1028	861		
Bandwidth	0.658	0.637	0.751	0.653		
VARIABLES	Teacher's Age	Teacher's Experience	Temporary Teachers	Teacher's Workload	Teacher's Gender	Teacher's Tenure
RD_Estimate	-0.902	-0.175	-0.004	0.922	-0.043	-0.761
	(0.699)	(0.617)	(0.036)	(0.811)	(0.074)	(0.975)
Effective n. of schools	856	871	1245	776	1111	831
BW	0.639	0.654	0.879	0.578	0.812	0.625

Results from the fuzzy RD regression. *IPVS* is the running variable and *ALE* is the treatment variable. Bandwidth and effect. n. of schools chosen using CALONICO et al. (2014). Standard errors in parentheses. $***$ p<0.01, $**$ p<0.05, $*$ p<0.1.

and (iii) the *IPVS* database, which shows the continuous index of the schools (non-public information). We use data from 2007 to 2012.[13]

Our identification strategy follows a standard fuzzy regression discontinuity design. As shown in Graph 2, *IPVS* generates a discontinuity on the probability of treatment at the threshold *IPVS*=3. By this set, some schools are induced to participate in *ALE* by the continuous *IPVS*, so it is possible to identify the local treatment effect parameter.

As a benchmark, we estimate the parameter of interest using the following model in an interval around $X = 3.0$.

$$Y_i = \beta_0 + \beta_1(X_i - 3) + \beta_2(X_i - 3)Z_i + \beta_3 W_i + u_i \tag{1}$$

$$W_i = \alpha_0 + \alpha_1(X_i - 3) + \alpha_2(X_i - 3)Z_i + \alpha_3 Z_i + e_i \tag{2}$$

Y_i is the turnover rate of a given school i; W_i is the treatment assignment (binary) variable; $Z_i = 1(X_i \geq 3.0)$ is the instrumental variable; X_i is the running variable (the school's IPVS); u_i and e_i are error terms. Our parameter of interest is β_3.

In order to identify this estimand as a causal impact of treatment, we must assume that both $E[Y(0)|X = x]$ and $E[Y(1)|X = x]$ are continuous functions at $X = 3$ (HAHN, TODD, & KLAAUW, 2001). We run local linear regressions with triangular kernel and choose the bandwidths according to CALONICO, CATTANEO, and TITIUNIK (2014).[14] The RDD strategy provides reliable empirical evidence of the casual effect of *ALE* on turnover, as long as there is no other discontinuity of the schools' characteristics at the threshold value (*IPVS* $= 3$).

In order to check the continuity assumption, we test whether there was manipulation on the running variable, and whether treatment and control schools were similar in the pre-treatment period around the

[13] Unfortunately, there is no data on teacher turnover prior to 2007; after 2012, the *ALE* assignment changed and other programs were introduced in vulnerable schools.

[14] As robustness checks, we also run regressions with higher-order polynomial functions, different kernels and bandwidths.

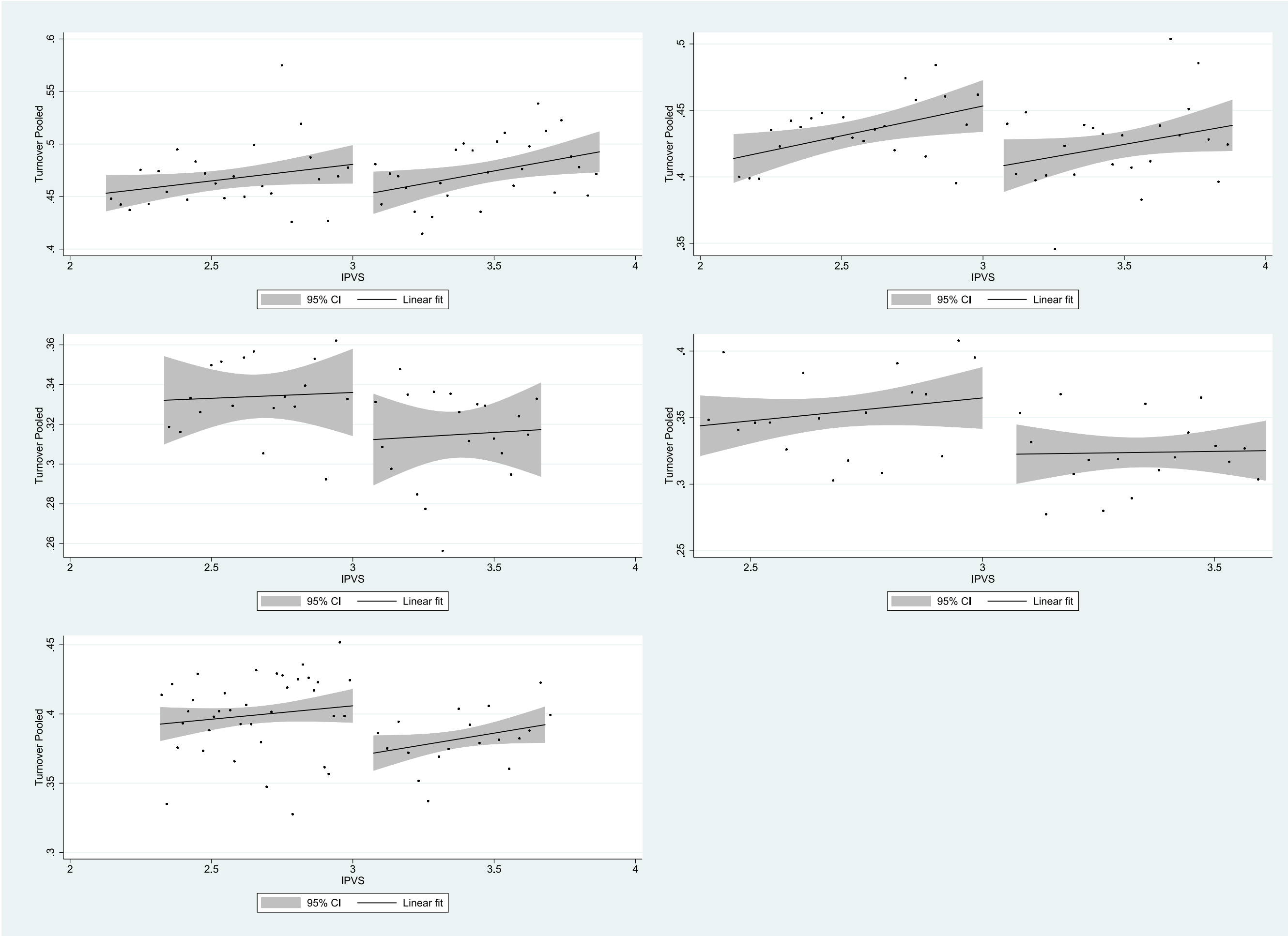

Fig. 4. Linear fit of post-treatment turnover rates by *IPVS*

cutoff point. Manipulation of the running variable is always a concern for RD designs. One can argue that schools had incentives to inflate their own *IPVS* scores in order to be included in *ALE*, so that school staff could earn higher wages. However, that is an implausible concern. The State Statistical Bureau (fully responsible for the index) is independent of the State Department of Education, and therefore has no relationship with schools. *IPVS* consists of variables beyond schools' control. Finally, the index was calculated in 2002, six years before its adoption as a criterion for *ALE*. Nevertheless, we test for manipulation. The histogram of schools by IPVS is presented at Figure 3. There is no density discontinuity at the cutoff. Moreover, the test proposed by CATTANEO, JANSSON, and A. (2015) shows no evidence of differences in the density of running variable at the cutoff (p-value 0.5475).

We also test for differences in the pre-treatment characteristics around the cutoff. Table 2 shows turnover rate; proportion of students retained in a grade, students with poor performance in Reading and Math in the *SARESP* exam; students' gender; parental schooling; and teachers' age, education, work status, and workload. The estimates show that treatment and control schools close to the threshold had similar turnover rates just before the treatment (2007-2008) and the difference is not significant.[15] Schools were also similar in their students' characteristics and in terms of teachers' profiles in 2007.[16]

All in all, the data suggest that the continuity assumption around the cutoff is valid.

5. Results

5.1. Impacts on Turnover

First, we evaluate the impacts of *ALE* on teacher turnover, which is the target effect of the policy. Figure 4 and Table 3 show the impacts of *ALE* on turnover rates during the first four years of the policy. The table shows the results for each year following the policy and also all years pooled. In general, the policy significantly impacted the turnover rates of treated schools in the years following its introduction. The additional wage brought by *ALE* reduced the turnover rate in 2008-2009 by 4.8 p.p. (optimal bandwidth). The effect becomes higher in the following years (except 2010-11), reaching 8.3 p.p. in 2011-2012. Pooling all years after the introduction of the policy, *ALE* diminished turnover by 5 p.p.

Considering that the average turnover rate for schools in the sample prior to the policy was 48.2%, *ALE* was responsible for reducing school turnover by 10.4% on average (pooled sample) and 17.2% in the last year of our sample. To put these figures into perspective, a 17.2% decrease in turnover corresponds to 3.2 fewer teachers leaving each treated school in a given year. Since the average monetary stimulus increased wages by 28%, the average turnover elasticity was around 0.4.

CLOTFELTER et al. (2008) find a larger turnover elasticity (around 4). It suggests that turnover in São Paulo is less sensitive to pay compared to North Carolina. Nevertheless, our results reinforce the idea that bonus payments for teachers in underprivileged schools have a

[15] Although the policy started in 2008, teachers could react to it only during the teacher allocation period for the next year after the end of school activities.

[16] See also Graphs A2 to A4 in the appendix.

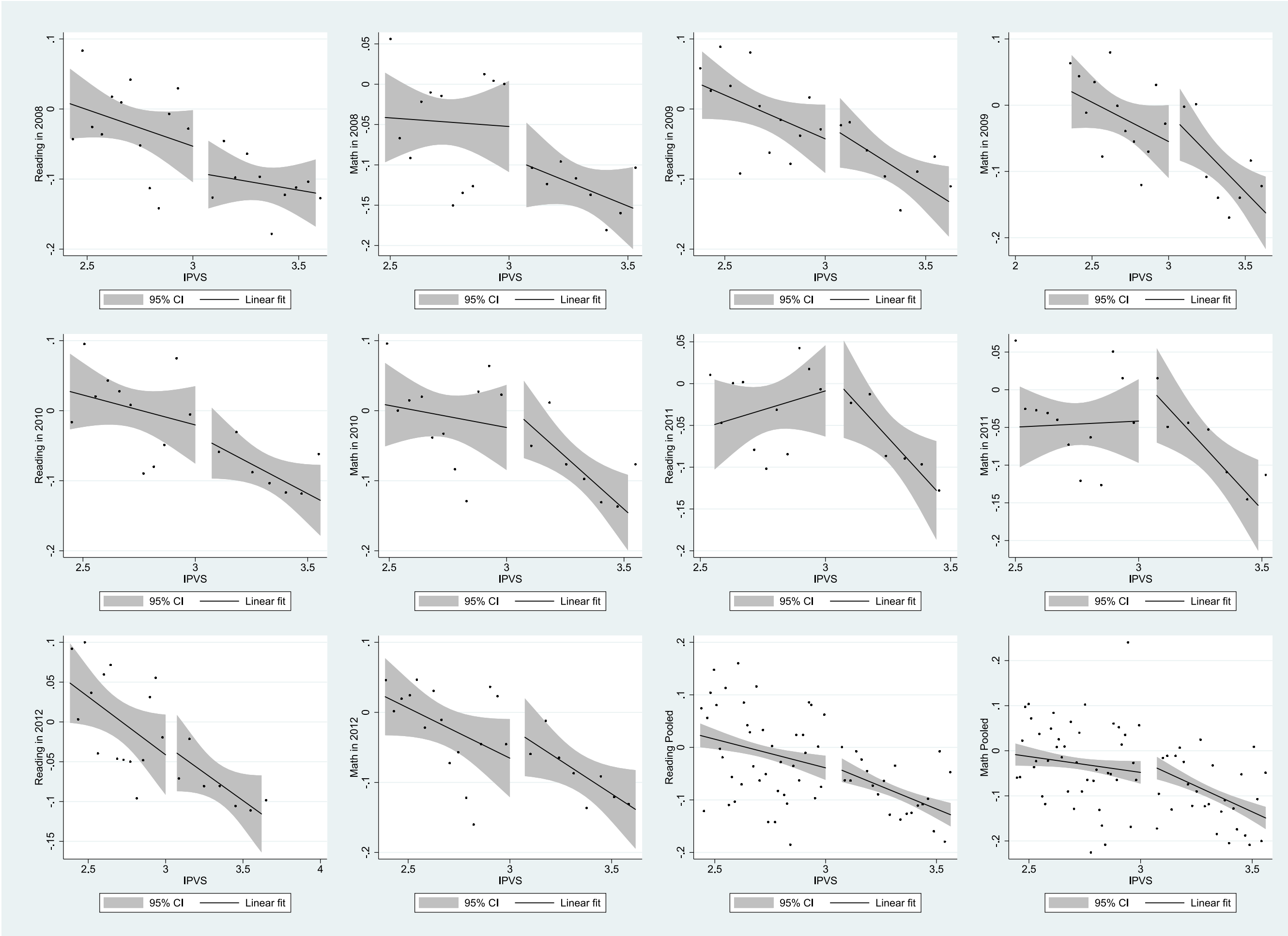

Fig. 5. Linear fit of post-treatment students' proficiency scores by *IPVS*

Table 3
Estimates of the impact of *ALE* on teacher turnover rates

VARIABLES	(1) 08_09	(2) 09_10	(3) 10_11	(4) 11_12	(5) Pooled
RD_Estimate	-0.048*	-0.064**	-0.028	-0.083**	-0.050**
	(0.026)	(0.028)	(0.030)	(0.034)	(0.024)
Effective n. of schools	1236	1249	887	821	3640
Bandwidth	0.873	0.882	0.668	0.611	0.682

Results from the fuzzy RD regression. *IPVS* is the running variable and *ALE* is the treatment variable. Bandwidth and effect. n. of schools chosen using CALONICO et al. (2014). Standard errors in parentheses. Clustered standard errors at the school level for pooled regression $***$ p<0.01, $**$ p<0.05, $*$ p<0.1.

meaningful impact on reducing turnover rates, even in a developing country context where the school environment is usually more adverse. Moreover, since teachers' payments in developing countries are lower, using wages to reduce turnover is likely to be less expensive compared to the United States or Western Europe.

5.2. Impacts on Proficiency

We further verify whether *ALE* had any impacts on student performance. Figure 5 and Table 4 show the reduced-form impact of *ALE* on student performance, measured as the average proficiency scores in Reading and Math on the state assessment. The impacts on average proficiency are small and statistically insignificant for both Reading and Math scores.

Based on the *SARESP* score, a student is classified as a low, basic, adequate or high performer in each subject (Math and Reading). A student's category is defined if her/his score is within specific score ranges in each grade and subject. For instance, a student is identified as a lower performer if his/her achievement in the *SARESP* exam is below a certain minimum level[17]. The proportion of low-performing students of a school in a given subject is calculated by the proportion of students in all grades that scored below the minimum proficiency threshold.

Turnover may be especially harmful for students with learning difficulties. We test this conjecture by running the RD approach on the proportion of low performing students in the school. Figure 6 and Table 5 show that *ALE* caused a reduction in the proportion of low performers in Reading and Math by 5.4 and 6.8 p.p., respectively. This impact is sizeable. The pre-treatment proportion of low performers in the treatment group was 27% in Reading and an astonishing 53% in Math. Therefore, *ALE* reduced the proportion of low-performing students by around 20% and 13% in Reading and Math, respectively. To put these effects in perspective, they represent a reduction in 46 (out of approximately 350) low-performing students in Math per school on average. If we extrapolate this calculation for the entire system, it would represent almost 250,000 fewer students with poor math performance in

[17] The minimum levels of proficiency for Reading are 125 in the 3rd grade, 150 in the 5th grade, and so on. The department of education sets the minimum standard of each grade and states that "students below that level demonstrate insufficient mastery of the contents and desirable skills for the school grade in which they are enrolled."

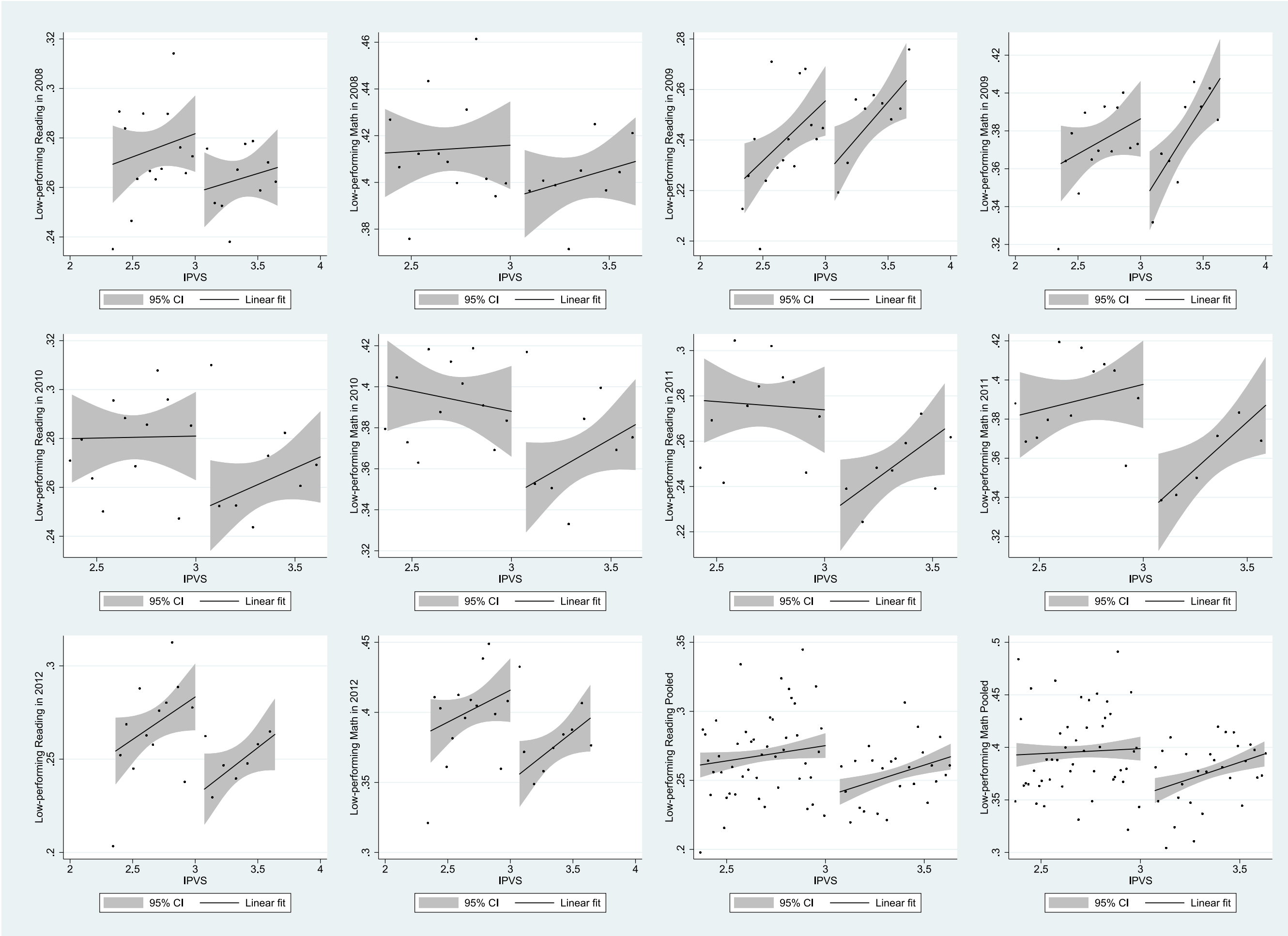

Fig. 6. Linear fit of post-treatment % of low performers by *IPVS*

Table 4
Estimates of the impact of *ALE* on students' average proficiency scores

	(1) 2008	(2) 2009	(3) 2010	(4) 2011	(5) 2012	(6) Pooled
	Math					
RD_Estimate	-0.090	0.110	-0.009	0.046	0.029	0.024
	(0.085)	(0.080)	(0.087)	(0.104)	(0.077)	(0.071)
Effective n. of schools	698	848	694	650	825	3762
Bandwidth	0.522	0.634	0.519	0.486	0.617	0.562
			Reading			
RD_Estimate	-0.063	0.053	-0.034	0.026	0.000	-0.002
	(0.073)	(0.073)	(0.075)	(0.107)	(0.068)	(0.064)
Effective n. of schools	782	826	748	576	829	3792
Bandwidth	0.581	0.616	0.558	0.444	0.621	0.567

Results from the fuzzy RD regression. *IPVS* is the running variable and *ALE* is the treatment variable. Bandwidth and effect. n. of schools chosen using CALONICO et al. (2014). Standard errors in parentheses. Clustered standard errors at the school level for pooled regression ∗ ∗ ∗ p<0.01, ∗ ∗ p<0.05, ∗ p<0.1.

a given year[18].

We further check whether *ALE* impacted the proportion of students in other categories (Basic, Adequate and High-performing). Table 6 shows the results for the pooled sample[19]. *ALE* reduces the proportion of students with low and basic performance and increased those with adequate and high performance. However, only the impact on the proportion of low-performing students is significant. The magnitude of the effect on the other categories is smaller and not significant.

Moreover, we tested for unequal impacts of *ALE* on different quantiles of the proficiency distribution. We followed the quantile regression discontinuity strategy proposed by FRANDSEN, FRöLICH, and MELLY (2012). We only find positive and significant impacts on schools at the bottom of the Math proficiency distribution (1st decile). Figure A5 in the appendix shows the *ALE* effects for different quantiles of the Math distribution.

Therefore, these findings could explain why the impact on % of low-performing is significant, but on average scores is not. Only schools with a high a fraction of students at the bottom of the score distribution seem to be significantly affected by the program. The impacts on proficiency seem to be more persuasive on schools with more students at the margin of the proficiency threshold between low and basic achievement.

It is hard to believe that *ALE* has a direct impact on low-performing students' effort, as it is aimed exclusively at teachers. Besides, we find no evidence of changes in the students' composition, since ALE has no impact on parental education.[20] However, raises in teacher salaries may impact students indirectly, via higher retention of teachers, attraction of better ones, or even improvements in their motivation. Therefore, the

[18] These findings go in line with those of RONFELDT et al. (2013) and HANUSHEK et al. (2016) that also show that turnover can be particularly harmful for disadvantaged students.

[19] The results of these regressions by year are available upon request.

[20] See Table A3 in the appendix.

Table 5
Estimates of the impact of *ALE* on the % of low performers

	Math					
	2008	2009	2010	2011	2012	Pooled
RD_Estimate	-0.023	-0.081**	-0.068*	-0.085**	-0.083**	-0.068**
	(0.033)	(0.039)	(0.039)	(0.042)	(0.040)	(0.033)
Effective n. of schools	844	854	836	798	853	4182
BW	0.633	0.636	0.623	0.594	0.640	0.624
			Reading			
RD_Estimate	-0.036	-0.050**	-0.051	-0.064*	-0.064**	-0.054**
	(0.026)	(0.025)	(0.032)	(0.035)	(0.031)	(0.025)
Effective n. of schools	873	865	840	747	852	4232
Bandwidth	0.659	0.649	0.628	0.558	0.639	0.634

Results from the fuzzy RD regression. *IPVS* is the running variable and *ALE* is the treatment variable. Bandwidth and effect. n. of schools chosen using CALONICO et al. (2014). Standard errors in parentheses. Clustered standard errors at the school level for pooled regression $***$ p<0.01, $**$ p<0.05, $*$ p<0.1.

Table 6
Estimates of the impact of *ALE* on the all categories of students - Pooled Sample

	(1) Low-performing	(2)	(3) Basic	(4)	(5) Satisfactory	(6)	(7) High-Performing	(8)
VARIABLES	Reading	Math	Reading	Math	Reading	Math	Reading	Math
RD_Estimate	-0.054**	-0.068**	-0.020	-0.028	0.015	0.039	0.015	0.014
	(0.025)	(0.033)	(0.017)	(0.019)	(0.023)	(0.028)	(0.012)	(0.010)
Effective n. of schools	4232	4182	4257	4247	4012	4082	6032	5972
Bandwidth	0.634	0.624	0.636	0.635	0.599	0.608	0.858	0.854

Results from the fuzzy RD regression. *IPVS* is the running variable and *ALE* is the treatment variable. Bandwidth and effect. n. of schools chosen using CALONICO et al. (2014). Standard errors in parentheses. Clustered standard errors at the school level for pooled regression $***$ p<0.01, $**$ p<0.05, $*$ p<0.1.

Table 7
Estimates of the impact of teachers' composition - characteristics

VARIABLES	(1) Teacher's Age	(2) Teacher's Experience	(3) Temporary Teachers	(4) Teacher's Workload	(5) Teacher's Gender	(6) Teacher's Tenure
RD_Estimate	-1.307**	-0.064	0.012	0.996	-0.022	0.348
	(0.551)	(0.458)	(0.031)	(0.799)	(0.059)	(0.726)
Effective n. of schools	1342	1322	1350	1082	1392	1370
Bandwidth	0.500	0.494	0.503	0.422	0.525	0.516

Results from the fuzzy RD regression. *IPVS* is the running variable and *ALE* is the treatment variable. Bandwidth and effect. n. of schools chosen using CALONICO et al. (2014). Standard errors in parentheses. Clustered standard errors at the school level for pooled regression $***$ p<0.01, $**$ p<0.05, $*$ p<0.1.

results correspond to the reduced-form impact of *ALE* on students.

5.3. Discussion

In the previous section, we show that a compensation policy for teachers in São Paulo's public schools reduces both teacher turnover and the proportion of low-performing students. When teachers leave a school, there is a loss of institutional memory related to teachers' knowledge about the school's routines, principals' and coordinators' methods of work, relationship with faculty, students' characteristics and other aspects that may affect teaching. However, *ALE* may also have impacted students' achievement through different channels. In this subsection, we discuss alternative explanations for the results and evidence that suggests these explanations are not plausible in our context. Considering the program's design, we discuss two alternative explanations for the reduction in the proportion of low-performing students: i) composition of teachers; ii) teachers' effort via wage change.

Monetary stimulus can also impact achievement through changes in the composition of teaching staff. In general, substitute/temporary teachers are less experienced and less effective than veteran ones (BOYD et al., 2008; CLOTFELTER et al., 2008; HANUSHEK et al., 2005; HANUSHEK et al., 2004). These teachers tend to be allocated to the most disadvantaged schools (BOYD et al., 2008; LOEB et al., 2012). CLOTFELTER et al. (2008) show that experienced teachers respond more to financial incentives than inexperienced ones. PUGATCH and SCHROEDER (2014) also find that a hardship allowance program in Gambia attracted more qualified teachers to treated schools. ADNOT, DEE, KATZ, and WICKOFF (2017) also show that teacher turnover can have positive effects on learning if low-performing teachers can be replaced.

Therefore, we first investigate if the program affected some demographic characteristics of the teachers such as age, tenure, workload or contract type. We further analyze whether *ALE* impacted the composition of teachers in treated schools beyond observable characteristics. We calculated the average teacher's value-added, as a measure of teacher quality, and estimate the impact of the *ALE* program on it. Although the methods for doing this are controversial (ROTHSTEIN, 2017; CHETTY, FRIEDMAN, & ROCKOFF, 2014; CHETTY, FRIEDMAN, & ROCKOFF, 2017), we follow some of the most traditional models, adapted to available data:

$$\overline{A}_{tcgi} = \lambda \overline{A}_{cg-2i} + \alpha_1 \overline{C}_{tcgi} + \alpha_2 H_{tcgi} + \mu_i + \epsilon_{tcgi}$$

in which $\overline{A}_{tcgi}$ is the average student test score in class c where teacher t

Table 8
Estimates of the impact of teachers' composition - value added

VARIABLES	(1) Teacher's VA - Pooled	(2) Moving Teacher's VA - Pooled
RD_Estimate	-0.009	0.149
	(0.077)	(0.133)
Effective n. of schools	1502	1472
Bandwidth	0.562	0.580

Results from the fuzzy RD regression. *IPVS* is the running variable and *ALE* is the treatment variable. Bandwidth and effect. n. of schools chosen using CALONICO et al. (2014). Standard errors in parentheses. Clustered standard errors at the school level $***$ p<0.01, $**$ p<0.05, $*$ p<0.1.

teaches grade g at school i. $\overline{A}_{cg-2i}$ is the counterpart for the last score assessed of the same class of students two years earlier,[21] $\overline{C}_{tcgi}$ is a vector of averaged student characteristics (gender, age, parents' education), H_{tcgi} are teachers' characteristics (age, tenure, type of contract and career level), and μ_i is a school fixed effect. The value-added is measured by the normalized residual ($\widehat{\epsilon_{tcgi}}$) of this regression. The available databases allow the estimation of this model for one year only, 2007, so we obtain the value-added for each teacher t in this year. Then, we assign these 2007 values for the same teachers in the following years, 2008 and 2009. Finally, we take the average of this variable by school to obtain a measure of (average) quality of teachers working in that school.[22]

As the state official exam focuses on math and language, the measure of teacher quality we use refers to the school average value-added of language and math teachers found in the databases from 2007 to 2009. In this sense, we are not considering the whole population of teachers. On the other hand, since we evaluate effects on students' performance in language and math, this measure of quality should, in fact, be most associated to our achievement variable.

Table 7 shows the impact *ALE* on teachers' characteristics pooling 2008 and 2009.[23] *ALE* did not significantly change the average teacher's experience, workload, gender and tenure. However, treated schools received younger teachers on average (about 1.3 year younger). Since average age is around 43 years old, this effect represents a decrease in 3% in average teacher's age. Table 8 shows the absence of impact of the *ALE* program on teachers' value-added[24]. We also tested if *ALE* impacted moving teachers differently, but we could not find significant results either. It reinforces the idea that the program affected students' proficiency only through turnover.

All in all, it is hard to infer much about the quality of the teachers that remain in treated schools. So far, we can only say that *ALE* attracted slightly more young teachers. Therefore, it is even harder to believe that those small changes in age are responsible for the reduction in low-performing students.

ALE may impact teachers' effort due to the increase in their earnings. Better-paid teachers may dedicate more time to prepare classes and other learning activities. Although there is no strong evidence of this channel in the literature (e.g. REE et al., 2018), we test whether wage increases have a direct impact on achievement. We also check if the wage increase impacted the teachers' workload as a measure of effort dedicated to the treated schools. We explore differences on the importance of *ALE* on teachers' earnings to check the validity of this possible explanation. As discussed, the weight of *ALE* on teachers' earnings is determined by tenure and career. We run a reduced-form regression including a variable that captures how important *ALE* was in a teacher's remuneration:

Table 9
Estimates of the impact of *ALE* on % low performers and teacher's workload through changes in teacher salary

	(1) Reading	(2) Math	(3) Teacher's Workload
V_i (*ALE weight on wage*)	-0.527	-0.678	102.124***
	(0.375)	(0.506)	(11.867)
$\mathbf{Z_i \times V_i}$ (*interaction*)	**-0.250**	**-0.339**	**-3.968**
	(0.496)	(0.646)	(17.793)
$Z_i(1_{IPVS\geq 3})$	0.044	0.045	0.179
	(0.130)	(0.169)	(4.656)
Constant	0.318**	0.467***	-2.203
	(0.125)	(0.167)	(3.998)
R-squared	0.099	0.134	0.307
Effective n. of schools	2010	1945	1579
Bandwidth	0.526	0.508	0.422

Results from the fuzzy RD regression. Pooled sample. Non-moving teachers only. *IPVS* is the running variable, Z_i is the indicator variable that *IPVS* ≥ 3 Bandwidth and effect. n. of schools chosen using CALONICO et al. (2014). V_i is the average proportion of ALE received over wages. Clustered standard errors by school in parentheses. Additional Controls: Teachers' age, gender, experience, tenure, and if s/he has temporary contract $***$ p<0.01, $**$ p<0.05, $*$ p<0.1.

$$E_i = \beta_0 + \beta_1(X_i - 3) + \beta_2(X_i - 3)Z_i + \beta_3 Z_i + \beta_4 V_i + \beta_5 Z_i V_i + \gamma' M_i + \upsilon_i.$$

where E_i is the proportion of low-achievers or the teachers' workload. To capture the effect of the wage increase on E_i, we include in the regression the variable V_i, the average value (as a percentage of the base wage) of *ALE* received by the school's teachers. As teachers' characteristics determine this value[25], we can calculate it even for control schools' teachers. In this sense, V_i measures an average potential weight of *ALE* on the wage of a school's teachers. Therefore, the interaction Z_iV_i captures the effective weight of *ALE* for schools above the threshold of the running variable.

The idea is that if (unconditional) money were important for teachers' effort and students' achievement, it would affect treated schools differently based on the average amount received by their teachers[26]. Therefore, our specification measures how different the *ALE* impact is to different schools with varying values of V_i.

A major concern is that tenure, and consequently wages, is correlated to a teacher's experience. Additionally, CLOTFELTER et al. (2008) show that experienced teachers tend to respond more to financial incentives to reduce turnover. Therefore, we restrict the sample to teachers that had not moved across schools between 2008 and 2009, so that we avoid confounding with any composition/turnover effect.[27] We also directly control for the teachers' characteristics by adding the vector M_i that includes age, gender, experience, tenure and if s/he is a temporary teacher.

In this exercise, we are interested in β_5. If this parameter is significant, we have evidence that, controlling for the characteristics that

[21] We exclude the 3rd graders of the sample and gather data from the 2005 *SARESP* to calculate $\overline{A}_{cg-2i}$.

[22] We did the same procedure for teachers in following years of our sample (2010 to 2012). However, since the value-added measure is calculated in 2007, it becomes outdated. Besides, since many new teachers enter in the system, the average school valued added is calculated in smaller samples as time goes by. Nevertheless, the results are qualitatively the same and are available upon request.

[23] We only have information on teachers' characteristics in 2007, 2008 and 2009.

[24] The results are robust to different specifications of equation 5.3. Excluding the vector of teachers' or students' characteristics does not change the main conclusions. Results are available upon request.

[25] We calculate it based on 2007 characteristics (before *ALE*).

[26] Actually, we are testing if the extra payment proportionally to the base salary affects effort.

[27] Since the value of the compensation depends on workload in treated schools, another concern is that *non-movers* could reallocate their workload toward those schools. We do not find any evidence of that. The impact is statistically insignificant and numerically smaller than for the *movers* discussed above. This result is available upon request.

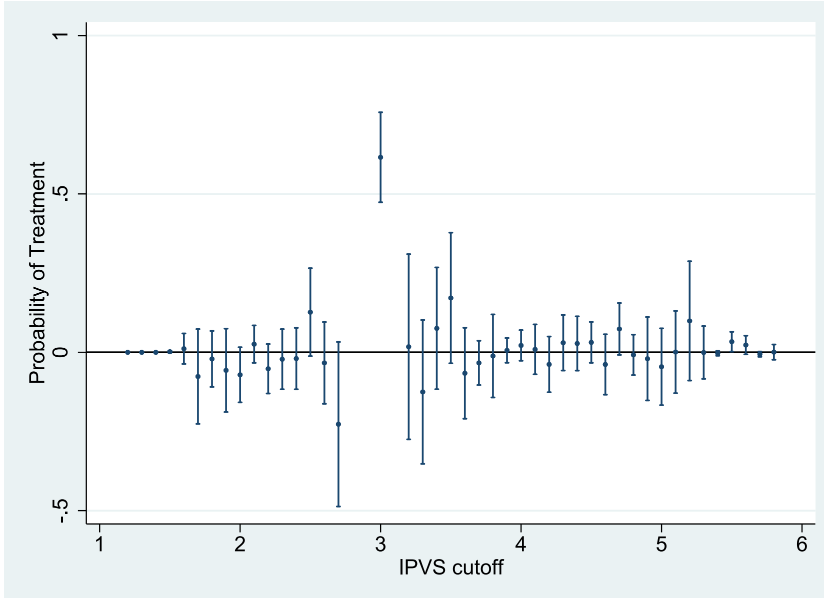

Fig. 7. Estimates and confidence intervals for discontinuities at probability of treatment by *IPVS* values

determine a teacher's wage, the value transferred by *ALE* impacts students' performance.

Table 9 shows results for the wage increase. Of interest first, the estimate of the coefficients associated to the interactions $Z \times V$ are insignificant in both years.[28] So, it does not seem that the reduction of low-performing students can be attributed to a wage increase. This result is not surprising, as the financial incentive received by the teacher is unrelated to any productivity measure. It is unlikely that teachers changed their effort levels because of the program.

One can argue that *ALE* raised the average numbers of hours teachers devote to treated schools, and that could be a measure of their productivity. However, this effect only indicates that teachers get more classes in treated schools, but it does not mean that students had more instructional time or even that teachers exerted more effort on those classes.

Another concern has to do with concurrent policies that may have affected schools at the same period of *ALE*. Indeed, two other programs were introduced to São Paulo Schools near *ALE*: (i) a pay-for-performance program, and (ii) a school management training program to principals.

In 2008, the State of São Paulo introduced a pay-for-performance scheme to its schools. The bonus payments were given to teachers, principals, and staff of all schools that reached determined targets based on an index that combines proficiency test scores and pass rates (*IDESP*). All schools were eligible to the incentive program. Moreover, the targets were based on previous *IDESP* levels in a continuous way.[29] Therefore, the performance-based incentive does not confound with *ALE* at the cutoff point. However, it possible that the existence of a pay-for-performance incentive may have boosted the *ALE* impact on proficiency. Unfortunately, since the performance-based program eligibility is universal, we cannot check whether it increased *ALE* effects.

In 2007, the State of São Paulo also started a school management training program to principals in low-performing schools. Treated schools' principals received managerial instructions and feedbacks from the Department of Education about their management practices (see TAVARES (2015) for details). The program was available to the bottom 5% schools in the *IDESP*. Therefore, the discontinuity associated with *ALE* is not related to the discontinuity of the management training program. Moreover, there is no treated schools of the management program in the effective samples of our exercises (within the optimal bandwidths suggested by CALONICO et al. (2014)).

In brief, we conclude that a compensation policy like *ALE*, that seeks to reduce teacher turnover in high-poverty schools, seems to, in fact, decrease the proportion of teachers moving from those schools. This impact seems to generate, at least in a medium term, reduction in the proportion of low-performing students. Our findings also suggest that the composition of teachers or the wage increase were not the cause of improved student performance. Therefore, it is plausible to conjecture that the reduction in the disruptive mechanism of teacher turnover is behind the improvement of low-performing students' proficiency in Math and Reading.

6. Robustness tests

We submit our estimates to a few kinds of robustness tests: (i) variations on econometric specification and sample; (ii) placebo tests; and (iii) external validity.

Tables A4 and A5 in the appendix reproduce the results of the paper on turnover and low performers using different bandwidth and kernel selection. Most results are robust to these changes in model specification. Qualitatively, the results hold to different specifications. Although some results become statistically insignificant for a narrower bandwidth choices, the magnitude of the impact is still sizeable.

Schools with an integer *IPVS* (exact 1, 2, 3, 4, 5, or 6) are on homogeneous regions (*IPVS* of census tracts around the school are identical). Therefore, the socioeconomic status of students enrolled in those schools may be more homogeneous and similar to the socioeconomic status of their teachers. These facts could impact turnover and students' performance. Therefore, we replicate our principal analysis, excluding schools with integer *IPVS*. There are 288 (9%) eligible schools with an integer *IPVS*. Table A6 in the appendix shows that the results are robust to the exclusion of such schools.

As for the placebo tests, Graph 7 shows estimates for the jump in the probability of treatment at a series of *IPVS* values, along with their confidence intervals. The only significant increase in treatment

[28] Notice that the numerical results in table 9 can not be directly compared to the previous average effects (Table 7), since here we present reduced-form regressions with additional control variables. The average effect on teachers' workload using a similar specification as in table 9 is -0.84.

[29] The target is a logistic function of the school's previous year *IDESP*.

Table 10
Dif-in-Dif Strategy for different intervals of *IPVS* - Impact on Turnover

VARIABLES	(1) IPVS ∈ [1.0-2.0]	(2) IPVS ∈ [2.0-3.0]	(3) IPVS ∈ [3.0-4.0]	(4) IPVS ∈ [4.0-5.0]	(5) IPVS ∈ [5.0-6.0]
$Post_tE_i$	-0.003	-0.006	-0.040***	-0.072***	-0.071***
	(0.016)	(0.008)	(0.009)	(0.011)	(0.019)
N. of schools	1,890	5,740	4,855	3,300	1,050

Results from Diff-in-diff regressions. Pooled sample. E_i indicates eligible schools. Clustered standard errors at the school level in parentheses *** p<0.01, ** p<0.05, * p<0.1

assignment is at 3.0, which reinforces the choice of this cutoff.

Finally, we use the eligibility rule (only schools in municipalities of the São Paulo metropolitan area) to estimate the impacts of *ALE* at different points of the *IPVS* distribution by a difference-in-difference strategy. We pool all years and schools (eligible and non-eligible) and run the following regression on different intervals of the *IPVS* distribution:

$$Y_{it} = \beta_0 + \beta_1 E_i + \beta_2 Post_t + \beta_3 Post_t E_i + \beta_4 IPVS_i + \beta_5' X_{it} + u_{it}$$

where E_i indicates if school i is eligible for *ALE* (i.e., it is located in an eligible municipality), $Post_t$ indicates if the year is 2008 or after, *IPVS* is the vulnerability index of school i in 2000, X_{it} is a vector of control variables, and u_{it} is an idiosyncratic error. Standard errors are clustered at the school level. The parameter of interest is β_3.

Table 10 shows the results of the difference-in-differences strategy for several intervals of the *IPVS*. As expected, only intervals where *ALE* was actually implemented show negative and significant impacts on turnover. Columns (1) and (2) show the results for the least vulnerable schools in the intervals [1-2] and [2-3] of the *IPVS*. In these intervals, very few eligible school received *ALE*. Seven schools in the first interval (2%) and forty schools in the second (5%) were in the program. We find no impact of the eligibility on turnover. Column (3) shows no impact of the eligibility for schools just after the cutoff of vulnerability ([3-4]). In this interval, 63 (9%) eligible schools did not receive the program. For this group, the impact on turnover is 4 p.p. and statistically significant throughout the period. Columns (4) and (5) show the impact at intervals where almost all eligible schools were treated - only 14 (3%) and 3 (2%) eligible schools did not receive the program in the [4-5] and [5-6] intervals, respectively. In these cases, the negative impact is also significant. Moreover, the magnitudes are similar to those found in the RDD strategy. The outcomes are qualitatively similar over all intervals, but increase (in absolute terms) for more vulnerable schools. The increase is probably related to the fact that there are proportionally more schools treated in the interval furthest away from the cutoff.

7. Conclusion

This paper evaluates the impacts of a wage compensation policy intended to reduce teacher turnover in São Paulo's state school system (the largest public school system in Brazil). Teachers allocated to disadvantaged schools received a wage premium. Beneficiary schools were chosen based on a socioeconomic index beyond the control of schools. This enables a regression-discontinuity design that allows us to estimate the causal impact of this policy.

We conclude, from our benchmark models, that this extra payment reduced teacher turnover by up to 8.3 percentage points, which means a drop of 16% compared to the pre-treatment average. Results are robust to other parametric and nonparametric specifications.

We also find evidence that this policy impacted students' performance. In a reduced-form model, the wage premium seems to impact the proportion of low performers in the treated schools. However, we have no evidence that this intervention attracted or retained better quality teachers in treated schools nor that the wage increase itself is responsible for the improvement on students achievement. Therefore, we suggest that the disruptive effect of turnover is reduced by the remuneration policy.

Unlike previous work, our analysis focuses on schools in vulnerable regions in a developing country. Students in impoverished schools typically come from disturbed families with low levels of parental education and commonly suffer from psychological and physical violence. In such an environment, the relationship between teachers and students is essential and goes beyond the learning process. The course of gaining confidence and building a trusting relationship takes time. Frequent changes in the teaching and support staff thwarts stable and reliable connections between teachers and students. Therefore, we conjecture that the disruptive effect in this setting is also critical for students' self-esteem and ultimately for their achievement.

Finally, our analysis is robust to different specifications. Moreover, external validity, which is a caveat in most studies that rely on RDD strategies, seem not to be a problem in our case. Using the eligibility rule by location allows us to estimate the impact of the program in different intervals of the running variable. The results are qualitatively similar and show the reduction on turnover for schools at different levels of vulnerability.

Therefore, this paper reinforces that policies of compensatory payments can influence the mobility of teachers, reducing turnover in deprived schools. We also conclude that the reduction of the disruptive effect of turnover is key for students' achievement.

CRediT authorship contribution statement

Rafael Camelo: Conceptualization, Methodology, Formal analysis, Investigation, Writing - original draft, Writing - review & editing. **Vladimir Ponczek:** Conceptualization, Methodology, Formal analysis, Investigation, Writing - original draft, Writing - review & editing.

Appendix

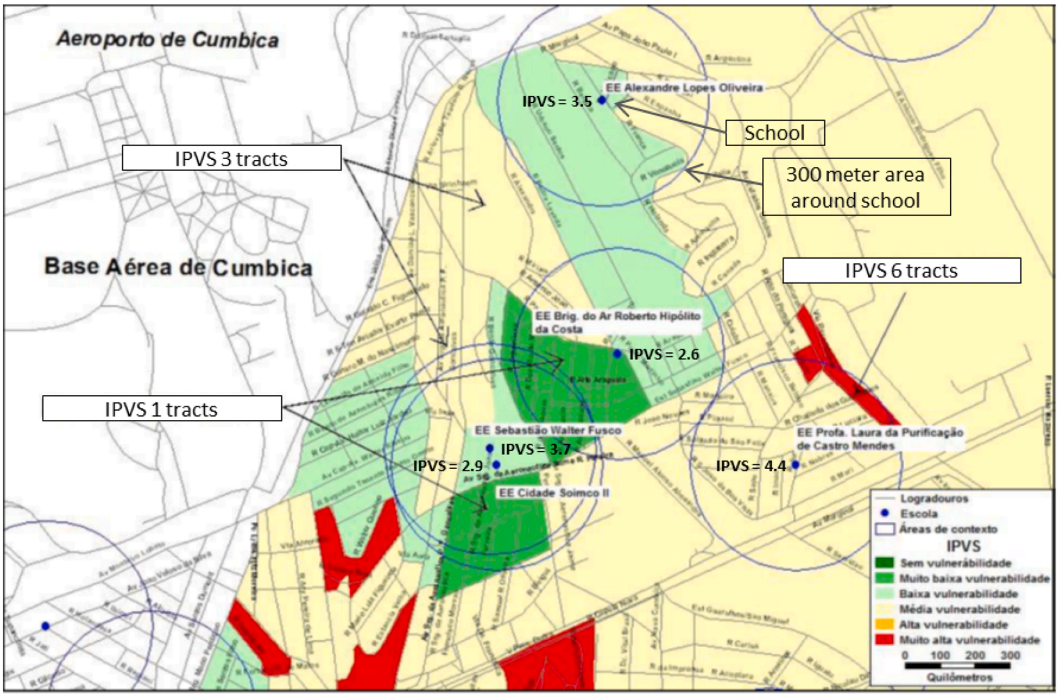

Fig. A1. Map containing census tracts, schools and their respective *IPVS* values

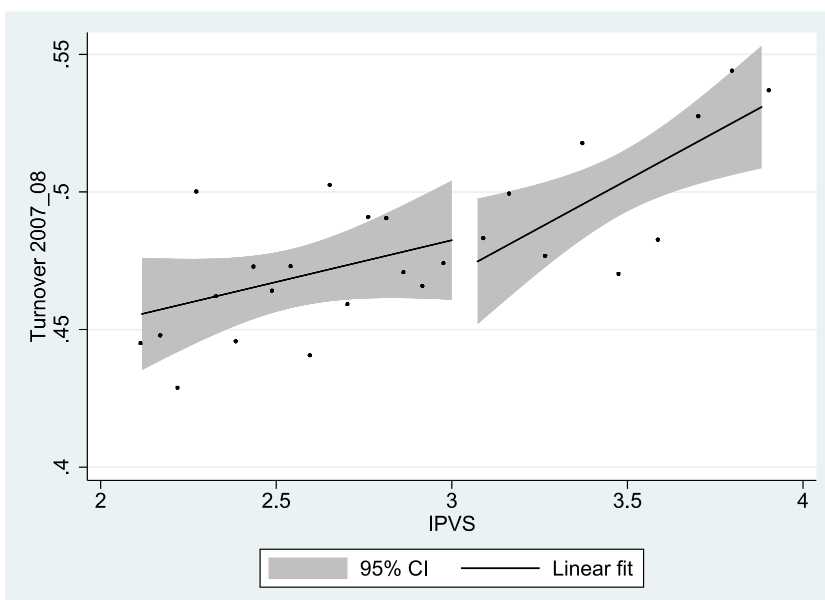

Fig. A2. Linear fit of pre-treatment turnover rate by *IPVS*

Table A1
Teachers' wages in 2008 US Dollars

Career levels	Elementary school teachers					Secondary and high school teachers				
Yrs. of tenure	I	II	III	IV	V	I	II	III	IV	V
[0, 5)	4.46	4.69	4.92	5.17	5.42	5.17	5.42	5.69	5.98	6.28
[5, 10)	4.69	4.92	5.17	5.42	5.69	5.42	5.69	5.98	6.28	6.59
[10, 15)	4.92	5.17	5.42	5.69	5.98	5.69	5.98	6.28	6.59	6.92
[15, 20)	5.17	5.42	5.69	5.98	6.28	5.98	6.28	6.59	6.92	7.27
[20, 25)	5.42	5.69	5.98	6.28	6.59	6.28	6.59	6.92	7.27	7.63
[25, 30)	5.69	5.98	6.28	6.59	6.92	6.59	6.92	7.27	7.63	8.01

For currency conversion, we use the 2008 average exchange rate: 0.5453 US$ /R$. Teachers in São Paulo public schools usually retire after 30 years of work, for men, and 25 years, for women. Less than 1% of teachers have more than 30 years of tenure.

Table A2
IPVS scale and vulnerability groups

IPVS scale	
1	No vulnerability
2	Very low vulnerability
3	Low vulnerability
4	Medium vulnerability
5	High vulnerability
6	Very high vulnerability

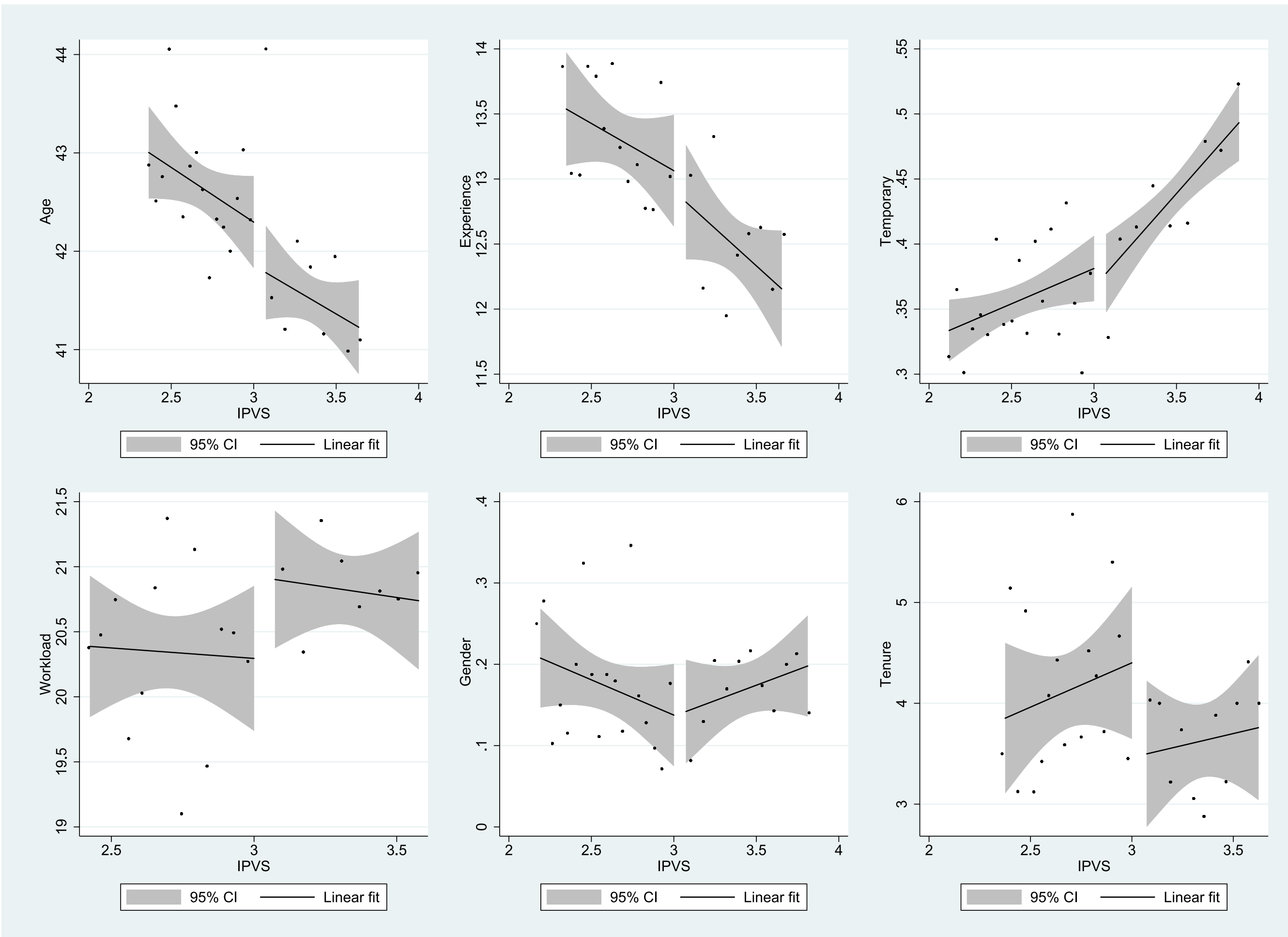

Fig. A3. Linear fit of pre-treatment teachers' characteristics (2007) by *IPVS*

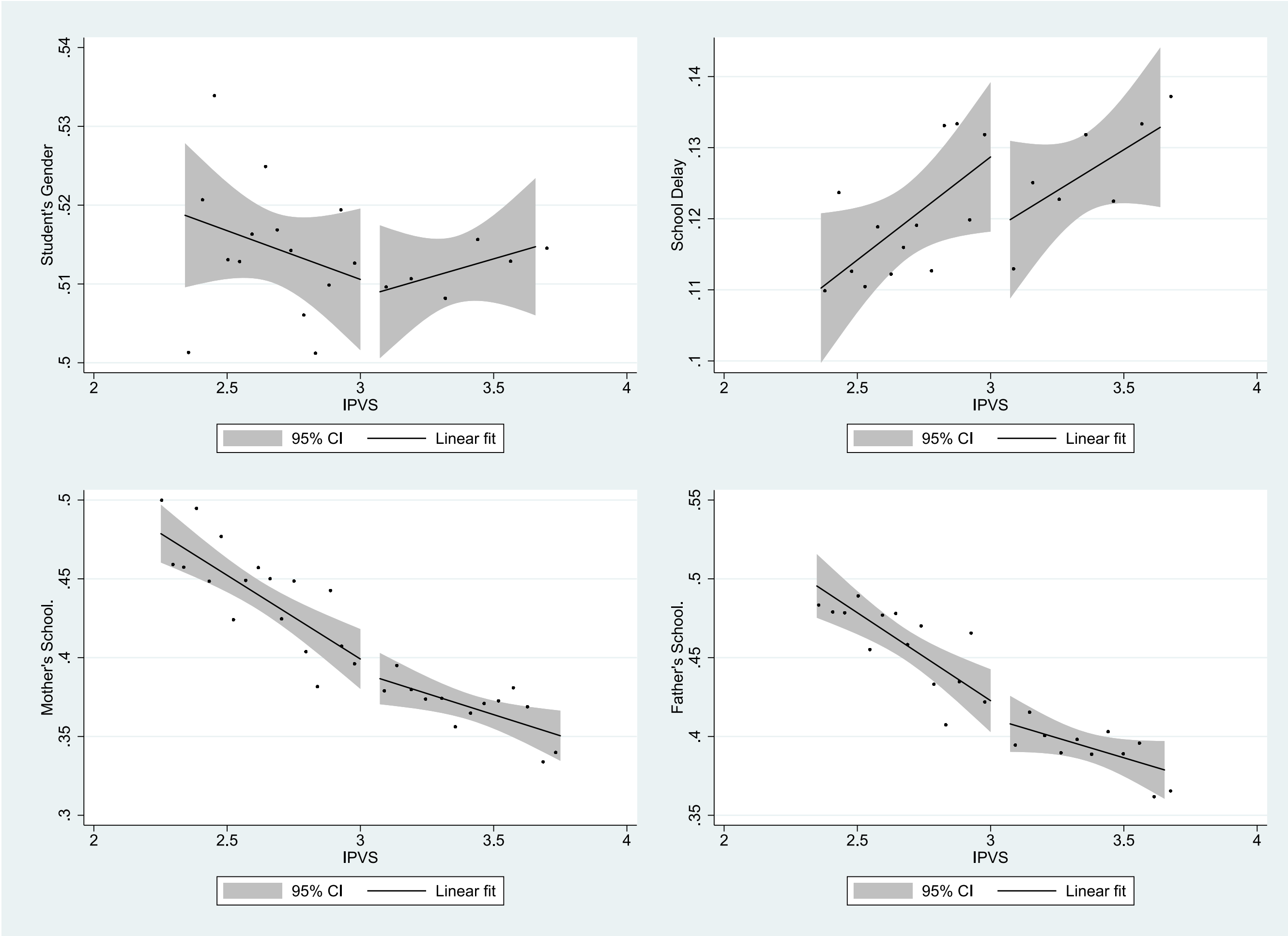

Fig. A4. Linear fit of pre-treatment students' characteristics (2007) by *IPVS*

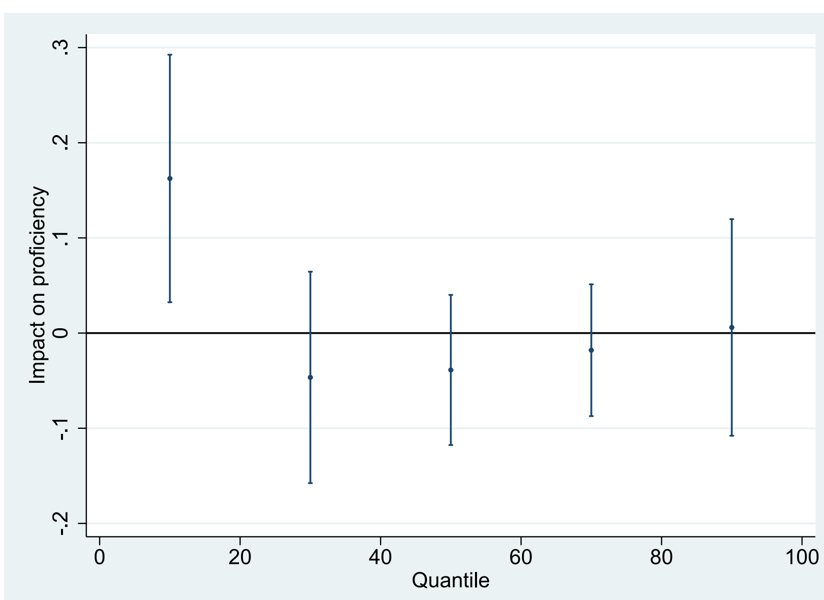

Fig. A5. Quantile Treatment Effect by Proficiency Score

Table A3
Estimates of the impact of ALE on Parental Education

	2008 Mother's Schooling	2008 Father's Schooling	2009 Mother's Schooling	2009 Father's Schooling	2010 Mother's Schooling	2010 Father's Schooling	2011 Mother's Schooling	2011 Father's Schooling	2012 Mother's Schooling	2012 Father's Schooling	Pooled Mother's Schooling	Pooled Father's Schooling
RD_Estimate	-0.016	-0.014	0.008	0.003	-0.002	-0.014	0.003	-0.012	0.018	0.002	0.003	-0.007
	(0.024)	(0.022)	(0.021)	(0.019)	(0.024)	(0.022)	(0.023)	(0.023)	(0.027)	(0.025)	(0.023)	(0.020)
Effective n. of schools	836	877	1167	1122	869	853	1038	852	862	813	5142	5290
Bandwidth	0.625	0.663	0.833	0.810	0.657	0.637	0.753	0.636	0.655	0.613	0.749	0.779

Results from the fuzzy RD regression. *IPVS* is the running variable and *ALE* is the treatment variable. Bandwidth and effect. n. of schools chosen using CALONICO et al. (2014) - CCK. Standard errors in parentheses. Clustered standard errors at the school level for pooled regression $***$ p<0.01, $**$ p<0.05, $*$ p<0.1.

Table A4
Estimates of the impact of *ALE* on teacher turnover rates - Alternative Specifications

	(1) 0.5 x CCK Pooled	(2) 2 x CCK Pooled	(3) CCK - Rect. Kernel Pooled
RD_Estimate	-0.050**	-0.050**	-0.052**
	(0.024)	(0.024)	(0.022)
Effective n. of schools	3640	3640	3504
Bandwidth	0.682	0.682	0.659

Results from the fuzzy RD regression. *IPVS* is the running variable and *ALE* is the treatment variable. Bandwidth and effect. n. of schools chosen using CALONICO et al. (2014). Standard errors in parentheses. Clustered standard errors at the school level $***$ p<0.01, $**$ p<0.05, $*$ p<0.1.

Table A5
Estimates of the impact of *ALE* on % of low performers - Alternative Specifications

VARIABLES	(1) 0.5 x CCK Pooled Reading	(2) 2 x CCK Pooled Reading	(3) CCK - Rect. Kernel Pooled Reading	(4) 0.5 x CCK Pooled Math	(5) 2 x CCK Pooled Math	(6) CCK - Rect. Kernel Pooled Math
RD_Estimate	-0.010	-0.061***	-0.063***	-0.009	-0.072***	-0.077**
	(0.050)	(0.015)	(0.024)	(0.068)	(0.019)	(0.032)
Effective n. of schools	1911	10102	3792	1876	10012	3702
Bandwidth	0.317	1.268	0.566	0.312	1.248	0.554

Results from the fuzzy RD regression. *IPVS* is the running variable and *ALE* is the treatment variable. Bandwidth and effect. n. of schools chosen using CALONICO et al. (2014). Standard errors in parentheses. Clustered standard errors at the school level $***$ p<0.01, $**$ p<0.05, $*$ p<0.1.

Table A6
Estimates of the impact of *ALE* - Non-Integer *IPVS* schools only

VARIABLES	(1) Turnover	(2) Low-performing Reading	(3) Low-performing Math
RD_Estimate	-0.051**	-0.053**	-0.066*
	(0.024)	(0.025)	(0.035)
Effective n. of schools	3752	4007	3942
Bandwidth	0.699	0.598	0.584

Results from the fuzzy RD regression. Schools with an integer *IPVS* were dropped. *IPVS* is the running variable and *ALE* is the treatment variable. Bandwidth and effect. n. of schools chosen using CALONICO et al. (2014). Standard errors in parentheses. Clustered standard errors at the school level $***$ p<0.01, $**$ p<0.05, $*$ p<0.1.

References

ADNOT, M., DEE, T., KATZ, V., & WICKOFF, J. (2017). Teacher turnover, teacher quality, and student achievement in DCPS. *Educational Evaluation and Policy Analysis, 54–76*.39-1

ALLENSWORTH, E., PONISCIAK, S., & MAZZEO, C. (2009). The schools teachers leave: Teacher mobility in chicago public schools. Chicago: Consortium on Chicago School Research - University of Chicago.

BOYD, D., GROSSMAN, P., LANKFORD, H., LOEB, S., & WYCKOFF, J. (2008). Who leaves? teacher attrition and student achievement. *NBER Working Paper*, 14022.

CABRERA, J. M., & WEBBINK, D. (2019). Do higher salaries yield better teachers and better student outcomes? *Journal of Human Resources*.Forthcoming

CALONICO, S., CATTANEO, M. D., & TITIUNIK, R. (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica, 82*, 2295–2326.

CARVER-THOMAS, D., & DARLING-HAMMOND, L. (2017). Teacher turnover: Why it matters and what we can do about it. palo alto. *CA: Learning Policy Institute*.

CASTRO, J., & ESPOSITO, B. (2018). The effect of bonuses on teacher behavior: A story with spillovers. *Working paper 104, Peruvian Economic Association*.

CATTANEO, M., JANSSON, M., & A., X. M. (2015). Simple local regression distribution estimators with an application to manipulation testing. *Working Paper*.

CLOTFELTER, C., GLENNIE, E., LADD, H., & VIGDOR, J. (2008). Would higher salaries keep teachers in high-poverty schools? evidence from a policy intervention in north carolina. *Journal of Public Economics, 92*, 1352–1370.

COWAN, J., & GOLDHABER, D. (2018). Do bonuses affect teacher staffing and student achievement in high poverty schools? evidence from an incentive for national board certified teachers in washington state. *Economics of Education Review, 65*, 138–152.

DEE, T., & WICKOFF, J. (2015). Incentives, selection, and teacher performance: Evidence from IMPACT. *Journal of Policy Analysis and Management, 34*(2), 267–297.

FIGLIO, D., & KENNY, L. (2007). Individual teacher incentives and student performance. *Journal of Public Economics, 91*, 901–914.

FRANDSEN, R. B., FRöLICH, M., & MELLY, B. (2012). Quantile treatment effects in the regression discontinuity design,. *Journal of Econometrics*, 382–395.168-2

HAHN, J., TODD, P., & KLAAUW, W. V. D. (2001). Identification and estimation of treatment effects with a regression discontinuity design. *Econometrica, 69*, 201–209.

HANUSHEK, E., KAIN, J., O'BRIEN, D., & RIVKIN, S. (2005). The market for teacher quality. *NBER Working Paper*, 11154.

HANUSHEK, E., KAIN, J., & RIVKIN, S. (2004). Why public schools lose teachers. *Journal of Human Resources, 39*, 326–354.

HANUSHEK, E., RIVKIN, S., & SCHIMAN, J. (2016). Dynamic effects of teacher turnover on the quality of instruction. *Economics of Education Review, 55*, 132–148.

HEMPHILL, C., & NAUER, K. (2009). The new marketplace: How small-school reforms and school choice have reshaped new york city high schools. new york: Center for new york city affairs. *The New School for Management and Urban Policy*.

INGERSOLL, R. M. (2001). Teacher turnover and teacher shortages: An organizational analysis. *American Educational Research Journal, 38*(3), 499–534.

JOHNSON, S. M., BERG, J. H., & DONALDSON, M. L. (2005). Who stays in teaching and why? a review of the literature on teacher retention. Washington, DC: National Retired Teachers Association.

KASMIRSKI, P. (2012). Mobilidade de professores na rede estadual paulista. *Retrieve from The Digital Library of Theses and Dissertations of the University of SÃo Paulo*.

LOEB, S., BETEILLE, T., & KALOGRIDES, D. (2012). Effective schools: Teacher hiring, assignment, development, and retention. *Education Finance and Policy, 7*, 269–304.

SIMON, N, & JOHNSON, S (2015). Teacher turnover in high-poverty schools: What we know and can do. *Teachers College Record, 117*(3).

MARINELL, W. H., & COCA, V. M. (2013). Who stays and who leaves? findings from a three part study of teacher turnover in NYC middle schools. *New York: The Research Alliance for NYC Schools*.

PUGATCH, T., & SCHROEDER, E. (2014). Incentives for teacher relocation: Evidence from the gambian hardship allowance. *Economics of Education Review, 41*, 120–136.

ROSEN, SHERWIN (1986). The theory of equalizing differences. In Orley Ashenfelter, & Richard Layard (Eds.), *1. Handbook of Labor Economics* (pp. 641–692). Elsevier.

REE, J, MURALIDHARAN, K, PRADHAN, M, & ROGERS, H (2018). Double for nothing? Experimental Evidence on an Unconditional Teacher Salary Increase in Indonesia. *The Quartely Journal of Economics, 133*(2), 993–1039.

CHETTY, R, FRIEDMAN, J, & ROCKOFF, J (2017). Measuring the Impacts of Teachers: Reply. *American Economic Review, 107*(6), 1685–1717.

ROTHSTEIN, J (2017). Measuring the Impacts of Teachers: Comment. *American Economic Review, 107*(6), 1656–1684.

RONFELDT, M., LOEB, S., & WYCKOFF, J. (2013). How teacher turnover harms student achievement. *American Educational Research Journal, 50*, 4–36.

SCAFIDI, B., SJOQUIST, D., & STINEBRICKNER, T. (2007). Race, poverty, and teacher mobility. *Economics of Education Review, 26*, 145–159.

TAVARES, P. A. (2015). The impact of school management practices on educational performance: Evidence from public schools in são paulo. *Economics of Education Review, 48*(C), 1–15.

CHETTY, R, FRIEDMAN, J, & ROCKOFF, J (2014). Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates. *American Economic Review, 104*(9), 2593–2632.