Communication within Organizations: Evidence from Teacher Teams

Juan Facal — Columbia University
Gonzalo Ferres — UC Berkeley

Last updated: September 7, 2026

Abstract

We study how reducing coordination costs reshapes communication flows and affects workers productivity within public schools. We exploit quasi-experimental variation in the number of teacher meetings induced by the eligibility for a school program in Uruguay. We find that the exposure to more meetings increases student test scores by 0.21 standard deviations. Using novel data on the use of a digital learning platform, we find that students increase their engagement with the school by 0.18 standard deviations. Permanent teachers, those directly exposed to the meetings, show a positive effect of 0.2 standard deviations in their digital engagement, whereas substitute teachers show no improvements. We find suggestive evidence consistent with enhanced intra-organizational knowledge flows: teachers ex-ante less engaged drive the performance gains, with no adverse effects on their ex-ante highly engaged peers. This implies a Pareto-improving reallocation of information within the team. Furthermore, treatment effects are concentrated in more homogeneous teams, suggesting that reduced coordination costs are most effective when the marginal costs of peer interaction are low. Our results show that reducing fixed coordination costs, even among public organizations with a lack of financial incentives, can increase the overall productivity.

JEL Codes: D23, D83, M54, J24, H11.

1 Introduction

In developing countries, the expansion of schooling access has not been matched by equivalent gains in student learning (WB, 2018). A crucial aspect to improving educational quality is that some teachers lack sufficient knowledge to utilize educational inputs productively (Glewwe and Muralidharan, 2016). Within public school systems, feedback is scarce, and mission alignment is weak, allowing large performance gaps to persist within the same organization (WB, 2023). A broad literature on organizations argues that communication across team members can lower the barriers to adopting effective practices and diffusing knowledge from more- to less-productive workers (Hoffman and Stanton, 2024). Nevertheless, there is virtually no evidence of its role in enhancing performance among this group of scarcely incentivized public workers. Specifically, it is unclear whether organizational failures such as high coordination costs severely limit communication between teachers, and therefore lead to a socially inefficient path for their performance and ultimately, students' learning.

In this paper, we explore the role that communication flow within the school organization has in improving teachers' performance. We provide suggestive evidence of the social learning channels through which an expansion of teacher team meetings can lead to raises in productivity. In addition, we show how coordination costs constitute an organizational friction that stagnates teachers' productivity. We focus on public elementary schools in Uruguay, where we exploit a natural experiment that dramatically expanded the number of paid teacher meetings for some schools. Leveraging quasi-experimental variation in the eligibility criteria for treatment, we estimate how a salient change in the number of teacher meetings potentially altered communication flows within the school social network, and thus, explains the observed increases in teacher performance.

In 2011, the Uruguayan education system introduced the APRENDER program, whose salient initiative increased the number of annual teacher meetings from two to ten for eligible schools. These meetings are held on one Saturday each month during the academic year and last 4 hours. All teachers from the school attend, and the agenda is driven by pedagogical and organizational matters. The APRENDER program is intended for public elementary schools that are both urban and have four-hour instructional time, which are almost half of the country's public schools. The national primary school authority (ANEP) computes a continuous school poverty index every 5 years, which determines eligibility for the program.

Schools falling into the first two quintiles of this distribution are assigned to APRENDER. Since the index is updated every 5 years, schools are reclassified, and some schools can enter or exit the program.

Our identification strategy leverages the sharp discontinuity in treatment eligibility at the threshold of the continuous poverty index. We estimate the local average treatment effect of the expanded meetings using a Regression Discontinuity Design (Cattaneo et al., 2020). Under the identifying assumption that, near the cutoff, the only difference between schools is the as-good-as-random discrete jump in the number of teacher meetings, we are able to leverage the true effects of implementing monthly teacher meetings on teachers' performance. The main strength of this empirical strategy relies on the nature of the running variable. Manipulation into treatment by schools cannot occur, as the index is built on the socioeconomic, predetermined characteristics of their students' households and not on past performance. We estimate both the average causal effect across time and the effect by year to study the dynamics of communication flows and teacher performance. We run alternative specifications to ensure our results are driven by the increase in the number of meetings, and not contaminated by the other two smaller school initiatives in the APRENDER program.

Our empirical analysis draws upon a rich array of administrative data. We utilize standardized math and language test scores for third and sixth-grade students from cross-sectional surveys conducted in 2017, 2020, and 2023. To capture dimensions of teacher effort and student engagement that test scores frequently omit, we incorporate novel, granular data from the CREA digital learning and management platform spanning from 2020 to 2024. This allows us to track the number of days connected to the platform, the number of actions performed, the number of posted comments, and the number of homework tasks assigned/submitted. For teachers' data, we only see observations at the teacher level each year, so we cannot account for group-specific metrics. To avoid non-classical measurement error, we focus our attention on teachers working at a single school each year.

Teachers in treated schools report having more than six additional meetings each year. We document substantial improvements in both student and teacher outcomes resulting from the exogenous expansion of meetings. On the student side, in our most conservative specification, we find significant jumps in academic test scores at the program eligibility cutoff of 0.21 standard deviations. These cognitive gains are accompanied by a striking change in

their engagement. Using an engagement index measuring the number of days connected to the digital platform and the number of actions performed, we find a significant positive effect of 0.18 SD for students in treated schools. We also find increases in the number of homework tasks submitted and the number of comments posted. In line with the fact that permanent teachers are the ones directly affected by the treatment, we observe a pronounced increase in their engagement on the digital platform. We find an increase in their engagement of 0.20 SD. Permanent teachers exhibit an increase of approximately 11 days connected to the platform and 160 additional platform actions compared to their control-school counterparts. Notably, we only find significant effects for permanent teachers. Substitute teachers show no significant jumps, highlighting that effects occur only for those regularly exposed to meetings.

We implement several tests to ensure our results are robust and stable. First, we show that there is no differential sorting based on ex-ante teacher performance or student socioeconomic background, ensuring that the effects stem from changes within the organization's social network rather than from new workers or students. Second, we address external validity concerns by showing that significant effects on teacher engagement persist when using the full universe of teachers, including those working across multiple schools. Third, we run a series of standard robustness exercises for the Regression Discontinuity Design. We run several tests, including density tests of the running variable, placebo analyses, pre-treatment differences analyses, and sensitivity tests. Overall, the results are persistent and robust across all these batteries of tests.

We document three relevant heterogeneities. First, effects are driven entirely by teachers who are ex ante less engaged with the platform within a school. Ex-ante highly engaged teachers show no detectable change. This pattern - improvement at the bottom of the within-team distribution without losses at the top - is consistent with a Pareto-improving reallocation of information in which high performers mentor low performers at low cost to themselves. The effect is also concentrated among the youngest teachers, plausibly those facing the highest pre-existing barriers to initiating productive contact with more experienced peers. Taken together, these effects are consistent with a knowledge-diffusion channel operating within the team rather than with general productivity shocks (Sandvik et al., 2020). Still, there can be alternative stories, such as an expansion of the community of learning among novice teachers, who are more willing to innovate. Second, effects are larger in teams that are more

homogeneous in age and ex-ante engagement with the platform, consistent with the prediction that, even after fixed communication costs fall, marginal interaction costs remain higher in more dispersed teams (Prat, 2002; Lyons, 2017; Hjort, 2014). Third, treatment effects are concentrated in schools that are ex-ante less engaged. Organizations furthest from the production frontier benefit most from loosening the communication constraint, consistent with the interpretation that teachers in such schools had the largest pre-existing wedge between private and social returns to communication.

Related Literature Our study speaks to three different strands of literature. First, we speak to a broad literature of management practices in organizations. There is evidence that good managerial practices matter not only in traditional firms but also in schools (Bloom et al., 2015, 2019; Muñoz and Prem, 2024). In particular, we speak to the literature of knowledge flows in workplaces (Battiston et al., 2021; Hortaçsu et al., 2024). Sandvik et al. (2020) studies the impact of improving knowledge flows between coworkers on firm productivity. They run an RCT in a sales firm where they reduce information friction through different treatments and find positive effects on sales of above 15%. Papay et al. (2020) is the closest work to ours, as they run an experiment to test the role of targeted coaching in improving the performance of existing low-quality teachers. Their intervention led to a 0.12 standard deviation gain in student achievement among those taught by those teachers. Our paper contributes to this literature by showing how team meetings can matter in settings with potentially high coordination costs. In our case, we find that expanding meetings might be positively affecting the degree of knowledge transfer within the social network, as ex-ante low-performing teachers experience sizable performance gains. Knowledge transfer may not come at the expense of high-performing teachers, as they do not experience a significant loss in productivity. What is novel is that in our setting there is no imposition on how workers are paired. Teachers endogenously choose with whom to interact after the exogenous reduction in their coordination costs by expanding the number of meetings.

Second, we add to the peer effect literature (Jackson and Bruegmann, 2009; Hamilton et al., 2003; Cai and Szeidl, 2018). Mas and Moretti (2009) found evidence that a worker's productivity depends on high-quality coworkers' proximity using data from a large supermarket chain. We draw on the literature that seeks to understand the mechanisms by which peers generate spillovers. Proximity enables learning from others and, hence, increases pro-

ductivity (Ellison and Glaeser, 1999; Davis and Dingel, 2019; Chaurey et al., 2025). In our case, we contribute by exploring when peer effects are the largest. First, we add evidence that the quality of peers matters the more workers are exposed to the organizational social network. In particular, by increasing the number of meetings while holding peers' quality constant, we find sizable gains in teacher performance. Second, we show that peer effects are larger among those ex ante more affected by coordination costs, as novice and low-performing teachers are the ones driving the results. Third, by showing that the effects are larger among less dispersed teams in terms of age, we provide additional evidence that peer effects are larger when the marginal costs of interactions are lower.

Third, we contribute to a broader literature on increasing bureaucrat performance in developing countries (Besley and Ghatak, 2005; Besley et al., 2022; Muñoz and Prem, 2024). Public school teachers are highly relevant for the delivery of a salient public service and thus for state capacity. Unlike private firms, which have incentives for improving productivity, officials and teachers in most public-sector settings may have few extrinsic incentives (Muralidharan and Singh, 2023; Lemos et al., 2024). We contribute to this literature by showing how increases in the number of paid teacher meetings can lead teachers to spend more time interacting with peers and, in turn, improve educational outcomes in public schools. To the best of our knowledge, little is known about how to enhance more cooperative contracts in teams of bureaucrats with few extrinsic incentives. We find suggestive evidence that, even in this situation, reducing coordination costs increases productivity.

Outline The paper proceeds as follows. Section 2 describes the institutional setting and the APRENDER program. Section 3 presents the data. Section 4 develops the empirical strategy. Section 5 reports the main results. Section 6 documents the robustness of our estimates. Section 7 analyzes heterogeneity and discusses the mechanism. Section 8 concludes.

2 Background

2.1 Uruguayan public primary education

Schooling can be provided by private centers or by the State. State Education is universal and free, and in 2019 enrolled 82.7% of the total primary school students. Schooling is highly concentrated in the capital, Montevideo, as 32.0% of students attended schools located there

in 2019. Urban schools accounted for 95% of the public supply of primary education. The standard option within urban schools is standard urban schools - the *Urbanas Comunes* schools-, which offer a pedagogical time of 4 hours and represented 70.2% of the students attending urban schools in 2019.[1] The APRENDER program selects from this pool of schools, and, in 2019, accounted for 43.5% of the students attending standard urban schools (Ministerio de Educación y Cultura, 2020). Primary school comprises six grades (1st through 6th) for children aged 6 through 11. Enrollment is first-come, first-served with no geographic or merit restrictions.

Schools are staffed by a principal, supporting administrators, and teachers.[2] Three features of the teacher labor market are relevant for our analysis. First, public-school teachers face no performance-based financial incentives; dismissal is rare in practice, and monitoring by inspectors is infrequent. Second, teachers are divided into permanent, substitute, and intern positions. Permanent teachers are those who hold a teacher certificate and have passed a national contest based on tests and merit that gave them the chance to be tenured at a school center. Within this category, there can be promotions, but those are mainly experience-driven. Substitute teachers are those who work temporarily in a school center to fill a vacancy during the academic year, but they are not part of the school's permanent staff. A teacher can have a temporary position as a substitute in a school, even if they have a permanent certification. Intern teachers are those who can remain at the school for only one year and do not need any certification. Generally, these teachers are students who have not yet obtained their degree. While permanent and substitute teachers earn the same wage (given their seniority), interns receive a lower wage. Teachers can work in multiple groups within and across schools.[3]

All teachers at standard urban schools are required to attend paid, full-staff meetings twice per academic year. The meetings last four hours, are held on Saturdays (leaving regular instructional time unaffected), and follow an agenda set by the principal and teachers rather than by the education ministry. Meetings are recorded in digitized transcripts.

[1] Teachers work twenty hours a week.
[2] There can be special teachers part of specific national programs, as Appendix C describes.
[3] Appendix C provides details on how the centralized teacher market works.

2.2 The APRENDER school program

Introduced in 2011, APRENDER targets the most socioeconomically disadvantaged standard urban schools. The stated objectives of the program are to raise student achievement, consolidate teacher communities, and strengthen school-family ties (CEIP, 2014). Its defining feature is an increase in paid full-staff meetings from two to ten per academic year.[4]

Beyond the meeting expansion, APRENDER bundles two smaller, later-in-the-year components: in-kind transfers and remedial instruction. First, schools in the program can apply for funds for specific projects, disbursed in July as a one-time payment whose modal value is approximately 0.3% of per-student annual expenditure. The second initiative takes place in the final quarter of the year, where specialized teachers are dispatched to support students lagging in their learning. This initiative is intended for specific grades, determined annually by the authorities.

Eligibility. Assignment to APRENDER follows a sharp rule on a continuous poverty index computed by the national primary school authority. The index aggregates predetermined household characteristics of the school's students collected through a representative family survey: household educational level, household socioeconomic level, and social integration level (Appendix C). Schools in the two lowest quintiles of this distribution are eligible; compliance is perfect. Two features of the index make it well-suited for an RD design. First, it is constructed from predetermined socioeconomic inputs outside the school's control, while data collection is centralized by the national authority, precluding manipulation. Second, the cutoff is defined in *relative* terms (the second quintile). The poverty index is reset every five years, so schools can enter or exit the program across revisions. We exploit both editions relevant during our sample period (2017–2021 and 2022–2024).

3 Data

3.1 School Characteristics - Panel 2017-2024

We use administrative data from the primary school national authority (ANEP), which includes information on school characteristics. This database includes data for the whole

[4]The additional eight meetings of four hours correspond to a 5% increase in teachers' monthly earnings.

universe of standard urban schools. Excluded are schools with extended class time, rural schools, and private centers. The resulting sample includes 570 centers, with observations available for each year from 2017 through 2024. Panel A in Table 1 shows summary statistics for these variables. On average, a school has 273 students enrolled in a specific year. Primary school is split into 6 grades, which are typically split into multiple groups. The average number of students in a specific group is 21.

Ordered by the school poverty index, the standard urban schools belonging to the two poorest quintiles are selected as eligible for the program. Compliance is perfect. During our sample period, two editions of the poverty index were in place: one for 2017-2021 and a second for 2022-2024. As expected, from one edition of the poverty index to the next one, most schools maintain the same treatment assignment condition. However, 15% of schools changed their eligibility status during the 2022 recategorization. We rescale both editions of the poverty index so that they are centered at the eligibility cutoff and have a standard deviation of 1. 47% of the observations correspond to a treated school, where assignment depends on the continuous poverty index constructed by the primary school national authority.

3.2 Digital Platform Use - Panel 2020-2024

We obtained data on the use of a digital learning and management platform (CREA) by teachers and students. This dataset is supported by CEIBAL, a digital technology center for education innovation in Uruguay. Having this novel data allows us to capture measures of engagement with the courses for both teachers and students, complementing data on test scores. While test scores measure knowledge at a specific point in time, our digital engagement measures stem from all-year indicators that can shed light on how engaged both the teachers and students are with the school (Jackson, 2018). While student engagement on the digital platform cannot be interpreted as a single predictor of academic performance, we view it as a positive correlate and use it extensively in our analysis, given the rich data we have on it. Similarly, we interpret teacher engagement as a positive correlate of their performance.

Table 1: **Summary Statistics**

	Mean	SD	N
Panel A: Admin 2017–2024			
Montevideo region (dummy)	0.30	0.46	4080
Total enrolled students	272.67	144.97	4080
Students per group	20.58	5.14	4080
APRENDER treatment (dummy)	0.47	0.50	4080
Running variable (standardized)	-0.07	1.00	4075
Unique School IDs	570		
Panel B: Teacher Platform Use 2020–2024			
Number of days connected	62.82	56.92	31609
Number of actions	595.75	896.79	31609
Number of homework tasks assigned	2.43	8.11	31609
Number of comments posted	13.70	41.94	31609
Percentage of weeks connected	0.40	0.26	31609
Permanent teacher (dummy)	0.83	0.38	32146
Unique Teacher IDs	12363		
Unique School IDs	504		
Panel C: Student Platform Use 2020–2024			
Number of days connected	31.71	37.00	697378
Number of actions	97.37	145.02	697378
Number of homework tasks submitted	18.61	40.72	697378
Number of comments posted	3.36	9.08	697378
Unique Student IDs	249838		
Unique School IDs	504		
Panel D: Survey Subsample			
Female (dummy)	0.50	0.50	19159
Ever transferred (dummy)	0.38	0.49	16997
Math score	292.52	49.19	16773
Language score	296.24	49.18	16770
Socioeconomic index	-0.21	0.74	16542
Unique Student IDs	19171		
Unique School IDs	208		

Notes: Each panel describes a dataset and reports mean, standard deviation, and number of observations. Panel-specific footer rows report unique IDs used in each sample.

3.2.1 Teachers

The dataset for teachers covers the period 2020-2024 and includes all teachers who worked in any school during each year. For each year, we restrict our main sample to teachers who worked at a single school (teaching either one or multiple groups), the "1 School Sample", leaving out those who taught in multiple schools. We do this because the dataset has a limitation: platform usage metrics are aggregated at the teacher level each year, without distinguishing metrics for each school or group when a teacher works in more than one. Analyzing teachers who work in multiple schools in the same year without being able to separate the outcomes by school would introduce contamination risk, as schools may have

different treatment status. Table B.1 in Appendix B shows descriptive statistics for teachers in the "1 School Sample" relative to the universal dataset. Our analysis dataset tends to include more observations per teacher, and teachers are slightly older (0.74 years). Overall, no major differences are observed between the two samples. As part of our robustness analysis, to assess external validity, we will test whether our main results differ when using the universal sample. In these cases, for teachers working with multiple groups at the same time, we will impute the values for each group by simple division or weighting by the number of students in the group.

Even though the dataset for teachers is available at the weekly level, we mainly analyze the data aggregated by year, in line with the school calendar. However, we will make use of the weekly observations for robustness analysis. Panel B in Table 1 shows summary statistics on the use of the digital platform for the teacher-year pairs in the "1 School Sample"[5]. Teachers connect to the platform, on average, 63 days per year, and perform 596 actions, including any activity (posting comments, assigning new tasks, answering posts, sending messages to other teachers). On average, they post 14 comments per year and assign 2 homework tasks. They connect to the platform on average 40% of the weeks during the year.

Given that we currently do not observe each teacher's role in a school, we define a proxy for being a permanent teacher in line with the institutional rules described in Section 2. We define the permanent teacher dummy as 1 if the teacher stayed in the same school for at least 2 consecutive years. Given that we will focus our analysis on exposure to teacher meetings, testing heterogeneity by teacher role will be pivotal, as permanent staff are those most highly exposed to the treatment. This is the reason why, in spite of having the data for the years 2019 and 2025, we only use them to observe whether a teacher stayed at least two consecutive years in the same school: we use the 2019 data to see which teachers are "stayers" in a school in 2020, and the 2025 data to see which teachers will stay at the same school in 2025.

In our analysis section, we will define the Engagement Index as our primary outcome for the digital platform, which we define as the average of the standardized number of actions and the number of days connected.

[5]We remove from the sample the teacher-year pairs that present any of the measures on the digital use of the platform above the 99th percentile in that year.

3.2.2 Students

The dataset for students includes all students who were enrolled in any school during each year. For consistency, we use the same period as for the teachers dataset, covering 2020 to 2024. Panel C in Table 1 shows summary statistics on the use of the digital platform for the student-year pairs in the sample.[6] Students connect to the platform, on average, 32 days per year, and perform 97 actions, including any kind of activity (same as for teachers). On average, they post 3 comments per year and submit 19 homework tasks.

Similar to the teachers, we will define the Engagement Index as our primary outcome for the use of the digital platform, which we define as the average of the standardized number of actions and number of days connected.

3.3 Learning & Context Survey - Cross-Section 2017, '20, '23

We obtain data on standardized math and language scores from the *Aristas* study, a cross-sectional survey conducted by INEEd in 2017, 2020, and 2023, using a nationally representative sample of schools. Only students in third and sixth grades were included in the study. In addition to cognitive outcomes, this study includes responses to questionnaires from students, parents, teachers, and principals. Panel D in Table 1 shows summary statistics for this dataset. The sample includes observations for 208 unique schools and more than 19,000 students. On average, 50% of the students are females. This database contains a socioeconomic and cultural context index at the student level, which measures the vulnerability of each student and school in the sample. This index is built using a methodology similar to that of the poverty index measured by the national primary school authority (INEEd, 2018, 2021). Its purpose is only descriptive.

In our analysis section, we will define the Academic Index as our primary outcome for cognitive test performance, defined as the average of standardized scores on the math and language tests.

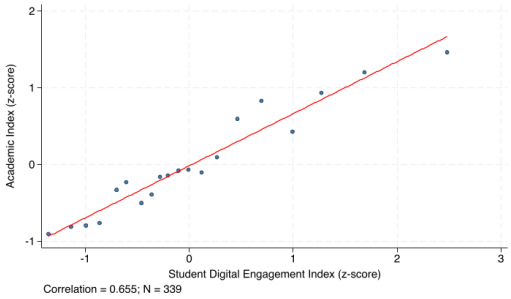
[6] We remove from the sample the student-year pairs that present any of the measures on the digital use of the platform above the 99th percentile in that year. We do the same for teachers as well.

3.4 Correlation between Digital Platform Use and Academic Scores

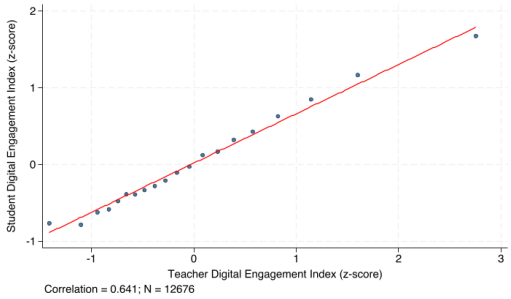
In Figure 1, we include correlations for the Engagement Index of students and teachers. Figure 1a shows the scatter plot between the Student Engagement Index and the Academic Index at the grade-year level.[7] Students who perform better in academic tests tend to be more engaged with the digital platform, with a correlation coefficient of 0.66. Figure 1b shows the scatter plot between the Student Engagement Index and the Teacher Engagement Index at the grade-year level. The correlation coefficient is 0.64, showing how these two variables also tend to move together. Additionally, Figure A.1 in Appendix A shows a correlation of 0.38 between teacher digital engagement and the students' academic index.

Figure 1: Correlation between Digital Platform Use and Academic Scores

(a) Academic Index and Student Engagement



(b) Teacher and Student Engagement



Notes: Panel (a) presents a scatter plot of the Academic Index and the Digital Engagement Index for students. Panel (b) shows a scatter plot of the Digital Engagement Index for teachers and the Digital Engagement Index for students. The correlation coefficient and number of observations are in the footnotes of each graph.

4 Empirical Strategy

To identify the causal impact of the expansion of teacher meetings on school outcomes, we exploit the abrupt change in the probability of being an APRENDER school between the second and third quintiles. As detailed in Section 2, eligibility for the program is determined by a continuous poverty index. We implement a sharp regression discontinuity design (RDD) centered on the eligibility threshold, where schools in the two most vulnerable quintiles (index values above zero in our rescaled metric) are assigned to the treatment.

Our primary approach involves estimating a local linear regression within an optimal bandwidth of the cutoff (Cattaneo et al., 2020). We estimate the following equation:

[7] Due to confidentiality reasons, we are not able to merge the digital platform data and the test scores at the individual level.

$$Y_{i,t} = \beta \text{APRENDER}_{s(i,t),t} + f(x_{s(i,t),t}) + \mathbf{X_1}'_{m(i,t)}\gamma_1 + \mathbf{X_2}'_{s(i,t)}\gamma_2 + \mu_t + \mu_{g(i,t)} + \varepsilon_{i,t} \quad (1)$$

where $Y_{i,t}$ represents the outcome for student or teacher i in year t; $\text{APRENDER}_{s(i,t),t}$ is a dummy variable taking the value of 1 if the school's rescaled poverty score $x_{s(i,t),t}$ is greater than zero; and $f(x_{s(i,t),t})$ is a smooth function of the school poverty index, which we allow to differ on either side of the cutoff. $\mathbf{X_1}_{m(i,t)}$ and $\mathbf{X_2}_{s(i,t)}$ are vectors of group-level and school-level controls, respectively: group size[8], program eligibility status according to the previous edition of the poverty index (past treatment status), assignment to the community teachers program and supporting teachers programs described in Appendix C, and a dummy of whether the school is in Montevideo (the capital of Uruguay). μ_t and $\mu_{g(i,t)}$ capture year and grade fixed effects. We will include these in those specifications where the outcome is not standardized by year and grade, respectively. $\varepsilon_{i,t}$ is the error term, with standard errors clustered at the school (treatment) level to account for within-school correlation across grades and years.

Our primary specification includes all available years for the outcome of interest. We also run estimates by year to study dynamic effects. Our benchmark specifications employ a triangular kernel and an MSE-optimal bandwidth selection criterion. The tables report the conventional RD point estimates (β). For inference, we use robust-bias-corrected standard errors and their corresponding p-values to ensure the statistical validity of our results. We focus on the linear local polynomial fit and also report results from the quadratic one in the plots.

Under the identification assumption that $\varepsilon_{i,t}$ does not change abruptly at the program eligibility cutoff, β provides an unbiased estimate of the average causal effect of the increase in the number of teacher meetings for teacher and student outcomes across time. This identification assumption relies on three complementary assumptions. First, we assume that schools are not sorted around the threshold. We test this assumption via density tests and by showing that school-level covariates are balanced and smooth at the cutoff. Second, we assume neither teachers nor students are selecting into treated schools based

[8]In our main analysis, where we use the "1 School" teacher sample, we control for the number of groups and students she teaches in a given year.

on ex-ante quality or characteristics that correlate with $Y_{i,t}$. We provide evidence that ex-ante teacher engagement and student socioeconomic background remain continuous at the eligibility threshold. Third, we assume that the other two program initiatives (remedial education and in-kind transfers) do not drive the treatment effects, and that these effects are explained solely by the increase in the number of teacher meetings.

We test this last assumption by estimating our main specifications, restricting the sample to observations unaffected by them. Observing estimates that are virtually similar provides support for our assumption. First, for the student outcomes, we run our specifications excluding observations from grades directly targeted by the remedial education initiative, isolating as much as possible the impact of the increased teacher meetings. Second, for the teacher outcomes, as we have weekly measures of the use of the digital platform available, we estimate effects on their engagement, excluding observations from the second semester (as both extra initiatives take place then).

As we explain in Section 3, in our main analyses we focus on the "1 School" sample, which for each year includes only teachers who taught at a single school. In these cases, the teacher may have taught one or multiple groups in that school. Our grade fixed effects μ_g incorporate this feature of the dataset, capturing the different grade combinations the teacher is working with[9]. We test whether our results are robust to using the universe of teachers using the imputed outcome values described in Section 3.

We also run further robustness analyses for our specifications. We run pre-treatment tests to predict outcomes for the period 2017-2021 using the poverty index for the following period (2022-2024), to show that there is no anticipation of effects or pre-trends. In addition, we provide balance tests of school characteristics to show their smoothness at the cutoff. Further, for our main estimates, we study effects using alternative specifications (local vs global polynomial, polynomial order, kernel, bandwidth selection, additional controls) to ensure results are not sensitive to our specification choice.

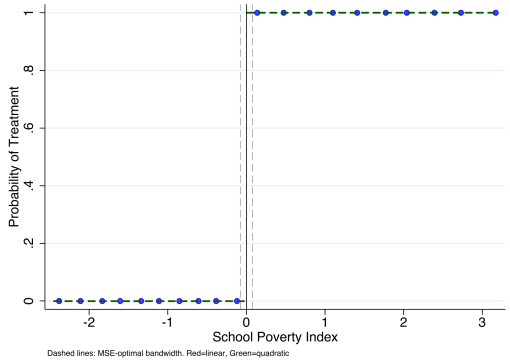
[9]In our robustness specifications using the universe of teachers, where we impute the outcome values for each group, we use standard grade fixed effects.

5 Results

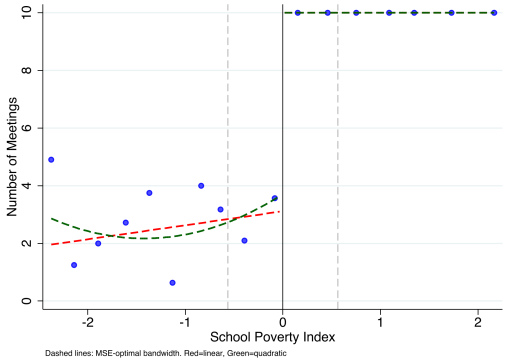
First Stage Figure 2a plots the compliance into treatment based on the school poverty index assignment criteria described in Section 2. There is perfect compliance in the selection into the APRENDER program. Figure 2b plots the number of meetings reported by the teachers in the 2023 *Aristas* survey. Interestingly, while there is no variation among treated schools in the fact that all have monthly meetings - ten during the academic year - teachers vary in their response among standard urban schools. Theoretically, standard urban schools should have two meetings per academic year. Nevertheless, at the cutoff, the average frequency is almost 4 meetings per year. Still, there is a clear jump of more than 6 additional meetings per year due to treatment.

Figure 2: First Stage Compliance

(a) Assignment to APRENDER (2017-2024)



(b) Reported Number of Meetings



Notes: This figure presents the first stage results. Panel (a) shows the assignment into the APRENDER program for the period 2017-2024. This period covers two editions of the poverty index: one for 2017-2021 and a second for 2022-2024. Panel (b) utilizes survey data from the *Aristas* survey to show the reported number of teacher meetings. All estimates are centered around the program eligibility cutoff.

Academic Outcomes Figure 3 shows the program's effects on the student's Academic Index. Students experience a significant increase of 0.32 standard deviations in their test performance. Treated schools on the cutoff score on average higher than even what a one-standard-deviation less vulnerable school would score. This implies that treatment makes students in APRENDER schools that are in the second quintile of the distribution perform as well as those in the fourth quintile. Table B.3 presents the decomposition of such an effect. The effect is larger and more precisely estimated for the language tests than for the math scores. Students experience a 0.34 SD increase in language scores and a 0.20 SD increase in math scores.

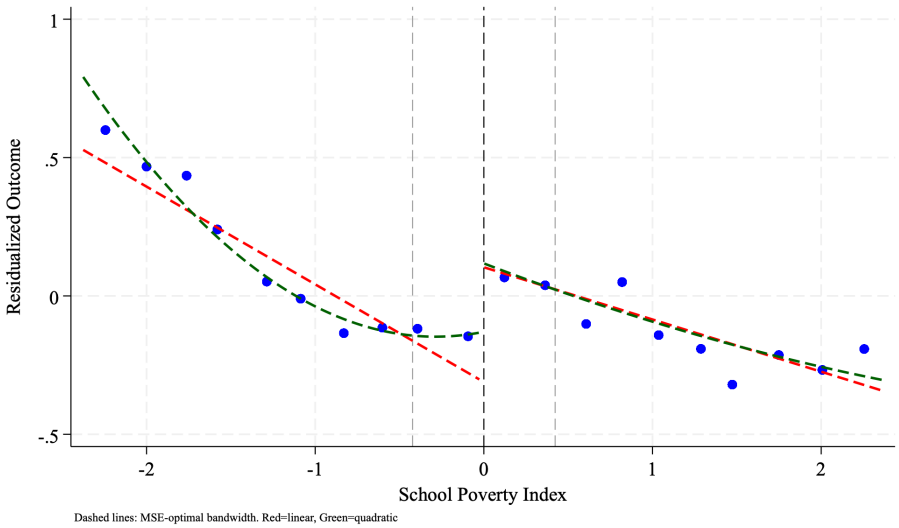


Figure 3: Effect on Academic Index

Digital platform use - Students We also analyze effects on the students' engagement with the digital platform, using the outcomes described in Section 3. As long as there is no differential crowd-out from traditional learning, increases in the Engagement Index will reflect that students are making greater overall effort. Figure 4 presents the effects of expanding teacher meetings on student engagement with the digital platform. We find a significant increase in students' engagement of 0.16 SD. As the plot shows, schools that are almost 1 standard deviation less vulnerable have average engagement values similar to those of APRENDER schools in the cutoff. Students are 0.15 SD more days on the platform and perform 0.17 SD more actions annually than their counterfactuals in similar control schools, as Table 2 shows. Interestingly, we also observe a 0.1 SD increase in the number of comments they post on the platform, along with a corresponding 0.22 SD increase in the number of homework assignments they submit. As discussed in Section 3, these positive effects complement the increases found in the academic test scores, as these measures are closely related to students' effort and motivation in their learning process.

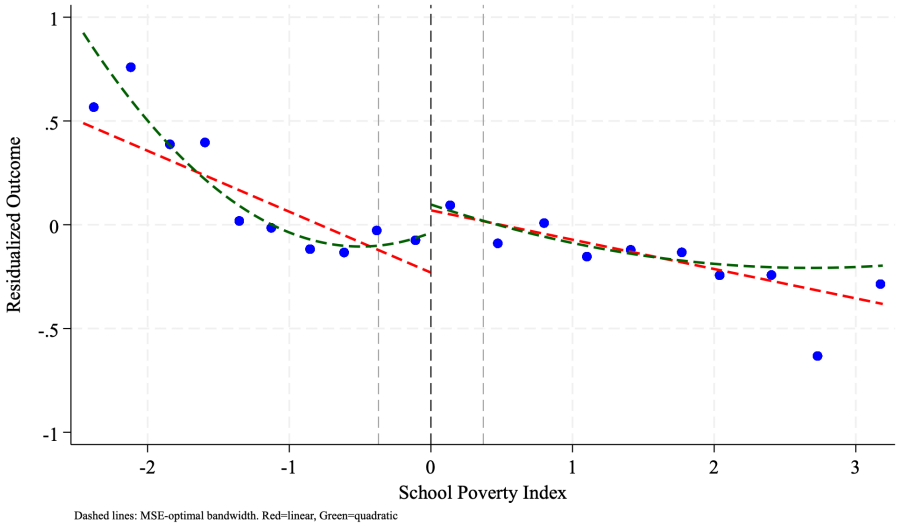


Figure 4: Effect on Students' Engagement with the digital platform

Table 2: **Effect of APRENDER on CREA usage outcomes at Student Level - All Teachers**

	CREA usage - Students				
	Engagement Index	Days Connected	Actions	HWS Submitted	Comments Posted
RD Estimate	0.163**	0.145**	0.173***	0.220***	0.111***
	(0.061)	(0.061)	(0.059)	(0.059)	(0.038)
	[0.072]	[0.071]	[0.070]	[0.070]	[0.044]
Controls:					
Year FE	No	No	No	No	No
Region FE	Yes	Yes	Yes	Yes	Yes
Grade FE	No	No	No	No	No
Bandwidth	.452	.459	.449	.423	.444
Mean Dep. Var.	-0.004	0.000	-0.010	-0.005	-0.006
SD Dep. Var.	0.940	0.927	0.954	1.015	0.967
Observations	697,378	697,378	697,378	697,378	697,378
Eff. Obs. at the Left ; Right	98427 ; 117936	99355 ; 121881	96541 ; 116852	88365 ; 112486	93296 ; 116375
Eff. Schools at the Left ; Right	131 ; 129	132 ; 132	126 ; 128	119 ; 123	126 ; 127

Notes: The table reports the conventional RD point estimate. Conventional standard errors are in parentheses; robust-bias-corrected standard errors are in brackets. Inference is based on the robust-bias-corrected standard error following Cattaneo, Idrobo, and Titiunik (2020). The school poverty index measures the vulnerability of the area where the school is located according to socio-economic indicators; see text for details. It is rescaled so that a school is eligible to be APRENDER if its index is higher than zero. All specifications implement a sharp RD design due to perfect compliance. Clusters at the school level are taken in every specification. All specifications use a triangular kernel, MSE-optimal bandwidth selection, and a local polynomial of order 1. The mean and standard deviation of the dependent variable are measured inside the estimation bandwidth among control schools. Effective observations are those within the local bandwidth. ***, **, and * indicate significance at the 1%, 5%, and 10% levels based on robust-bias-corrected inference.

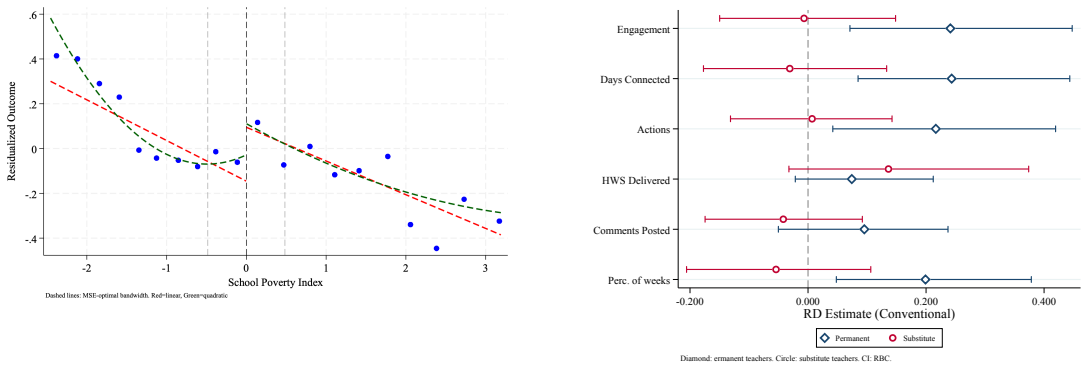
Digital platform use: Teachers Since teachers are directly affected by the expansion of meetings, we explore whether there are any changes in their engagement as a channel to rationalize the increases in students' academic achievement and engagement. Figure 5a shows that teachers are 0.19 SD more engaged with the digital platform. On the cutoff, engagement levels for teachers in treated schools are similar to the average teacher in schools more than 1 standard deviation below the cutoff in terms of vulnerability. Table B.4 indicates that teachers are 0.21 SD more days connected and perform 0.17 SD more actions on the platform. The effects on the number of comments posted and the number of homework tasks assigned are positive but not significant. The fact that teachers are not posting more

homework tasks, while students are submitting more, as Table 2 shows, may indicate that the increase in students' homework submission is not mechanical but rather reflects greater student engagement.

Figure 5: Effect on Teachers' Engagement with the digital platform

(a) All Teachers

(b) Permanent vs Substitute Teachers



Notes: This figure presents the main effects of the expansion of meetings on teacher performance. Panel (a) shows the plot from estimating Equation 1 using the teacher engagement index as the dependent variable. We analyze the effect on all "1 School" teachers. The engagement index is the average of the standardized number of days connected and the number of teacher actions on the digital platform. Panel (b) plots the coefficient plots from running the same equation but splitting between permanent and substitute teachers. The period goes from 2020 to 2024.

We explore the heterogeneity of these effects by teacher type. Permanent teachers are exposed to an increased number of meetings, while substitutes are substantially less exposed or not exposed at all due to the temporary nature of their job. [10] As Figure 5b portrays, permanent teachers are those who drive the effect with a 0.22 SD increase in engagement, which aligns with the fact that they are those in the teaching staff who are regularly affected by the monthly meetings. On the other hand, there are no significant differences between substitute teachers in treated and control schools. This test serves as a placebo analysis, since substitute teachers should not be affected by APRENDER, given their low or non-existent exposure to meetings. For the remainder of the paper, we will focus on analyzing teacher engagement among permanent teachers.

Dynamic effects We exploit the yearly dimension of our data to examine whether effects exhibit cumulative behavior. Answering this question has significant policy implications: if the program had an immediate, non-cumulative impact, it would be sufficient to treat schools once rather than repeatedly. Surprisingly, the academic and engagement indices

[10] Substitute teachers can attend meetings, but since their job is temporary, the likelihood of being affected is low.

show different behaviors. On the one hand, Figure A.3 shows that effects on test scores are significant in the recategorization year and the following year, but not in the third year after the recategorization (we don't observe years 2 or 4). On the other hand, Figure A.4 shows that the student engagement effects appear only after the third year relative to the recategorization. Similarly, Figure A.5 shows that teacher engagement effects start being significant after the second year relative to the recategorization. Both figures suggest a cumulative story for the engagement with the school from teachers and students. This finding aligns with the idea that organizational transformations may unfold over time. It can take up to three years to see significant changes in teachers and four years in students.

Moving away from the Bundle As Section 4 refers, our main specification exploits the effect of being assigned to the APRENDER program. Even though the most salient initiative of the program is the expansion in the number of meetings, the other two, and in particular the remedial education one, can be conflated with and affect the true effects of the increased number of meetings. We run additional specifications to show that the bundle results presented in the previous paragraphs capture the desired causal effect of the increase in the number of meetings. In particular, we run specifications only for observations unaffected or least affected by the remaining two initiatives.

As we explain in Section 4, for student test scores, given the nature of the survey data and the timing of the tests at the end of the year, we can only control for the remedial education initiative. We perform a conservative test by removing the observations from the affected grades for each year. Despite losing precision, we still observe that student achievement is higher in APRENDER schools by 0.21 SD, as shown in Figure A.6. We perform the same exercise for student engagement on the digital platform. Figure A.7 outlines that there is virtually no difference between the bundle and the no-remedial-education coefficients: we observe a significant 0.18 SD increase in student engagement.

Our most exhaustive exercise can be done using teacher usage data, since we have weekly-level observations. Hence, we can not only turn off the effect of remedial education by removing the last quarter of the year, but also by removing the effect of the in-kind transfers, which would be after the transfers happen in the second semester of the year. Figure A.8 plots three different specifications: (i) the bundle effect, (ii) removing the second semester (turning off the in-kind transfers initiative and remedial education from the current year), and (iii)

same as (ii) but also removing grades that could have been affected by remedial education the previous year. The size and precision of the effects persist for these specifications, indicating that the bundle effect already captures the effect of the increased number of meetings. We observe a significant 0.20 SD increase in specification (iii), which is the most conservative.

Overall, these tests rule out the concern that the effects of the bundle were driven by the other initiatives. The lack of change in the coefficients when isolating the observations least affected by them supports our main identifying assumption that the bundle effect is capturing the effect of expanding teacher meetings. We reckon that, as a limitation of the paper, we cannot impose more strict restrictions on student academic test scores or student digital platform usage, as we do with teacher performance, due to the nature of the data. However, conditional on the implementation of feasible restrictions, the coefficient remains virtually stable, providing additional confidence in those results. Also, our rich teacher dataset allows us to be more precise in studying the first-order effects of the meetings: teacher effort in school, which is our main focus in the paper. We show that the coefficients associated with teacher performance are robust and stable to the most conservative specifications that only consider the observations none or least affected by the other two initiatives.

6 Robustness

Manipulation of the Running Variable We test whether schools are differentially sorting into treatment or control around the cutoff. In theory, it is unlikely that schools manipulate the running variable given its nature. The school poverty index is centrally measured by the primary school national authority and is derived from aggregating several household characteristics from representative samples of school families. In addition, instead of a fixed value, the eligibility cutoff is the second quintile of the index. As long as the index is continuous and based on schools' relative vulnerability, there is little scope for differential density around the cutoff. Panel (a) in Figure A.9 confirms these expectations by running a density test for the main sample. We also run the test for the *Aristas* survey subsample, as here there could be differential selection to the survey on either side of the cutoff. There is no manipulation of schools around the cutoff in either the full sample or the subsample.

Differential Sorting Based on Quality Another big threat to our results is that, since the treatment is at the school level, the arrival of new teachers who were ex-ante differentially better might be explaining the results. In other words, the effects we observe could be due not to the expansion of communication flows and higher diffusion in the social network, but to a differential attraction of ex-ante better teachers, driven by the 5% earnings premium or other factors. In Table B.5, we test whether new teachers in the treated schools are, ex ante, more engaged with the digital platform around the cutoff. There are no significant differences in the ex-ante quality of teachers entering APRENDER and standard urban schools near the cutoff in any of the outcomes.

In a similar vein, students who select into APRENDER might come from a better background. We use our rich survey data on students' SES background and see no discontinuity around the cutoff in Table B.6. The students' vulnerability index, the likelihood of coming from another school, or being female, are smooth around the cutoff. These verify that there is no differential sorting of students around the cutoff. Overall, this set of findings rules out the possibility that differential teacher and student selection based on ex-ante performance or SES background is contaminating our results. Our findings are robust to this, shedding light on the fact that the positive treatment effects do not come from outside but from changes within the organization.

Alternative Teacher Samples Our main results may also be subject to a potential limitation. As Section 3 outlines, given the data-aggregation level, we only consider teachers who taught at a single school each year to obtain true measures at the school-year level. However, these teachers represent only 54% of the school team on average, as shown in Table B.2. Even though the descriptive statistics for the "1 School" teachers and the universe look similar, there may be unobservable characteristics that differ between the two groups. If this were the case, our main results on teacher engagement could present external validity concerns. To address this, we estimate our main regression using all permanent teachers' imputed values instead. As we explain in Section 3, the variables that can be imputed are the number of actions, the number of homework tasks assigned, and the number of comments. The imputations can be made by weighting by the number of students per group, or by uniformly splitting across the number of groups taught by the teacher. As Figure A.10 plots, the effect on the number of actions performed is also significant for all

permanent teachers in the school, although it becomes smaller in size. While the number of actions rises by approximately 0.20 SD for the "1 School" teachers, when considering the entire sample of permanent teachers in the school, the effect drops to 0.09 SD. The effect remains relatively stable when we run the specifications that turn off the in-kind resources and remedial education initiatives. The effects are virtually the same regardless of the imputation technique. Even though these results come from imputed measures, and those from the "1 School" teachers are more trustworthy, these findings can still be informative. We still see significant effects on the number of teacher actions, which suggests that our main results are not substantially biased relative to the universe of teachers in the school.

Falsification Tests First, to ensure that our main findings are not driven by spurious jumps or misspecification of the underlying functional form, we conduct a series of placebo cutoff tests. We estimate Equation 1 at various false thresholds where no change in program eligibility occurs. Table B.7, Table B.8, and Table B.9 show the same pattern overall for the test scores, student, and teacher engagement: an absence of significant discontinuities at placebo cutoffs. At $c = 0.5$ there is a significantly negative effect on the teacher engagement, as Table B.9 shows. Still, the effect has the opposite direction, alleviating concerns that the true cutoff is the result of a spurious jump. Overall, these tests reinforce the validity of our RD design and suggest that the observed effect at the true APRENDER cutoff is due to the program's expansion rather than noise.

Second, our main findings could be driven by ex-ante differentials rather than the treatment effect. We present two tests that rule out that possibility. The first piece of evidence is on our dynamic regression discontinuity plots. Figure A.3, Figure A.4, and Figure A.5, show no pretrends in our main outcomes. Moreover, Table B.10, Table B.11, and Table B.12 test whether there is any discontinuity of pre-treatment outcomes on the cutoff for the test scores sample, the students platform sample, and the teachers platform sample, respectively. We find that the pre-treatment outcomes are smooth around the cutoff, ruling out that schools' past trajectories could explain our treatment effects.

Third, the main effects might be correlated with other programs or school characteristics that can be discontinuous at the cutoff. We test this by running Equation 1 without controls for school characteristics and treating them as dependent variables. **??** outlines that there is no difference at the cutoff in the likelihood of having school lunch or breakfast (for offering

lunch, the effect is negatively significant at 10%). Also, the chances of having a multidisciplinary team at the school, specialized classes or teachers, or English teachers are smooth at the cutoff. Classroom sizes and the number of immigrants at schools are also continuous at the cutoff. The same pattern arises in Table B.14, which assesses whether there is any difference in the type of teachers schools have. There is no significant jump at the cutoff for almost all types of special teachers. The exception is that APRENDER schools are 20 percentage points more likely to have community teachers. However, such a variable is already included in Equation 1, which reinforces our model specification. Overall, no other programs, school characteristics, or teacher types show a discontinuous jump at the cutoff, reinforcing the validity of our effects.

Alternative Specifications We run different versions of Equation 1 to ensure our results are robust and stable to different alternatives. We do the following modifications to our main specification: (i) uses a quadratic polynomial function for the running variable, (ii) uses a CER-optimal bandwidth criterion, (iii) uses a uniform kernel, and (iv) uses a global linear polynomial estimation method instead of the local one. As Table B.15 points out, the benchmark estimation for student academic scores is robust to these alternatives. The effect is stable and significant across almost all specifications. Still, there is a drop when using the global linear estimation in the last column, as the effects go from 0.32 SD to 0.19 SD.

Table B.16 shows that student engagement is also robust and stable across almost all specifications. However, it loses significance when using the CER-optimal bandwidth. We are not concerned about this, as CER-optimal bandwidths penalize the bias from moving away from the cutoff when estimating, which improves inference, but at the cost of a less reliable point estimate than the MSE-optimal bandwidth (Cattaneo et al., 2020). The loss of significance stems from the fact that the point estimate is lower than the benchmark, while the standard errors are not much different from those of the quadratic polynomial in the second column. Hence, we do not interpret the findings in the third column as troublesome; we should only care about the inference, not the point estimate, when using the CES-optimal bandwidth. Apart from that, even though the global polynomial estimator is significant, the effect is reduced by almost a third relative to the benchmark.

Table B.17 describes how robust the effect on teacher engagement is to these alternative specifications. The effects are overall robust. When using the global polynomial in the last

column the effect decreases considerably, from 0.23 sd to 0.14 sd. Still, the effect of more meetings on teacher engagement is sizable and relatively stable.

Furthermore, our results might be confounded by school characteristics. We modify Equation 1, adding in each column of Table B.18, Table B.19, and Table B.20 for the test scores, student, and teacher engagement samples, respectively, to ensure that our main results are robust to saturating our main specification with more controls. In Table B.18, we control for whether the school offers lunch and breakfast, the elementary school size, and the student vulnerability index reported in the survey. The effects are robust and persistent across cumulative additions of these controls. For student and teacher engagement, we add controls for whether the school offers lunch and breakfast, school size, number of permanent teachers in the group, and total number of permanent and substitute teachers in the school. As Table B.19 and Table B.20 show, the effects on student and teacher engagement, respectively, are robust and stable when such controls are added. A limitation of these two last models is that we are unable to control for students' or teachers' SES background using administrative data. Nevertheless, knowing that the effects are still significant for test scores, using the rich survey data and controlling for their vulnerability level, can be a partial way to ensure it is a minor concern.

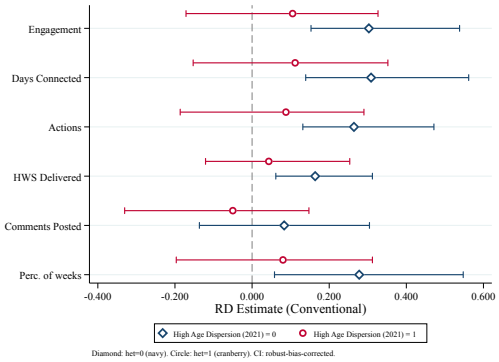
7 Heterogeneous Analysis

Identifying which schools and which teachers drive the average treatment effect is informative for two reasons. First, it tells us which team configurations are most responsive to the communication margin. Second, and more important for this paper, the pattern of within-team heterogeneity can help distinguish between competing mechanisms - in particular, between a knowledge-diffusion story (in which low performers improve by learning from high performers) and general productivity effects (in which all teachers improve roughly equally). Because assignment to treatment depends on a poverty index that is plausibly uncorrelated with team composition or ex-ante performance near the cutoff, we can interact the RD with these margins to analyze heterogeneities at the school and teacher levels. We focus throughout this section on teacher engagement, the first-order outcome most directly tied to the increase in the number of meetings.

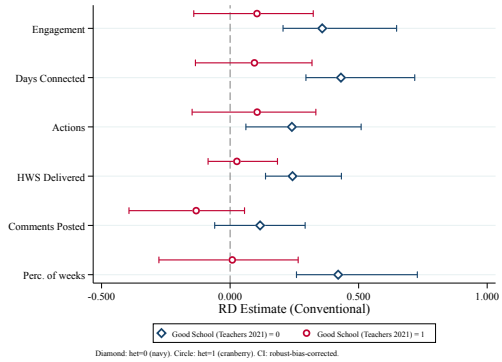
Team composition: dispersion Figure 6 reports our main heterogeneity results. Figure 6a splits schools by the coefficient of variation of teacher age in the 2021 teaching staff (conditional on each school's 2022 poverty index). Treatment effects are concentrated in less age-dispersed teams, and high-age-dispersion teams show no detectable response. Figure A.11a shows the same pattern when we split by the dispersion in the ex-ante teacher number of actions: more homogeneous teams drive the effect. The pattern is consistent with models in which the marginal cost of productive peer interaction rises with team heterogeneity (Prat, 2002; Lyons, 2017; Hjort, 2014). Lowering fixed communication costs (by increasing the number of meetings) is therefore most effective when the remaining marginal costs of within-team interaction are lowest.

Figure 6: Heterogeneity by Team Characteristics and Intra-School Differences

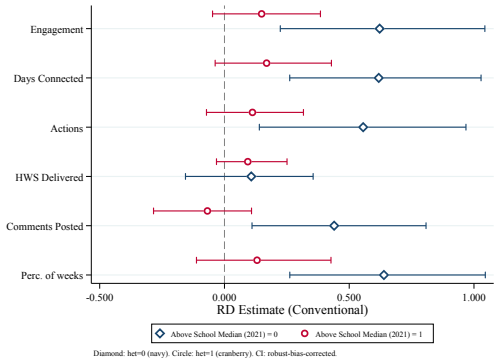
(a) School Age Dispersion



(b) School Ex-Ante Number of Actions



(c) Intra-School Ex-Ante Number of Actions



Notes: This figure presents the main heterogeneous analysis. The first row presents heterogeneity exploiting school characteristics. Panel (a) shows the heterogeneous effects on schools, split between those with highly dispersed teacher ages (red circle) and those without (blue diamond). High age-dispersed schools are those whose coefficient of variation of age for the entire teacher team in 2021 - before the recategorization - is above the average for their vulnerability index in 2022. Panel (b) plots the effects differentiating by the ex-ante quality of schools in 2021, previous to the reclassification in 2022. We obtain the average number of actions done by permanent teachers in every school in 2021 and determine which are above or below their average, given their vulnerability index in 2022. Then we plot the ex-ante good schools, given their vulnerability index (red circle), and bad schools (blue diamond), based on past teacher performance. In the second row, we present heterogeneity within schools. In panel (c) we compare the effects between teachers that in 2021 were above the median in terms of number of actions from the whole teacher team (red circle), and those below the median (blue diamond).

School ex-ante performance Figure 6b splits schools by their 2021 average teacher number of actions, conditional on their 2022 vulnerability index. Treatment effects are concentrated in ex-ante low-performing schools; ex-ante high-performing schools, conditional on vulnerability, show no statistically detectable gain.[11] Organizations furthest from the production frontier benefit most from loosening the communication constraint. Our rationale is that ex-ante low-performing schools have the largest gap between private and social returns to intra-team communication.

Within-team heterogeneity The sharpest test of the knowledge-diffusion hypothesis is at the within-school level. Figure 6c splits permanent teachers by whether their 2021 platform activity placed them above or below the median of their own school's teaching staff. Teachers below the within-school median drive the effects; those above the median show no statistically detectable change. Within a given treated school, therefore, the meeting expansion does not raise all teachers' productivity equally, but it raises productivity only at the bottom of the within-team distribution, leaving the top unchanged. This pattern is exactly what a knowledge-diffusion mechanism would generate: by lowering coordination costs, the intervention enables ex-ante low performers to establish productive contact with higher-performing peers, from whom they learn effective practices. High performers show no detectable gains in their number of actions. Still, the absence of a decline at the top also implies that the transfer of knowledge does not detectably impose a productivity cost on them. Taken together, treated schools move to a Pareto superior equilibrium on this margin. We observe a parallel pattern when we split teachers by their position in the within-school age distribution, as shown in Figure A.11c. The youngest tertile of teachers within each school drives the treatment effect; older tertiles show smaller, less precise responses. Young teachers are disproportionately novice and, consistent with the WB (2023) evidence that novices face the highest communication frictions, are the most responsive to treatment.

An important caveat to our findings is that the observed heterogeneity may reflect localized productivity shocks rather than top-down knowledge diffusion. By increasing meeting frequency, the organization fosters a collaborative environment that encourages pedagogical innovation. These improvements, therefore, may not necessarily originate from a transfer of

[11] Figure A.11b replicates this split using ex-ante student engagement in 2021. The split is less sharp, though the point estimates are larger on the number of days and percentage of weeks connected for ex-ante low-performing schools.

best practices from high-performing teachers, but rather from the provision of dedicated time for collective experimentation. Novice teachers, who inherently face the lowest marginal costs of adaptation, naturally exhibit the largest treatment effects when coordination frictions are relaxed.

8 Conclusion and Next Steps

In this paper, we provide novel evidence of how changes affecting a latent organizational friction can improve performance among public workers. In particular, how reducing coordination costs, by increasing the number of teacher meetings, can foster teacher performance. Exploiting a quasi-exogenous, sizable increase in the number of teacher meetings in elementary schools, we find noticeable positive effects on students' achievement and effort. More importantly, we find increases in teacher performance and engagement, which align with the fact that they are those directly affected by the treatment.

Our analysis on heterogeneous effects finds suggestive evidence of an incremental presence of knowledge flows in the treated schools, as the effect is driven by the ex-ante low-performing teachers. The expansion of meetings seems not to be causing general improvements, but the diffusion of knowledge from high-performing to low-performing teachers might. Yet we cannot rule out that the effects come from localized productivity gains, rather than top-to-bottom knowledge diffusion. Ongoing work seeks to tackle that. Moreover, the effects seem to be larger on more homogeneous teams in terms of age and ex-ante quality.

We highlight the importance of understanding how changes in organizational structures can affect workers' performance. In particular, it is of utmost significance to assess how alleviating organizational frictions, such as coordination costs, can increase communication levels within teams and thereby foster performance. We do not claim that additional meetings are a sufficient condition for increased performance. Nonetheless, in organizations with poor communication and high coordination costs, the social marginal return of increasing team meetings from semi-annual to monthly is relevant, as it can enable greater interactions between peers and, therefore, foster overall productivity.

In future work, we seek to shed more light on the mechanism at play and whether teachers are enhancing their communication in spaces outside staff meetings. We plan to conduct interviews with teachers to understand their experiences and beliefs about the relevance

of meetings and other communication spaces. In addition, we plan to conduct three analyses with new data we have accessed this summer and yet to get data. First, we have signed a form to get data of the interactions between teachers on the digital platform, allowing us to have a proxy of the social network within the organization. By exploiting these data, we will evaluate whether there is a change in teachers' social networks, by measuring whether ex-ante low-performing teachers interact with better peers, and whether the central nodes in the network are more performance-oriented. Second, this summer we have collected more than 25000 digitalized meetings transcripts. By analyzing them, which is our current stage, we plan to examine the connections between the topics discussed and performance improvements. In particular, we are interested in understanding whether schools that discuss pedagogical strategies more are driving the effects. Third, we recently have collected data on principals' tenure test scores, and thus, we are looking forward to evaluating the heterogeneous treatment effects accounting for the quality of team managers.

References

ANEP (2016). Relevamiento de características socioculturales de las escuelas públicas del consejo de educación inicial y primaria 2015. Reporte técnico, Dirección Sectorial de Planificación Educativa, División de Investigación, Evaluación y Estadística, Montevideo, Uruguay.

Battiston, D., Blanes I Vidal, J., and Kirchmaier, T. (2021). Face-to-Face Communication in Organizations. *The Review of Economic Studies*, 88(2):574–609.

Besley, T., Burgess, R., Khan, A., and Xu, G. (2022). Bureaucracy and Development.

Besley, T. and Ghatak, M. (2005). Competition and Incentives with Motivated Agents. *American Economic Review*, 95(3):616–636.

Bloom, N., Brynjolfsson, E., Foster, L., Jarmin, R., Patnaik, M., Saporta-Eksten, I., and Van Reenen, J. (2019). What Drives Differences in Management Practices? *American Economic Review*, 109(5):1648–1683.

Bloom, N., Lemos, R., Sadun, R., and Van Reenen, J. (2015). Does Management Matter in schools? *The Economic Journal*, 125(584):647–674.

Cai, J. and Szeidl, A. (2018). Interfirm Relationships and Business Performance*. *The Quarterly Journal of Economics*, 133(3):1229–1282.

Cattaneo, M. D., Idrobo, N., and Titiunik, R. (2020). *A Practical Introduction to Regression Discontinuity Designs: Foundations.* Elements in Quantitative and Computational Methods for the Social Sciences. Cambridge University Press.

CEIP (2014). *Haciendo memorias: Programa A.Pr.En.D.E.R.* Consejo de Educación Inicial y Primaria (CEIP), ANEP, Montevideo, Uruguay. Compiladora: Luisa Ayerza. Coordinadora: Nancy Picotti.

Chaurey, R., Nayyar, G., Sharma, S., and Verhoogen, E. (2025). Social learning among urban manufacturing firms: Energy-efficient motors in bangladesh. Working paper.

Davis, D. R. and Dingel, J. I. (2019). A Spatial Knowledge Economy. *American Economic Review*, 109(1):153–170.

Ellison, G. and Glaeser, E. L. (1999). The Geographic Concentration of Industry: Does Natural Advantage Explain Agglomeration? *American Economic Review*, 89(2):311–316.

Glewwe, P. and Muralidharan, K. (2016). Improving Education Outcomes in Developing Countries. In *Handbook of the Economics of Education*, volume 5, pages 653–743. Elsevier.

Hamilton, B. H., Nickerson, J. A., and Owan, H. (2003). Team Incentives and Worker Heterogeneity: An Empirical Analysis of the Impact of Teams on Productivity and Participation. *Journal of Political Economy*.

Hjort, J. (2014). Ethnic divisions and production in firms. *The Quarterly Journal of Economics*, 129(4):1899–1946.

Hoffman, M. and Stanton, C. T. (2024). People, Practices, and Productivity: A Review of New Advances in Personnel Economics. *NBER Working Paper*.

Hortaçsu, A., Natan, O. R., Parsley, H., Schwieg, T., and Williams, K. R. (2024). Organizational Structure and Pricing: Evidence from a Large U.S. Airline. *The Quarterly Journal of Economics*, 139(2):1149–1199.

INEEd (2018). *Anexos. Aristas 2017: Informe de resultados de tercero y sexto de educación primaria*. Instituto Nacional de Evaluación Educativa (INEEd), Montevideo, Uruguay. Accessed: 2026-04-17.

INEEd (2021). Aristas 2020: Primer informe de resultados de tercero y sexto de educación primaria. Technical report, Instituto Nacional de Evaluación Educativa (INEEd), Montevideo, Uruguay. Accessed: 2026-04-17.

Jackson, C. K. (2018). What Do Test Scores Miss? The Importance of Teacher Effects on Non–Test Score Outcomes. *Journal of Political Economy*, 136(5).

Jackson, C. K. and Bruegmann, E. (2009). Teaching Students and Teaching Each Other: The Importance of Peer Learning for Teachers. *American Economic Journal: Applied Economics*, 1(4):85–108.

Lemos, R., Muralidharan, K., and Scur, D. (2024). Personnel Management and School Productivity: Evidence from India. *The Economic Journal*, 134(661):2071–2100.

Lyons, E. (2017). Team production in international labor markets: Experimental evidence from the field. *American Economic Journal: Applied Economics*, 9(3):70–104.

Mas, A. and Moretti, E. (2009). Peers at Work. *American Economic Review*, 99(1):112–145.

Ministerio de Educación y Cultura (2020). Anuario estadístico de educación 2019. Technical report, Dirección de Educación, MEC, Montevideo, Uruguay. Accessed: 2026-04-17.

Muralidharan, K. and Singh, A. (2023). Improving Public Sector Management at Scale? Experimental Evidence on School Governance in India. *Working Paper*.

Muñoz, P. and Prem, M. (2024). Managers' productivity and recruitment in the public sector. *American Economic Journal: Economic Policy*, 16(4):223–53.

Papay, J. P., Taylor, E. S., Tyler, J. H., and Laski, M. E. (2020). Learning Job Skills from Colleagues at Work: Evidence from a Field Experiment Using Teacher Performance Data. *American Economic Journal: Economic Policy*, 12(1):359–388.

Prat, A. (2002). Should a team be homogeneous? *European Economic Review*, 46(7):1187–1207.

Sandvik, J. J., Saouma, R. E., Seegert, N. T., and Stanton, C. T. (2020). Workplace Knowledge Flows. *The Quarterly Journal of Economics*, 135(3):1635–1680.

WB (2018). Learning to Realize Education's Promise.

WB (2023). MAKING TEACHER POLICY WORK.

Appendix A Appendix Figures

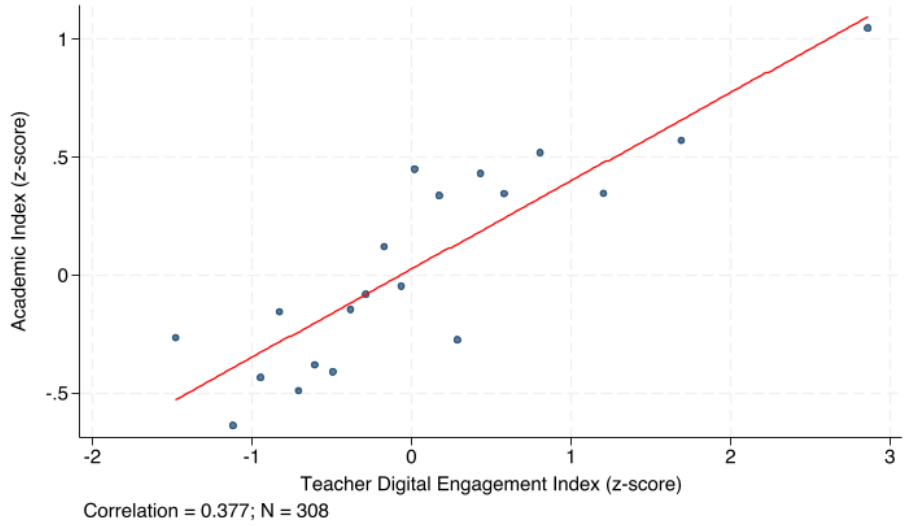
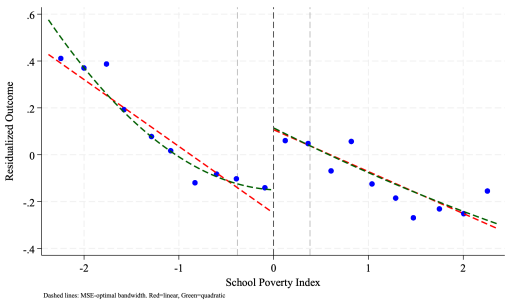


Figure A.1: Academic Index and Teacher Digital Engagement

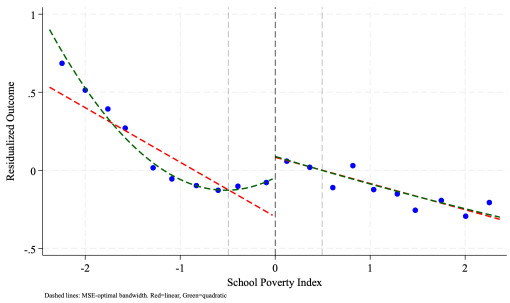
Notes: This figure presents a scatter plot of the Academic Index and the Digital Engagement Index for teachers. The correlation coefficient and number of observations are in the footnotes of each graph.

Figure A.2: Cognitive Outcomes

(a) Language Scores



(b) Math Scores



Notes: This figure presents the effect of the APRENDER program on Math and Language scores.

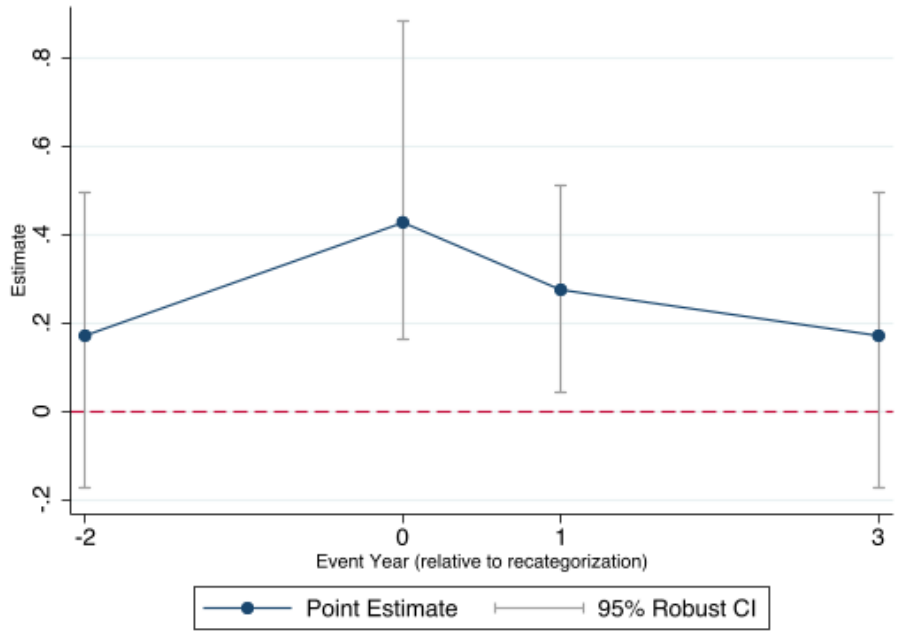


Figure A.3: Dynamic Effect on Students' Academic Index

Notes: This figure presents the dynamic effect of the APRENDER program on the students' academic index.

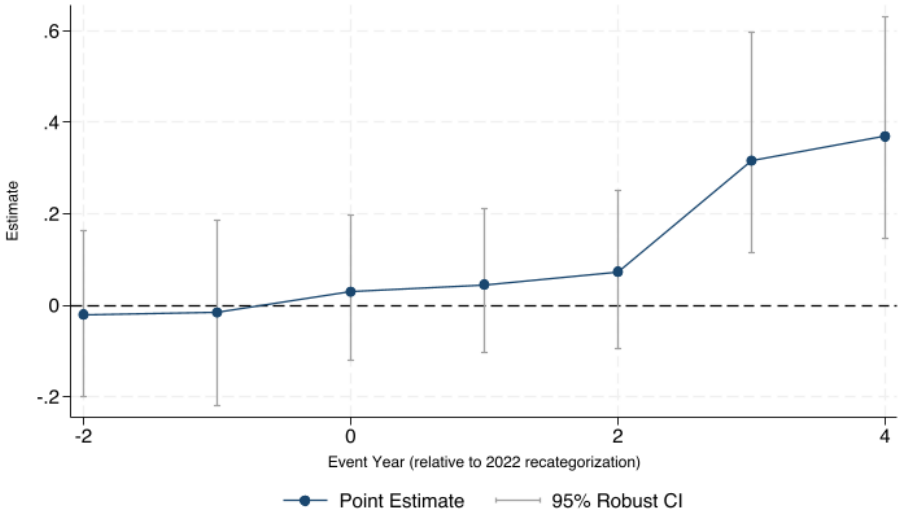


Figure A.4: Dynamic Effect on Students' Engagement with the digital platform

Notes: This figure presents the dynamic effect of the APRENDER program on students' engagement with the digital platform.

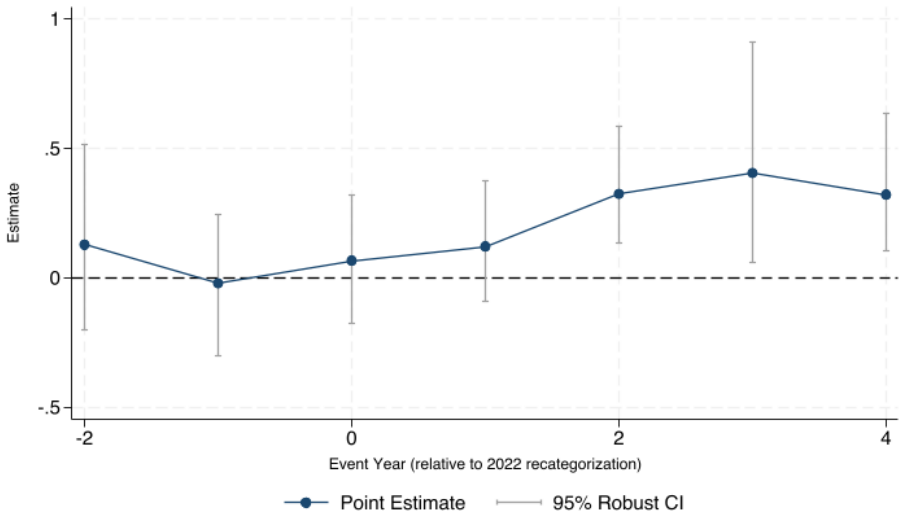


Figure A.5: Dynamic Effect on Permanent Teachers' Digital Engagement Index

Notes: This figure presents the dynamic effect of the APRENDER program on teachers' engagement with the digital platform.

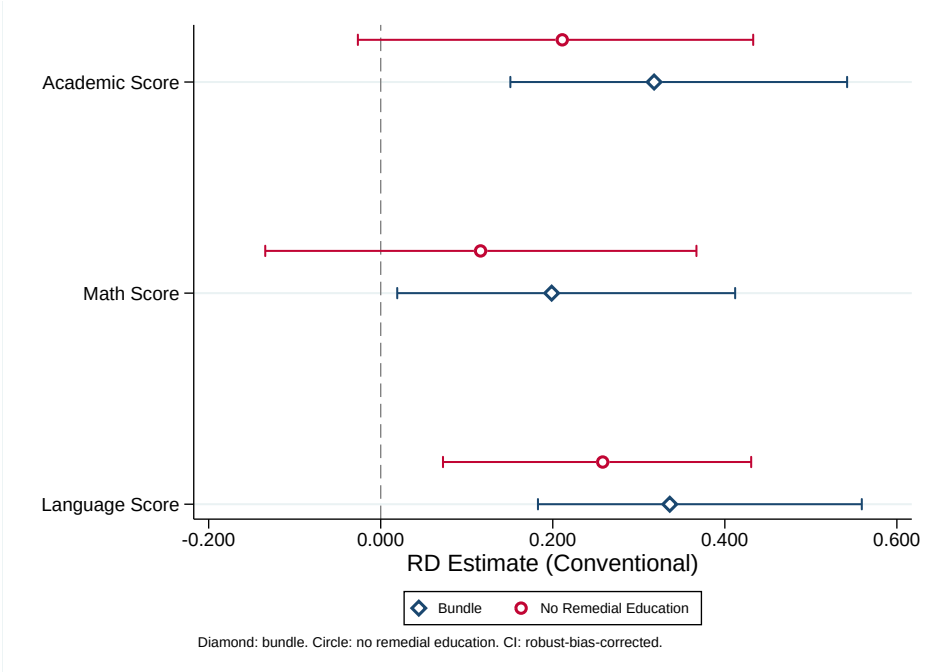


Figure A.6: Moving away from the bundle: Effect on the Academic Index

Notes: This figure presents the results of running Equation 1 in the benchmark bundle estimation and the one removing current remedial education for the test scores sample.

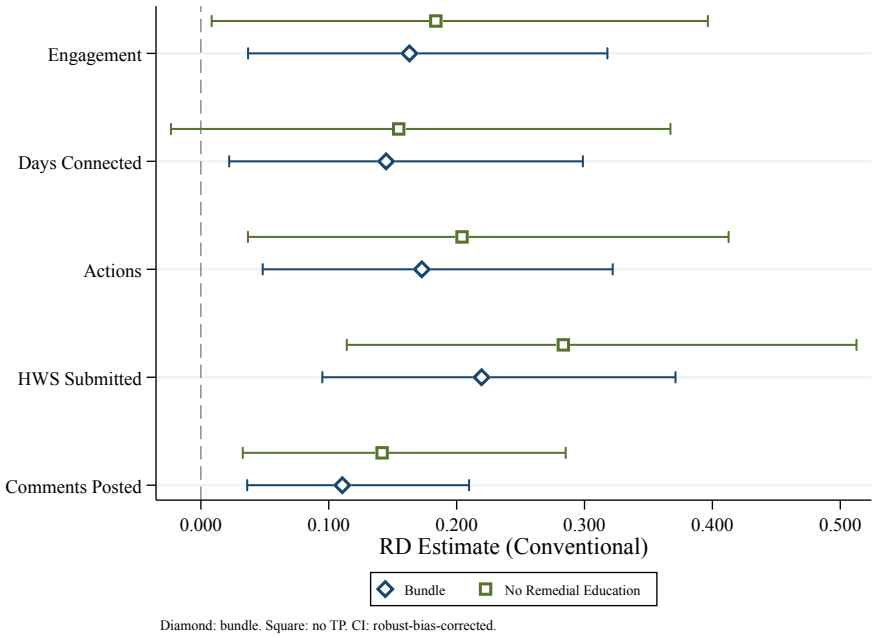


Figure A.7: Moving away from the bundle: Effect on Students' Engagement with the digital platform

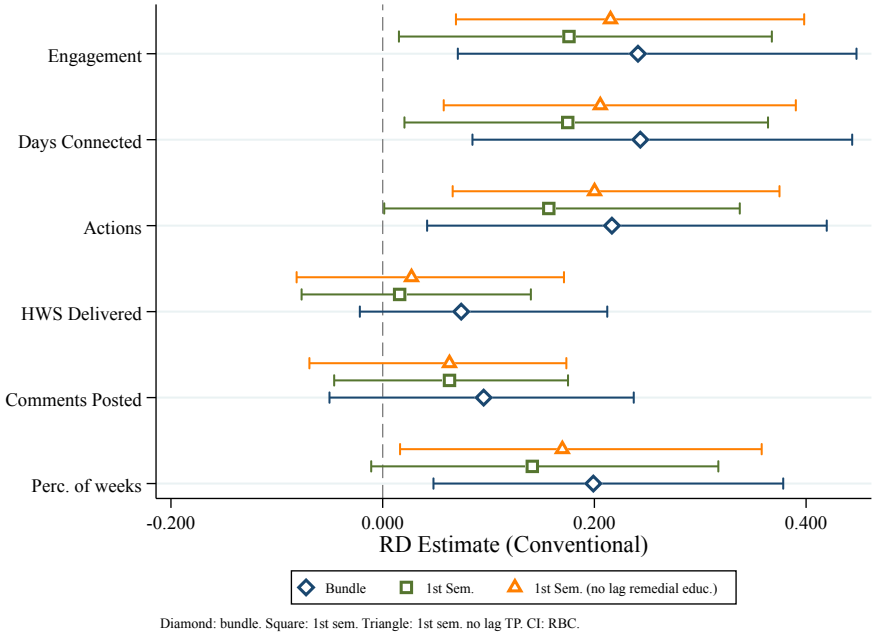
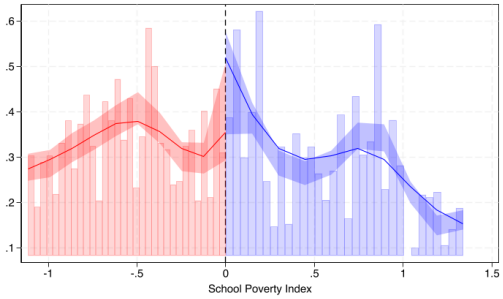


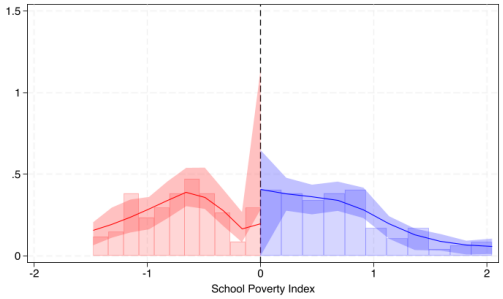
Figure A.8: Moving away from the bundle: Effect on Teachers' Engagement with the digital platform

Figure A.9: Density Test

(a) All Schools



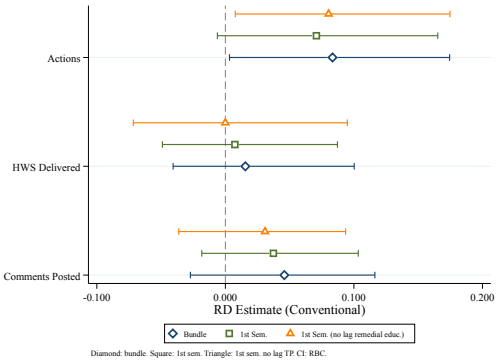
(b) *Aristas* sub-sample



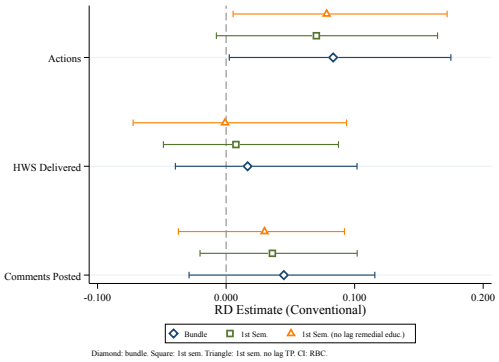
Notes: This figure presents the McCrary test for the poverty index.

Figure A.10: Effects on Teacher Performance using All Permanent Teachers and Imputations

(a) By Number of Students in each Group



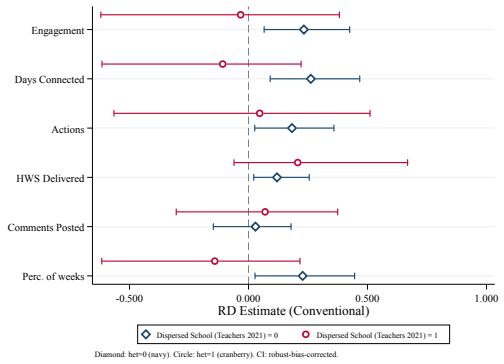
(b) By Number of Groups Taught



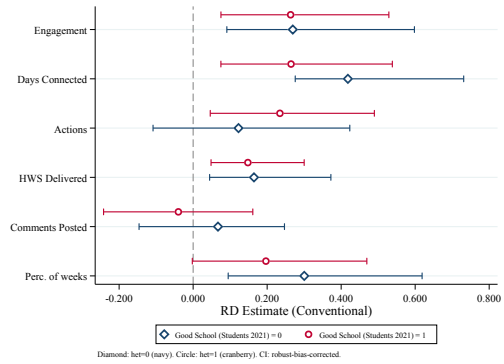
Notes: This figure presents the coefficient plots from running Equation 1 on all permanent teachers in the school using imputed variables. For each outcome variable, we have three specifications, from more general to the more causal on the effects of the expansion of meetings: (i) the bundle effect, (ii) the one restricting the second semester, and (iii) restricting the second semester and removing grades that had students who could have been affected by remedial education the previous year. Panel (a) imputes teacher annual measures weighting by the number of students in each group relative to the total number of students taught by the teacher in a given year. Panel (b) imputes teacher annual measures uniformly across all the groups taught, just by dividing by the number of groups taught in a given year.

Figure A.11: Heterogeneity by Team Characteristics and Intra-School Differences

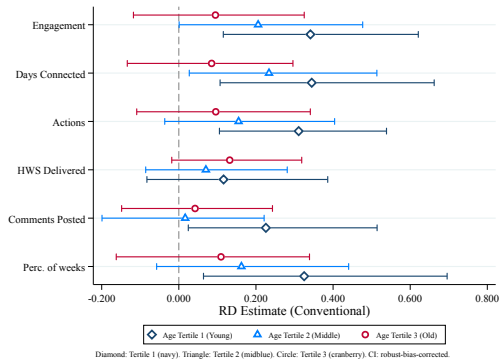
(a) School Number of Actions Dispersion



(b) School Ex-Ante Engagement (Students)



(c) Intra-School Age Tertiles



Notes: This figure presents the main heterogeneous analysis. The first row analyzes heterogeneity exploiting school characteristics. Panel (a) shows heterogeneous effects on schools, split between schools where teachers are highly dispersed in terms of performance in 2021 (red circle) and those that are not (blue diamond). Highly dispersed schools in performance are those whose coefficient of variation of the number of actions done in the digital platform by the entire teacher team in 2021, before the re-categorization, is above the average given their vulnerability index in 2022. Panel (b) plots the effects differentiating by the ex-ante quality of schools in 2021, previous to the reclassification in 2022. We obtain average student engagement on the digital platform for each school in 2021 and determine whether each is above or below its average, given each school's 2022 vulnerability index. Then we plot the ex-ante good schools, given their vulnerability index (red circle), and bad schools (blue diamond) based on past teacher performance. In the second row, we present heterogeneity within schools. In panel (c), we compare the effects across age tertiles within the same school.

Appendix B Appendix Tables

Table B.1: **Teacher Summary Statistics**

	(1)	(2)	(3)	(4)	(5)
	Universe		1 School		
	Mean	SD	Mean	SD	Diff vs Univ.
Panel A: All Teachers					
Age	41.82	10.14	42.56	9.65	0.74
Female (dummy)	0.93	0.25	0.94	0.24	0.00
Permanent Teacher (dummy)	0.77	0.42	0.83	0.38	0.06
Years in Sample	3.70	1.38	3.83	1.35	0.13
Number of Schools per Year	1.40	0.99	1.00	0.00	-0.40
Number of Groups per Year	3.09	3.95	1.78	1.86	-1.30
Unique Teacher IDs	14614		12363		
Unique School IDs	504		504		
Panel B: Permanent Teachers					
Age	42.72	9.77	43.18	9.33	0.45
Female (dummy)	0.94	0.23	0.94	0.23	0.00
Permanent Teacher (dummy)	1.00	0.00	1.00	0.00	0.00
Years in Sample	4.04	1.17	4.11	1.15	0.07
Number of Schools per Year	1.28	0.82	1.00	0.00	-0.28
Number of Groups per Year	2.83	3.73	1.77	1.86	-1.06
Unique Teacher IDs	10539		9341		
Unique School IDs	504		504		
Panel C: Substitute Teachers					
Age	38.82	10.72	39.54	10.56	0.72
Female (dummy)	0.90	0.30	0.90	0.30	-0.00
Permanent Teacher (dummy)	0.00	0.00	0.00	0.00	0.00
Years in Sample	2.58	1.40	2.46	1.40	-0.12
Number of Schools per Year	1.79	1.34	1.00	0.00	-0.79
Number of Groups per Year	3.93	4.49	1.87	1.88	-2.06
Unique Teacher IDs	7345		4535		
Unique School IDs	504		504		

Notes: Columns (1)-(2) report the universe mean and standard deviation. Columns (3)-(4) report means and standard deviations for the 1-school sample. Column (5) reports differences with respect to the universe mean. The bottom of each panel reports unique teacher and school IDs by sample.

Table B.2: **Summary Statistics at the School level**

Summary Statistics at the School level

	Mean	SD	N
Number of teachers	23.95	11.12	2458
Number of 1-school teachers	13.08	6.18	2458
Number of 1-group teachers	10.41	4.88	2303
Number of permanent teachers	16.38	7.78	2456
Number of students	287.95	150.83	2458
Students per teacher	12.22	4.18	2458
Number of teacher turnovers (t-1)	10.10	6.32	2458
Number of permanent teacher turnover (t-1)	4.24	3.72	2458
Mean teacher age	40.92	4.88	2458
Coefficient of variation of teacher age	0.24	0.05	2458
Unique School IDs	504		

Notes: The table reports mean, standard deviation, and sample size at the school-year level, using data from the digital platform panel. The footer row reports unique IDs used in the sample.

Table B.3: **Effect of APRENDER on Academic Outcomes**

	Academic Score			Math Score			Reading Score		
	All	Third	Sixth	All	Third	Sixth	All	Third	Sixth
RD Estimate	0.318***	0.313**	0.319***	0.199**	0.186	0.220*	0.336***	0.308**	0.334***
	(0.083)	(0.129)	(0.103)	(0.087)	(0.126)	(0.105)	(0.079)	(0.115)	(0.085)
	[0.100]	[0.160]	[0.118]	[0.100]	[0.156]	[0.118]	[0.096]	[0.142]	[0.099]
Controls:									
Year FE	No	No	No	No	No	No	No	No	No
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Grade FE	No	No	No	No	No	No	No	No	No
Bandwidth	.481	.453	.544	.58	.481	.616	.431	.434	.486
Mean Dep. Var.	-0.066	-0.041	-0.106	-0.044	0.015	-0.038	-0.076	-0.037	-0.076
SD Dep. Var.	0.933	0.953	0.919	0.923	0.941	0.922	0.962	0.977	0.951
Observations	15,103	7,287	7,816	16,773	8,141	8,632	16,770	8,090	8,680
Eff. Obs. at the Left ; Right	1977 ; 2885	864 ; 1307	1109 ; 1545	2524 ; 3483	1032 ; 1550	1669 ; 1921	1840 ; 2877	863 ; 1415	1157 ; 1647
Eff. Schools at the Left ; Right	31 ; 43	29 ; 40	33 ; 43	36 ; 45	31 ; 43	40 ; 47	25 ; 38	25 ; 39	31 ; 42

Notes: The Table reports the RD specification estimates associating academic performance with the school poverty index. The Academic Score is the simple average of the standardized math and reading test scores from the *Aristas* survey. The tests were conducted in 2017, 2020, and 2023 among third and sixth graders. The table reports the conventional RD point estimate. Conventional standard errors are in parentheses; robust-bias-corrected standard errors are in brackets. Inference is based on the robust-bias-corrected standard error following Cattaneo, Idrobo, and Titiunik (2020). The school poverty index measures the vulnerability of the area where the school is located according to socio-economic indicators; see text for details. It is rescaled so that a school is eligible to be APRENDER if its index is higher than zero. All specifications implement a sharp RD design due to perfect compliance. Clusters at the school level are taken in every specification. All specifications use a triangular kernel, MSE-optimal bandwidth selection, and a local polynomial of order 1. The mean and standard deviation of the dependent variable are measured inside the estimation bandwidth among control schools. Effective observations are those within the local bandwidth. ***, **, and * indicate significance at the 1%, 5%, and 10% levels based on robust-bias-corrected inference.

Table B.4: **Effect of APRENDER on CREA usage at Teacher Level - All Teachers, One School**

	CREA usage - Teachers					
	Engagement Index	Days Connected	Frequency	Actions	HWS Delivered	Comments Posted
RD Estimate	0.193***	0.203***	0.176**	0.165**	0.062	0.067
	(0.068)	(0.068)	(0.066)	(0.064)	(0.047)	(0.045)
	[0.078]	[0.077]	[0.076]	[0.074]	[0.054]	[0.053]
Controls:						
Year FE	No	No	No	No	No	No
Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Grade FE	Yes	Yes	Yes	Yes	Yes	Yes
Bandwidth	.524	.508	.544	.547	.457	.73
Mean Dep. Var.	-0.027	-0.006	-0.014	-0.023	0.041	0.007
SD Dep. Var.	0.954	0.973	0.992	0.955	1.050	0.996
Observations	31,607	31,607	31,607	31,607	31,607	31,607
Eff. Obs. at the Left ; Right	5221 ; 6283	4916 ; 6010	5354 ; 6477	5475 ; 6512	4635 ; 5507	7146 ; 8050
Eff. Schools at the Left ; Right	145 ; 143	138 ; 141	147 ; 148	149 ; 149	132 ; 131	183 ; 178

Notes: The table reports the conventional RD point estimate. Conventional standard errors are in parentheses; robust-bias-corrected standard errors are in brackets. Inference is based on the robust-bias-corrected standard error following Cattaneo, Idrobo, and Titiunik (2020). The school poverty index measures the vulnerability of the area where the school is located according to socio-economic indicators; see text for details. It is rescaled so that a school is eligible to be APRENDER if its index is higher than zero. All specifications implement a sharp RD design due to perfect compliance. Clusters at the school level are taken in every specification. All specifications use a triangular kernel, MSE-optimal bandwidth selection, and a local polynomial of order 1. The mean and standard deviation of the dependent variable are measured inside the estimation bandwidth among control schools. Effective observations are those within the local bandwidth. ***, **, and * indicate significance at the 1%, 5%, and 10% levels based on robust-bias-corrected inference.

Table B.5: **Selection of Teachers using Past Performance Measures - Permanent Teachers, One School**

	CREA usage - Teachers					
	Engagement Index	Days Connected	Perc. of weeks	Homework Tasks	Comments Posted	Actions
RD Estimate	0.071	0.069	0.085	0.110	0.088	0.065
	(0.076)	(0.077)	(0.075)	(0.071)	(0.073)	(0.077)
	[0.089]	[0.089]	[0.087]	[0.083]	[0.085]	[0.090]
Controls:						
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Grade FE	Yes	Yes	Yes	Yes	Yes	Yes
Bandwidth	.539	.553	.624	.657	.634	.54
Mean Dep. Var.	0.029	0.060	0.067	-0.003	0.018	0.010
SD Dep. Var.	0.987	1.030	1.019	0.950	0.963	0.946
Observations	6,886	6,886	6,886	6,886	6,886	6,886
Eff. Obs. at the Left ; Right	1115 ; 1480	1143 ; 1486	1247 ; 1636	1337 ; 1729	1259 ; 1638	1121 ; 1480
Eff. Schools at the Left ; Right	134 ; 141	137 ; 142	150 ; 155	159 ; 160	153 ; 155	134 ; 141

Notes: The table reports the conventional RD point estimate. Conventional standard errors are in parentheses; robust-bias-corrected standard errors are in brackets. Inference is based on the robust-bias-corrected standard error following Cattaneo, Idrobo, and Titiunik (2020). The school poverty index measures the vulnerability of the area where the school is located according to socio-economic indicators; see text for details. It is rescaled so that a school is eligible to be APRENDER if its index is higher than zero. All specifications implement a sharp RD design due to perfect compliance. Grade and year fixed effects are implemented in every specification. Clusters at the school level are taken in every specification. All specifications use a triangular kernel, MSE-optimal bandwidth selection, and a local polynomial of order 1. The mean and standard deviation of the dependent variable are measured inside the estimation bandwidth among control schools. Permanent observations are those within the local bandwidth. ***, **, and * indicate significance at the 1%, 5%, and 10% levels based on robust-bias-corrected inference.

Table B.6: **Balance Tests - Student Characteristics**

	Student Characteristics		
	SES Index	Female	Ever transferred
RD Estimate	0.120	-0.036	-0.009
	(0.074)	(0.023)	(0.037)
	[0.087]	[0.027]	[0.044]
Controls:			
Year FE	Yes	Yes	Yes
Region FE	Yes	Yes	Yes
Grade FE	Yes	Yes	Yes
Bandwidth	.718	.567	.578
Mean Dep. Var.	-0.347	0.494	0.398
SD Dep. Var.	0.636	0.500	0.489
Observations	16,542	19,159	16,997
Eff. Obs. at the Left ; Right	3634 ; 4247	2828 ; 3973	2611 ; 3495
Eff. Schools at the Left ; Right	49 ; 55	35 ; 45	36 ; 45

Notes: The table reports the conventional RD point estimate. Conventional standard errors are in parentheses; robust-bias-corrected standard errors are in brackets. Inference is based on the robust-bias-corrected standard error following Cattaneo, Idrobo, and Titiunik (2020). The school poverty index measures the vulnerability of the area where the school is located according to socio-economic indicators; see text for details. It is rescaled so that a school is eligible to be APRENDER if its index is higher than zero. All specifications implement a sharp RD design due to perfect compliance. Grade, year, and region fixed effects are implemented in every specification. Clusters at the school level are taken in every specification. All specifications use a triangular kernel, MSE-optimal bandwidth selection, and a local polynomial of order 1. The mean and standard deviation of the dependent variable are measured inside the estimation bandwidth among control schools. Effective observations are those within the local bandwidth. ***, **, and * indicate significance at the 1%, 5%, and 10% levels based on robust-bias-corrected inference.

Table B.7: **Effect of APRENDER on Student Academic Score: Placebo Cutoff Tests**

	Control Region			Benchmark	Treatment Region		
	c=-0.75	c=-0.5	c=-0.25	c=0	c=0.25	c=0.5	c=0.75
RD Estimate	0.260	0.333	-0.170	0.318***	0.137	-0.124	0.160
	(0.190)	(0.142)	(0.171)	(0.083)	(0.107)	(0.064)	(0.115)
	[0.274]	[0.214]	[0.199]	[0.100]	[0.124]	[0.116]	[0.141]
Controls:							
Year FE	No	No	No	No	No	No	No
Grade FE	No	No	No	No	No	No	No
Bandwidth	.203	.147	.172	.481	.215	.139	.275
Mean Dep. Var.	0.118	0.059	0.079	-0.066	-0.131	-0.179	-0.302
SD Dep. Var.	0.980	0.945	0.913	0.933	0.936	0.883	0.901
Observations	7,873	7,873	7,873	15,103	7,230	7,230	7,230
Eff. Obs. at the Left ; Right	1050 ; 1209	1025 ; 999	796 ; 445	1977 ; 2885	1260 ; 1209	822 ; 457	1241 ; 1546
Eff. Schools at the Left ; Right	18 ; 18	15 ; 15	11 ; 7	31 ; 43	18 ; 20	14 ; 6	19 ; 20

Notes: The table reports the conventional RD point estimate. Conventional standard errors are in parentheses; robust-bias-corrected standard errors are in brackets. Inference is based on the robust-bias-corrected standard error following Cattaneo, Idrobo, and Titiunik (2020). The ICSC index measures the vulnerability of the area where the school is located according to socio-economic indicators; see text for details. It is rescaled so that a school is eligible to be APRENDER if its index is higher than zero; otherwise, it should be UC. All specifications implement a sharp RD design due to perfect compliance. Grade and year fixed effects are implemented in every specification. Clusters at the school level are taken in every specification. Column (4) is the benchmark with the true assignment cutoff at 0. Columns (1)–(3) use placebo cutoffs within the Control Region (poverty index lower than zero), so schools above the real cutoff are excluded. Columns (5)–(7) use placebo cutoffs within the Treatment Region (poverty index above zero), so schools below the real cutoff are excluded. For all placebo columns the running variable is re-centered at the placebo cutoff.

Table B.8: **Effect of APRENDER on Student CREA Engagement Index: Placebo Cutoff Tests**

	Control Region			Benchmark	Treatment Region		
	c=-0.75	c=-0.5	c=-0.25	c=0	c=0.25	c=0.5	c=0.75
RD Estimate	-0.052	0.136	0.019	0.163**	-0.087	-0.132	-0.047
	(0.106)	(0.173)	(0.173)	(0.061)	(0.083)	(0.105)	(0.105)
	[0.129]	[0.219]	[0.211]	[0.072]	[0.096]	[0.129]	[0.121]
Controls:							
Year FE	No	No	No	No	No	No	No
Grade FE	No	No	No	No	No	No	No
Bandwidth	.314	.155	.168	.452	.167	.143	.235
Mean Dep. Var.	0.195	0.061	0.152	-0.004	-0.026	-0.239	-0.288
SD Dep. Var.	1.013	0.929	1.060	0.940	0.953	0.754	0.716
Observations	349,168	349,168	349,168	697,378	348,210	348,210	348,210
Eff. Obs. at the Left ; Right	58803 ; 65275	34820 ; 35306	34656 ; 32101	98427 ; 117936	45787 ; 31786	34977 ; 31233	52858 ; 62967
Eff. Schools at the Left ; Right	81 ; 105	58 ; 56	52 ; 47	131 ; 129	61 ; 42	45 ; 39	64 ; 74

Notes: The table reports the conventional RD point estimate. Conventional standard errors are in parentheses; robust-bias-corrected standard errors are in brackets. Inference is based on the robust-bias-corrected standard error following Cattaneo, Idrobo, and Titiunik (2020). The ICSC index measures the vulnerability of the area where the school is located according to socio-economic indicators; see text for details. It is rescaled so that a school is eligible to be APRENDER if its index is higher than zero; otherwise, it should be UC. All specifications implement a sharp RD design due to perfect compliance. Grade and year fixed effects are implemented in every specification. Clusters at the school level are taken in every specification. Column (4) is the benchmark with the true assignment cutoff at 0. Columns (1)–(3) use placebo cutoffs within the Control Region (poverty index lower than zero), so schools above the real cutoff are excluded. Columns (5)–(7) use placebo cutoffs within the Treatment Region (poverty index above zero), so schools below the real cutoff are excluded. For all placebo columns the running variable is re-centered at the placebo cutoff.

Table B.9: **Effect of APRENDER on Teacher CREA Engagement Index: Placebo Cutoff Tests**

	Control Region			Benchmark	Treatment Region		
	c=-0.75	c=-0.5	c=-0.25	c=0	c=0.25	c=0.5	c=0.75
RD Estimate	-0.149	0.026	0.105	0.231***	-0.082	-0.327**	0.111
	(0.134)	(0.163)	(0.171)	(0.078)	(0.113)	(0.146)	(0.146)
	[0.162]	[0.200]	[0.218]	[0.090]	[0.130]	[0.175]	[0.169]
Controls:							
Year FE	No	No	No	No	No	No	No
Grade FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Bandwidth	.268	.179	.148	.507	.169	.138	.193
Mean Dep. Var.	0.215	0.133	0.170	0.054	0.110	-0.097	-0.138
SD Dep. Var.	1.057	1.029	1.056	0.974	1.006	0.827	0.854
Observations	13,152	13,152	13,152	26,137	12,985	12,985	12,985
Eff. Obs. at the Left ; Right	2093 ; 2058	1411 ; 1550	1046 ; 1086	3948 ; 4970	1792 ; 1244	1260 ; 972	1610 ; 1993
Eff. Schools at the Left ; Right	73 ; 87	63 ; 62	44 ; 43	138 ; 141	61 ; 42	44 ; 36	54 ; 60

Notes: The table reports the conventional RD point estimate. Conventional standard errors are in parentheses; robust-bias-corrected standard errors are in brackets. Inference is based on the robust-bias-corrected standard error following Cattaneo, Idrobo, and Titiunik (2020). The ICSC index measures the vulnerability of the area where the school is located according to socio-economic indicators; see text for details. It is rescaled so that a school is eligible to be APRENDER if its index is higher than zero; otherwise, it should be UC. All specifications implement a sharp RD design due to perfect compliance. Grade and year fixed effects are implemented in every specification. Clusters at the school level are taken in every specification. Column (4) is the benchmark with the true assignment cutoff at 0. Columns (1)–(3) use placebo cutoffs within the Control Region (poverty index lower than zero), so schools above the real cutoff are excluded. Columns (5)–(7) use placebo cutoffs within the Treatment Region (poverty index above zero), so schools below the real cutoff are excluded. For all placebo columns the running variable is re-centered at the placebo cutoff.

Table B.10: **Pre-treatment Tests - Cognitive Outcomes**

	Academic Score	Reading	Math
RD Estimate	-1.842	-1.981	-0.926
	(14.218)	(37.076)	(3.921)
	[16.678]	[43.631]	[4.614]
Controls:			
Year FE	Yes	Yes	Yes
Region FE	Yes	Yes	Yes
Grade FE			
Bandwidth	.555	.507	.55
Mean Dep. Var.	-0.090	-0.096	-0.120
SD Dep. Var.	0.842	0.963	0.920
Observations	10,034	11,191	11,178
Eff. Obs. at the Left ; Right	1814 ; 2282	1776 ; 2518	2019 ; 2531
Eff. Schools at the Left ; Right	111 ; 90	100 ; 86	111 ; 90

Notes: The table reports the conventional RD point estimate. Conventional standard errors are in parentheses; robust-bias-corrected standard errors are in brackets. Inference is based on the robust-bias-corrected standard error following Cattaneo, Idrobo, and Titiunik (2020). The school poverty index measures the vulnerability of the area where the school is located according to socio-economic indicators; see text for details. It is rescaled so that a school is eligible to be APRENDER if its index is higher than zero. All specifications implement a sharp RD design due to perfect compliance. Grade, year, and region fixed effects are implemented in every specification. Clusters at the school level are taken in every specification. All specifications use a triangular kernel, MSE-optimal bandwidth selection, and a local polynomial of order 1. The mean and standard deviation of the dependent variable are measured inside the estimation bandwidth among control schools. Effective observations are those within the local bandwidth. ***, **, and * indicate significance at the 1%, 5%, and 10% levels based on robust-bias-corrected inference.

Table B.11: **Pre-treatment Tests - CREA Outcomes - Students**

	Engagement Index	Days Connected	Homework Tasks	Comments	Actions
RD Estimate	-0.046	-0.052	-0.053	-0.026	-0.039
	(0.071)	(0.073)	(0.068)	(0.039)	(0.068)
	[0.083]	[0.085]	[0.080]	[0.044]	[0.080]
Controls:					
Year FE	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes
Grade FE	Yes	Yes	Yes	Yes	Yes
Bandwidth	.599	.589	.667	.837	.61
Mean Dep. Var.	-0.035	-0.044	-0.022	-0.023	-0.033
SD Dep. Var.	0.954	0.948	0.980	0.954	0.958
Observations	431,733	431,733	431,733	431,733	431,733
Eff. Obs. at the Left ; Right	71647 ; 101880	71647 ; 97796	83064 ; 110947	108064 ; 134189	73765 ; 101880
Eff. Schools at the Left ; Right	117 ; 91	111 ; 89	129 ; 102	161 ; 130	121 ; 92

Notes: The table reports the conventional RD point estimate. Conventional standard errors are in parentheses; robust-bias-corrected standard errors are in brackets. Inference is based on the robust-bias-corrected standard error following Cattaneo, Idrobo, and Titiunik (2020). The school poverty index measures the vulnerability of the area where the school is located according to socio-economic indicators; see text for details. It is rescaled so that a school is eligible to be APRENDER if its index is higher than zero. All specifications implement a sharp RD design due to perfect compliance. Grade, year, and region fixed effects are implemented in every specification. Clusters at the school level are taken in every specification. All specifications use a triangular kernel, MSE-optimal bandwidth selection, and a local polynomial of order 1. The mean and standard deviation of the dependent variable are measured inside the estimation bandwidth among control schools. Effective observations are those within the local bandwidth. ***, **, and * indicate significance at the 1%, 5%, and 10% levels based on robust-bias-corrected inference.

Table B.12: **Pre-treatment Tests - CREA Outcomes - Teachers - Effective Teachers**

	Engagement Index	Days Connected	Responsiveness	Homework Tasks	Comments	Actions
RD Estimate	-0.040	-0.073	-0.055	-0.021	-0.045	-0.003
	(0.115)	(0.115)	(0.099)	(0.076)	(0.099)	(0.113)
	[0.134]	[0.134]	[0.114]	[0.089]	[0.119]	[0.131]
Controls:						
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Grade FE	Yes	Yes	Yes	Yes	Yes	Yes
Bandwidth	.617	.629	.615	.858	.686	.596
Mean Dep. Var.	-0.002	0.010	0.013	0.033	-0.013	-0.017
SD Dep. Var.	0.992	1.009	1.008	1.069	0.971	0.977
Observations	11,926	11,926	12,170	11,926	11,926	11,926
Eff. Obs. at the Left ; Right	2221 ; 2874	2239 ; 2888	2241 ; 2913	3169 ; 3912	2543 ; 3179	2110 ; 2835
Eff. Schools at the Left ; Right	122 ; 93	124 ; 94	122 ; 93	163 ; 139	133 ; 104	117 ; 91

Notes: The table reports the conventional RD point estimate. Conventional standard errors are in parentheses; robust-bias-corrected standard errors are in brackets. Inference is based on the robust-bias-corrected standard error following Cattaneo, Idrobo, and Titiunik (2020). The school poverty index measures the vulnerability of the area where the school is located according to socio-economic indicators; see text for details. It is rescaled so that a school is eligible to be APRENDER if its index is higher than zero. All specifications implement a sharp RD design due to perfect compliance. Grade, year, and region fixed effects are implemented in every specification. Clusters at the school level are taken in every specification. All specifications use a triangular kernel, MSE-optimal bandwidth selection, and a local polynomial of order 1. The mean and standard deviation of the dependent variable are measured inside the estimation bandwidth among control schools. Effective observations are those within the local bandwidth. ***, **, and * indicate significance at the 1%, 5%, and 10% levels based on robust-bias-corrected inference.

Table B.13: **Balance Tests - School Characteristics**

	School Characteristics								
	Offers Lunch	Offers Breakfast	Multidisciplinary team	Inclusive Classes	Inclusive Teacher	Billingual Teacher	Perc. of Immigrants	Group Size	School Size
RD Estimate	-0.085*	0.045	0.029	-0.152	0.008	0.027	14.080	-0.465	8.322
	(0.051)	(0.086)	(0.026)	(0.102)	(0.022)	(0.079)	(12.924)	(0.881)	(29.006)
	[0.058]	[0.098]	[0.029]	[0.118]	[0.024]	[0.088]	[15.502]	[1.012]	[32.961]
Controls:									
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Bandwidth	.582	.776	.68	.585	.62	.525	.637	.616	.57
Mean Dep. Var.	0.919	0.638	0.037	0.466	0.963	0.158	32.267	20.230	263.126
SD Dep. Var.	0.273	0.481	0.189	0.499	0.188	0.364	142.292	5.069	141.055
Observations	4,075	4,075	4,072	4,075	4,075	4,075	3,428	4,075	4,075
Eff. Obs. at the Left ; Right	820 ; 781	1096 ; 1024	978 ; 902	822 ; 784	884 ; 825	732 ; 734	762 ; 720	884 ; 825	807 ; 775
Eff. Schools at the Left ; Right	171 ; 166	215 ; 201	197 ; 183	172 ; 167	178 ; 174	158 ; 154	180 ; 173	178 ; 174	169 ; 165

Notes: The table reports the conventional RD point estimate. Conventional standard errors are in parentheses; robust-bias-corrected standard errors are in brackets. Inference is based on the robust-bias-corrected standard error following Cattaneo, Idrobo, and Titiunik (2020). The school poverty index measures the vulnerability of the area where the school is located according to socio-economic indicators; see text for details. It is rescaled so that a school is eligible to be APRENDER if its index is higher than zero. All specifications implement a sharp RD design due to perfect compliance. Grade, year, and region fixed effects are implemented in every specification. Clusters at the school level are taken in every specification. All specifications use a triangular kernel, MSE-optimal bandwidth selection, and a local polynomial of order 1. The mean and standard deviation of the dependent variable are measured inside the estimation bandwidth among control schools. Effective observations are those within the local bandwidth. ***, **, and * indicate significance at the 1%, 5%, and 10% levels based on robust-bias-corrected inference.

Table B.14: **Balance Tests - Special Teachers and Programs**

	Special Teacher Roles and Programs					
	Vice-Principal	Dean	Support Teacher	Ceibal Support Teacher	Community Teracher	Remedial Education Teacher
RD Estimate	-0.004	0.042	-0.139	0.072	0.199***	-0.000
	(0.077)	(0.080)	(0.098)	(0.091)	(0.068)	(0.004)
	[0.089]	[0.093]	[0.109]	[0.106]	[0.078]	[0.005]
Controls:						
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Bandwidth	.601	.657	.625	.695	.607	.421
Mean Dep. Var.	0.155	0.676	0.305	0.412	0.698	0.010
SD Dep. Var.	0.362	0.468	0.460	0.492	0.459	0.101
Observations	4,075	4,075	4,075	4,075	4,075	4,075
Eff. Obs. at the Left ; Right	854 ; 805	948 ; 884	894 ; 830	1005 ; 925	873 ; 814	553 ; 601
Eff. Schools at the Left ; Right	173 ; 171	190 ; 181	181 ; 174	202 ; 188	175 ; 172	128 ; 131

Notes: The table reports the conventional RD point estimate. Conventional standard errors are in parentheses; robust-bias-corrected standard errors are in brackets. Inference is based on the robust-bias-corrected standard error following Cattaneo, Idrobo, and Titiunik (2020). The school poverty index measures the vulnerability of the area where the school is located according to socio-economic indicators; see text for details. It is rescaled so that a school is eligible to be APRENDER if its index is higher than zero. All specifications implement a sharp RD design due to perfect compliance. Grade, year, and region fixed effects are implemented in every specification. Clusters at the school level are taken in every specification. All specifications use a triangular kernel, MSE-optimal bandwidth selection, and a local polynomial of order 1. The mean and standard deviation of the dependent variable are measured inside the estimation bandwidth among control schools. Effective observations are those within the local bandwidth. ***, **, and * indicate significance at the 1%, 5%, and 10% levels based on robust-bias-corrected inference.

Table B.15: **Effect of APRENDER on Student Academic Score: Robustness to Specification**

	(1)	(2)	(3)	(4)	(5)
RD Estimate	0.318***	0.331***	0.309***	0.334***	0.189***
	(0.083)	(0.097)	(0.096)	(0.090)	(0.048)
	[0.100]	[0.109]	[0.108]	[0.105]	
Controls:					
Year FE	Yes	Yes	Yes	Yes	Yes
Grade FE	Yes	Yes	Yes	Yes	Yes
Bandwidth	.481	.794	.368	.398	.
Mean Dep. Var.	-0.066	0.006	-0.194	-0.140	0.248
SD Dep. Var.	0.933	0.944	0.931	0.927	1.007
Observations	15,103	15,103	15,103	15,103	15,103
Eff. Obs. at the Left ; Right	1977 ; 2885	3599 ; 4334	1020 ; 2135	1196 ; 2356	7873 ; 7230
Eff. Schools at the Left ; Right	31 ; 43	53 ; 63	17 ; 31	20 ; 33	110 ; 103

Notes: The table reports the conventional RD point estimate. Conventional standard errors are in parentheses; robust-bias-corrected standard errors are in brackets (except column (5)). Inference is based on the robust-bias-corrected standard error following Cattaneo, Idrobo, and Titiunik (2020). Column (5) reports conventional clustered standard errors only. The school poverty index measures the vulnerability of the area where the school is located according to socio-economic indicators; see text for details. It is rescaled so that a school is eligible to be APRENDER if its index is higher than zero. All specifications implement a sharp RD design due to perfect compliance. Grade and year fixed effects are implemented in every specification. Clusters at the school level are taken in every specification. Column (1) is the benchmark: triangular kernel, MSE-optimal bandwidth, and local linear polynomial. Column (2) uses a quadratic polynomial (p=2). Column (3) uses a CER-optimal bandwidth. Column (4) uses a uniform kernel. Column (5) uses a global linear polynomial estimated by OLS using the assignment indicator (x>0) as the treatment variable, exploiting perfect compliance; no robust-bias-corrected standard error is available for this specification.

Table B.16: **Effect of APRENDER on Student CREA Engagement Index: Robustness to Specification**

	(1)	(2)	(3)	(4)	(5)
RD Estimate	0.161**	0.176**	0.113	0.169***	0.103**
	(0.061)	(0.071)	(0.075)	(0.069)	(0.049)
	[0.072]	[0.081]	[0.081]	[0.075]	
Controls:					
Year FE	Yes	Yes	Yes	Yes	Yes
Grade FE	Yes	Yes	Yes	Yes	Yes
Bandwidth	.453	.737	.33	.341	.
Mean Dep. Var.	-0.004	0.034	-0.077	-0.074	0.277
SD Dep. Var.	0.940	0.951	0.893	0.894	1.126
Observations	697,378	697,378	697,378	697,378	697,378
Eff. Obs. at the Left ; Right	98427 ; 117936	151926 ; 181290	66362 ; 92189	68018 ; 95251	349168 ; 348210
Eff. Schools at the Left ; Right	131 ; 129	183 ; 182	95 ; 105	99 ; 108	307 ; 267

Notes: The table reports the conventional RD point estimate. Conventional standard errors are in parentheses; robust-bias-corrected standard errors are in brackets (except column (5)). Inference is based on the robust-bias-corrected standard error following Cattaneo, Idrobo, and Titiunik (2020). Column (5) reports conventional clustered standard errors only. The school poverty index measures the vulnerability of the area where the school is located according to socio-economic indicators; see text for details. It is rescaled so that a school is eligible to be APRENDER if its index is higher than zero. All specifications implement a sharp RD design due to perfect compliance. Grade and year fixed effects are implemented in every specification. Clusters at the school level are taken in every specification. Column (1) is the benchmark: triangular kernel, MSE-optimal bandwidth, and local linear polynomial. Column (2) uses a quadratic polynomial (p=2). Column (3) uses a CER-optimal bandwidth. Column (4) uses a uniform kernel. Column (5) uses a global linear polynomial estimated by OLS using the assignment indicator (x>0) as the treatment variable, exploiting perfect compliance; no robust-bias-corrected standard error is available for this specification.

Table B.17: **Effect of APRENDER on Teacher CREA Engagement Index: Robustness to Specification**

	(1)	(2)	(3)	(4)	(5)
RD Estimate	0.225***	0.251***	0.173*	0.262***	0.140***
	(0.077)	(0.084)	(0.088)	(0.083)	(0.051)
	[0.089]	[0.093]	[0.095]	[0.092]	
Controls:					
Year FE	Yes	Yes	Yes	Yes	Yes
Grade FE	Yes	Yes	Yes	Yes	Yes
Bandwidth	.513	.945	.373	.388	.
Mean Dep. Var.	0.053	0.121	-0.003	0.009	0.250
SD Dep. Var.	0.972	1.020	0.937	0.942	1.120
Observations	26,127	26,127	26,127	26,127	26,127
Eff. Obs. at the Left ; Right	3969 ; 4990	7471 ; 9049	2786 ; 3812	2978 ; 3924	13147 ; 12980
Eff. Schools at the Left ; Right	140 ; 141	215 ; 213	107 ; 111	109 ; 114	307 ; 267

Notes: The table reports the conventional RD point estimate. Conventional standard errors are in parentheses; robust-bias-corrected standard errors are in brackets (except column (5)). Inference is based on the robust-bias-corrected standard error following Cattaneo, Idrobo, and Titiunik (2020). Column (5) reports conventional clustered standard errors only. The school poverty index measures the vulnerability of the area where the school is located according to socio-economic indicators; see text for details. It is rescaled so that a school is eligible to be APRENDER if its index is higher than zero. All specifications implement a sharp RD design due to perfect compliance. Grade and year fixed effects are implemented in every specification. Clusters at the school level are taken in every specification. Column (1) is the benchmark: triangular kernel, MSE-optimal bandwidth, and local linear polynomial. Column (2) uses a quadratic polynomial (p=2). Column (3) uses a CER-optimal bandwidth. Column (4) uses a uniform kernel. Column (5) uses a global linear polynomial estimated by OLS using the assignment indicator (x>0) as the treatment variable, exploiting perfect compliance; no robust-bias-corrected standard error is available for this specification.

Table B.18: **Effect of APRENDER on Student Academic Score: Extra Controls**

	(1)	(2)	(3)	(4)
RD Estimate	0.311***	0.328***	0.328***	0.288***
	(0.083)	(0.085)	(0.084)	(0.077)
	[0.099]	[0.103]	[0.102]	[0.094]
Controls:				
Year FE	Yes	Yes	Yes	Yes
Grade FE	Yes	Yes	Yes	Yes
Bandwidth	.483	.474	.477	.495
Mean Dep. Var.	-0.066	-0.080	-0.066	-0.066
SD Dep. Var.	0.933	0.936	0.933	0.933
Observations	15,103	15,103	15,103	13,133
Eff. Obs. at the Left ; Right	1977 ; 2885	1855 ; 2885	1977 ; 2885	1778 ; 2579
Eff. Schools at the Left ; Right	31 ; 43	29 ; 43	31 ; 43	31 ; 43

Notes: The table reports the conventional RD point estimate. Conventional standard errors are in parentheses; robust-bias-corrected standard errors are in brackets. Inference is based on the robust-bias-corrected standard error following Cattaneo, Idrobo, and Titiunik (2020). The school poverty index measures the vulnerability of the area where the school is located according to socio-economic indicators; see text for details. It is rescaled so that a school is eligible to be APRENDER if its index is higher than zero. All specifications implement a sharp RD design due to perfect compliance. Grade and year fixed effects are implemented in every specification. Clusters at the school level are taken in every specification. All specifications use a triangular kernel, MSE-optimal bandwidth selection, and a local polynomial of order 1. The mean and standard deviation of the dependent variable are measured inside the estimation bandwidth among control schools. Effective observations are those within the local bandwidth. ***, **, and * indicate significance at the 1%, 5%, and 10% levels based on robust-bias-corrected inference. Controls are added cumulatively: (1) benchmark estimation including group size; (2) adds whether school offers lunch and milk; (3) elementary school size; (4) adds student socioeconomic vulnerability index.

Table B.19: **Effect of APRENDER on Student CREA Engagement Index: Extra Controls**

	(1)	(2)	(3)	(4)	(5)
RD Estimate	0.161**	0.162**	0.157**	0.158**	0.151**
	(0.061)	(0.059)	(0.060)	(0.061)	(0.061)
	[0.072]	[0.070]	[0.072]	[0.072]	[0.072]
Controls:					
Year FE	Yes	Yes	Yes	Yes	Yes
Grade FE	Yes	Yes	Yes	Yes	Yes
Bandwidth	.453	.464	.461	.461	.466
Mean Dep. Var.	-0.004	0.000	0.000	0.000	0.001
SD Dep. Var.	0.940	0.942	0.942	0.942	0.942
Observations	697,378	697,378	697,378	697,378	697,378
Eff. Obs. at the Left ; Right	98427 ; 117936	100008 ; 122819	100008 ; 121881	100008 ; 121881	100355 ; 122819
Eff. Schools at the Left ; Right	131 ; 129	134 ; 133	134 ; 132	134 ; 132	134 ; 133

Notes: The table reports the conventional RD point estimate. Conventional standard errors are in parentheses; robust-bias-corrected standard errors are in brackets. Inference is based on the robust-bias-corrected standard error following Cattaneo, Idrobo, and Titiunik (2020). The school poverty index measures the vulnerability of the area where the school is located according to socio-economic indicators; see text for details. It is rescaled so that a school is eligible to be APRENDER if its index is higher than zero. All specifications implement a sharp RD design due to perfect compliance. Grade and year fixed effects are implemented in every specification. Clusters at the school level are taken in every specification. All specifications use a triangular kernel, MSE-optimal bandwidth selection, and a local polynomial of order 1. The mean and standard deviation of the dependent variable are measured inside the estimation bandwidth among control schools. Effective observations are those within the local bandwidth. ***, **, and * indicate significance at the 1%, 5%, and 10% levels based on robust-bias-corrected inference. Controls are added cumulatively: (1) benchmark estimation; (2) whether school offers lunch and milk; (3) elementary school size; (4) number of effective (non-substitute) teachers in the group; (5) total number of teachers and substitute teachers.

Table B.20: **Effect of APRENDER on Teacher CREA Engagement Index: Extra Controls**

	(1)	(2)	(3)	(4)	(5)
RD Estimate	0.229***	0.223***	0.210**	0.215**	0.210**
	(0.077)	(0.076)	(0.077)	(0.079)	(0.079)
	[0.089]	[0.088]	[0.090]	[0.093]	[0.094]
Controls:					
Year FE	Yes	Yes	Yes	Yes	Yes
Grade FE	Yes	Yes	Yes	Yes	Yes
Bandwidth	.511	.519	.492	.503	.509
Mean Dep. Var.	0.052	0.040	0.055	0.053	0.053
SD Dep. Var.	0.971	0.967	0.973	0.972	0.972
Observations	26,130	26,130	26,130	26,130	26,130
Eff. Obs. at the Left ; Right	3969 ; 4969	4068 ; 5078	3914 ; 4969	3947 ; 4969	3947 ; 4969
Eff. Schools at the Left ; Right	140 ; 141	142 ; 141	138 ; 141	138 ; 141	138 ; 141

Notes: The table reports the conventional RD point estimate. Conventional standard errors are in parentheses; robust-bias-corrected standard errors are in brackets. Inference is based on the robust-bias-corrected standard error following Cattaneo, Idrobo, and Titiunik (2020). The school poverty index measures the vulnerability of the area where the school is located according to socio-economic indicators; see text for details. It is rescaled so that a school is eligible to be APRENDER if its index is higher than zero. All specifications implement a sharp RD design due to perfect compliance. Grade and year fixed effects are implemented in every specification. Clusters at the school level are taken in every specification. All specifications use a triangular kernel, MSE-optimal bandwidth selection, and a local polynomial of order 1. The mean and standard deviation of the dependent variable are measured inside the estimation bandwidth among control schools. Effective observations are those within the local bandwidth. ***, **, and * indicate significance at the 1%, 5%, and 10% levels based on robust-bias-corrected inference. Controls are added cumulatively: (1) benchmark estimation; (2) whether school offers lunch and milk; (3) elementary school size; (4) number of effective (non-substitute) teachers in the group; (5) total number of teachers and substitute teachers.

Appendix C Appendix: Institutional Details

Teacher hiring. Schools announce permanent-position vacancies on national and regional portals in the academic year preceding hiring. Teachers within the same region submit ranked preferences, and the system runs a Deferred Acceptance algorithm in which teachers are ordered primarily by seniority, making truthful revelation strategic. If positions remain unfilled among certified applicants, the process repeats within the pool of non-certified and intern teachers. The market for substitute positions operates similarly but year-round. Permanent teachers are subject to a two-year window of rigidity before they may request a transfer (absent critical circumstances such as cross-regional relocation).

Additional school-level programs. Beyond APRENDER, schools may host non-traditional specialized staff. *Community teachers* work to reconnect high-absenteeism students and their families with the school, through remedial education and home visits. It is intended for the most vulnerable schools, but assignment to this program is not determined by any specific rule. *Inclusion (or support) teachers* serve students with permanent or temporary learning difficulties, including disabilities. Eligibility for this program is also not determined by any specific rule. We control for participation in both programs in our main analysis to ensure

they do not bias our estimates.

Poverty index construction. The national primary-education authority computes the poverty index every five years using a representative survey of students' families. The index aggregates three dimensions: (i) *household educational level* (educational attainment of the student's mother); (ii) *household socioeconomic level* (e.g., overcrowding and access to drinking water); and (iii) *social integration level* (fraction of households with children not enrolled in school and number of families living in irregular settlements) (ANEP, 2016). These inputs are predetermined relative to school decisions and outcomes. Two editions of the index are relevant for our sample: the 2015 index (applied 2017–2021) and the 2022 index (applied 2022–2024). We rescale each edition so that its eligibility cutoff is at zero and its standard deviation is one.

Scale of the in-kind transfer component. Approved projects receive a fixed per-student amount disbursed as a one-time payment in July. In 2020, this corresponded to approximately USD 4.78 per student at the average exchange rate,[12]. Transfers must be used within the academic year, with a single optional one-year renewal.

[12] Annual average UYU/USD exchange rate from IMF (https://data.imf.org/).