Do More Public Sector School Resources Increase Learning Outcomes? Evidence from a Comprehensive Education Reform in Peruvian Secondary Schools

Jorge M. Agüero (†),[a] Marta Favara,[b] Catherine Porter,[c] Alan Sánchez[d]

Abstract

We evaluate a large-scale government reform in Peruvian public secondary schools that lengthened the school day and invested in pedagogy, staffing, and infrastructure. Using a fuzzy regression discontinuity design, we find that the program increased math and reading test scores by approximately 0.185 SD and 0.103 SD, respectively. For math only, we estimate that instructional time contributes approximately two-thirds of the effect, suggesting the importance of complementary inputs. The reform effectively enhanced school resources and increased students' overall study time. Relative to other education interventions in Low- and Middle-Income Countries, the program delivered above-average learning gains but was relatively expensive.

[a] *University of Connecticut, Department of Economics and El Instituto, USA.*
[b] *Oxford Department of International Development, University of Oxford, Oxford, UK. Email: marta.favara@qeh.ox.ac.uk)*
[c] *Lancaster University, United Kingdom. Email: catherine.porter@lancaster.ac.uk*
[d] *Oxford Department of International Development, University of Oxford, Oxford, UK, and Grupo de Análisis para el Desarrollo (GRADE), Lima, Peru. Corresponding author. Email: alan.sanchez@qeh.ox.ac.uk*

Data availability statement: This paper uses confidential data from the Ministry of Education in Peru (MINEDU). The data can be obtained by filing a request online directly with MINEDU at: https://enlinea.minedu.gob.pe/. The authors are willing to assist (Alan Sánchez, asanchez@grade.org.pe). Data from the Young Lives study is accessible via the UK data service at (https://ukdataservice.ac.uk/). Replication files are available from the authors upon request.

Disclosure statement: The authors have nothing to disclose. Favara, Porter and Sánchez were supported by ESRC/GCRF Grant "Inequality of opportunity in Peru: How can young people develop relevant skills and find decent employment in a rapidly changing labour market?" (ES/S004564/1). A previous version of this study by Agüero was supported by GRADE/FORGE.

Acknowledgements: This research is dedicated to the memory of Jorge Agüero, our friend and co-author who passed away, too soon, in 2023, having contributed hugely to the project. We thank the personnel of the Ministry of Education in Peru for facilitating access to the data and for answering all our questions about the *Jornada Escolar Completa* and providing useful comments. Thanks also to Paul Glewwe, Maria Padilla-Romo and Diego Vera Cossio for their valuable comments. We are grateful to Dianela Espinoza, Alessandra Hidalgo and Andrea Cornejo for excellent research assistance. We thank seminar participants at the Southern Economic Association, NBER Education Program Meeting, Universidad Pablo de Olavide, *Universidad de Piura*, *Universidad del Pacífico*, the Midwest Economic Development Conference, *Grupo de Análisis para el Desarrollo* (GRADE), UConn Neag School of Education, the Inter-American Development Bank, Newcastle University and Lancaster University.

1. Introduction

Despite years of investment and experimentation, the learning gap for students in Low- and Middle-Income Countries (LMICs) remains stubbornly large and significant. Programs to increase school inputs, particularly pedagogical inputs, have had positive effects, but there is wide variation. Success has been attenuated when projects are implemented at scale, especially within the public school system (Evans and Popova 2015; Ganimian and Murnane, 2016; Glewwe and Muralidharan, 2016; Evans and Yuan, 2022; Angrist et al., 2025). Recent studies have shown that in LMICs simply augmenting one aspect of school resources may have a limited impact on learning outcomes due to complementarities with other inputs that are not improved at the same time (Mbiti et al., 2019; Kerwin and Thornton, 2021). One input that has been a focus in LMICs is time: extending the school day in public schools (Holland et al., 2014; Ganimian and Murnane, 2016). Overall, the empirical evidence on the effectiveness of extending instructional time is not conclusive, and isolating its causal impact from other factors has proven difficult.[1,2]

Evidence from high-income contexts suggests that extended instructional time may be effective when bundled with complementary reforms. Fryer (2014) showed that academic achievement in low-performing public schools in the USA was significantly improved through the introduction of a package of best practices from high-performing charter schools, including increased instructional time as well as other pedagogical improvements. However, it remains an open question whether this approach could also be effective in an LMIC public school system. We study the *Jornada Escolar Completa* (*JEC*) government educational reform in Peru, an upper-middle-income country that has consistently underachieved in educational performance despite rapid macroeconomic growth and expanded school access.[3] The program added 2 instructional hours (of 45 minutes each) per day in (initially) 1,000 public secondary schools nationwide to match the number of instructional hours in private schools (Saavedra and Gutierrez, 2020). In addition to the increased instruction time, *JEC* aimed to improve IT resources, increase the number of teaching and non-teaching staff, and improve pedagogical resources for teachers and students.

To be eligible, schools were required to have eight or more sections.[4] The choice of eight sections as the cutoff was i) arbitrary, calculated based on budget constraints and number of classes, and ii) difficult to manipulate, since it was based on census data from two years prior to the intervention. We capitalize on both these aspects for our identification strategy.

We use two large administrative data sources to evaluate the effect of the program: the 2015 *Evaluación Censal de Estudiantes de Secundaria* (*ECE-S*), the official national math and reading assessment for eighth-graders; and the 2015-16 *Semáforo Escuela*, a comprehensive school survey administered by the Ministry of Education (MINEDU).[5] We further investigate mechanisms by analyzing Young Lives (YL) longitudinal data from 2016 (and before) for a cohort of adolescents aged 15, and the 2016 teacher survey, *Encuesta Nacional a Docentes* (*ENDO*).

We employ a fuzzy regression discontinuity design based on the administrative cutoff. We confirm a discontinuously higher probability of participating in the program at this cutoff, moving from near zero below to almost 60 percent above. Smoothness tests applied to a large set of pre-intervention variables validate the arbitrary nature of the threshold. Our main specification is a global polynomial regression discontinuity approach that exploits all available data. As our main robustness check, we use local randomization restricted to the narrowest window (only schools with 7 and 8 sections), allowing us to relax the assumption that the running variable is continuous. In this case, the discontinuity in treatment eligibility acts like a randomized experiment in the vicinity of the cut-off— whether a school has 7 or 8 sections is akin to a coin flip. We find that participating in *JEC* increased math and reading scores by 18.5 and 10.3 percent of a standard deviation (SD), respectively, after one year. This represents closing the observed gap between private and public schools by approximately 25 percent in math and 11 percent in reading. The point estimates vary across samples, empirical approach, and dataset, but remain robust to a battery of robustness checks, including when restricting the sample to urban schools, when using the local randomization approach, and using the Young Lives dataset. In addition, inference for math outcomes remains statistically significant when using bias-aware bounded second derivative (BSD) confidence intervals for local polynomial regressions, following Kolesár and Rothe (2018).[6]

We benchmark the effectiveness and cost-effectiveness of *JEC* vis-a-vis other interventions in LMICs using the Learning-Adjusted Years of Schooling (LAYS) methodology by Angrist et al. (2025), a unified measure of learning gains relative to the annual learning rate of a high-quality education environment. We estimate that the impact (averaged across math and reading) of 0.18 LAYS is in the upper half of the distribution of 200 education interventions evaluated in 52 countries, more effective than the typical general teacher training and input-only interventions, but below the typical top-end effects seen in cutting-edge pilot studies of targeted tutoring, and of *teaching at the right level*.

Investing in infrastructure, staff and pedagogy is costly, however. The total cost of *JEC* was approximately USD 571 per student (expressed in 2015 USD), translating into a cost-effectiveness of 0.032 LAYS per USD 100 spent. Relative to the subset of interventions with published cost data and statistically significant impacts (Angrist et al., 2025), this places *JEC* in the lower tail of estimated cost-effectiveness. Note that most cost-effective interventions tend to be small-scale NGO- or researcher-led pilots, and there are few purely government-led interventions to directly compare with. Furthermore, our cost-effectiveness estimate likely represents a lower bound, given several conservative assumptions—including the assumption that impacts last for only one year.[7] Notwithstanding these considerations, the cost-effectiveness of *JEC* is comparable to that of providing full scholarships for secondary education (Duflo et al., 2021).

Although *JEC* is an overall school reform, a key feature is the extended instructional time. We can isolate the impact of the extra instructional time for math by exploiting the fact that the reform increased instructional hours for math more than for reading, by one hour per week. We use a student-fixed-effects approach together with the discontinuity created by the reform (difference-in-discontinuities) and the assumption that the impact of each additional pedagogical hour (and any other JEC effect) is the same across the two subjects. Our results suggest that one hour of additional math instruction contributes an effect size of approximately 32% of the total improvement in math test scores, which is around 64% of the total effect of *JEC* on math, as instructional time for math increased by two hours per week.

We explore several other mechanisms that may help explain the impact of *JEC*. Since we do not observe all potential channels, the evidence should be interpreted as suggestive. We find evidence confirming that school resources improved in *JEC* schools. The extension of the school day was designed to be accompanied by increased pedagogical assistance to teachers, information-technology for the school overall (e.g., language labs for English instruction), and more staff. We show a significant improvement in the amount of pedagogical support received by *JEC* schools. We also find an increase in access to computers as well as higher numbers of support staff, such as school psychologists.

Estimates of the total impact of the reform on learning may also reflect compensating behavioral responses by students, parents, or teachers that either reinforce or offset the program mechanisms (Levin and Tsang, 1987; Todd and Wolpin, 2013). We rule out upward bias due to student (or teacher) sorting (Padillo-Romo, 2022). A further concern is that the impact of the program could be offset by teachers and students reducing their effort per hour. Students may also reduce the time they spend studying at home, and parents might change their behavior by helping their children less in response to longer school days.[8] We find little evidence of such offsetting behaviors. Teachers' behaviors and attitudes do not change systematically following the reform. Students spend around 30 minutes less per day studying at home—less than the daily increase in school hours—instead, time devoted to domestic chores and other leisure activities declines. We find partial evidence that parents spend less time explaining school topics and recommending books to their children, consistent with students spending fewer hours at home due to the reform. Overall, these behavioral adjustments are modest and are unlikely to attenuate learning gains.

Our findings contribute to the literature on how to improve learning outcomes in developing countries, and importantly in the public-school context, including by extending instructional time, as well as by introducing broader pedagogical improvements. Ganimian & Murnane (2016) conclude that improving resources only works when the student experience improves. Other specific LMIC studies (Agüero and Beleche, 2013; Bellei, 2009; Cerdán-Infantes and Vermeersch, 2007; Hincapié, 2016; Cabrera-Hernández, 2020; Dominguez and Ruffini, 2023; Padilla-Romo, 2022) also show that interventions targeting time in school alone typically produce modest and uneven gains. However, Padilla-Romo (2022) argues that in many of these country-specific studies, isolating the causal effect of the reforms has been difficult. Lavy (2015) uses cross-country data (PISA) from 50 countries to show that one additional instructional hour in each subject (mathematics, science, or reading) improves test scores by 0.06 SD of the test score distribution in that subject, holding overall instructional time constant. However, the effect falls to 0.025 SD for LMICs (see also Rivkin and Schiman, 2015). Focusing on the effect of one extra hour of instructional time in math (subject to our strong assumptions), the results for *JEC* are larger than those observed in other LMICs, and comparable to Lavy's (2015) finding across the PISA (higher income) sample. The total effect size in math and reading suggests that other investments also played a role.

A complementary body of research highlights the impact of pedagogy: Cilliers et al. (2020) show that teacher-focused reforms such as coaching can yield substantial learning improvements. Angrist et al. (2025) document that pedagogical interventions in LMICs have amongst the highest learning gains across a large sample. Our analysis of *JEC* and its components shows a clear change in children's experiences through improved access to facilities, increased staffing, and access to pedagogical support with a corresponding increase in student achievement, though we are unable to fully quantify the contribution of pedagogy to the overall program effect.

Our results show that a large-scale extension of instructional time can be effective when combined with complementary investments in pedagogy and school resources. In this sense, *JEC* operates as a bundled school reform, and not merely as a time-based intervention, altering multiple dimensions of students' learning environments. Benchmarking the effect sizes from *JEC* against the LAYS framework (Angrist et al., 2025) shows that, while less cost-effective than small-scale pilots, *JEC* achieves meaningful learning improvements at scale, illustrating the potential of government-led bundled reforms in LMICs.

In the next section, we describe the program structure and implementation timetable. In Section 3, we outline our data sources. In Section 4, we explain the identification strategy and empirical methodologies, then present results and explore mechanisms of impact in Section 5, accompanied by an exhaustive set of robustness checks. We conclude in the final section.

2. A comprehensive educational reform in Peru: *Jornada Escolar Completa*

The Peruvian state education system comprises compulsory elementary school (grades 1-6) and secondary school (grades 7-11). The academic year for public schools runs from mid-March to mid-December.[9] Starting in the 2015 academic year, Peru implemented a comprehensive educational reform, the *Jornada Escolar Completa* (*JEC*), in 1,000 public secondary schools nationwide.

As part of the reform package, *JEC* schools added two instructional hours (45 minutes each) per day, thereby increasing the instructional week from 35 to 45 hours (26.2 to 33.8 regular hours), an increase of almost 30 percent, and in line with private schools. However, the program's goal was to improve not just the quantity of schooling hours but also the quality of public secondary education and to replicate some of the features observed in private schools. Relative to regular public secondary schools, *JEC* schools also benefited from: (i) a pedagogical component; (ii) improved school management through more and better organized school personnel; (iii) improved physical infrastructure; and (iv) improved IT support (Ministry of Education, 2014). The pedagogical component included a new pedagogical support program for teachers ("*Acompañamiento pedagógico*"), and the availability of psychologists to meet with students twice a year, as well as English for students, and skills development for the job market. Teachers' and principals' salaries increased to account for the additional hours,[10] and online support was offered to subject coordinators.[11]

As illustrated in Table 1, the regular system (column (1)) assigns four instructional hours to both math and reading each week. In *JEC* schools (column (2)), math instruction was extended by two hours (from 4 to 6), reading by one (from 4 to 5), English by three hours (from 2 to 5), and science by two (from 3 to 5). Also, the program paid special attention to tutoring students, especially those falling behind.[12] The *JEC* reform was designed to be implemented in public secondary schools with only a morning shift (so that the extension in school hours would not affect afternoon shift students) and to be large enough: have eight or more sections and available space to accommodate, for example, a laboratory and a library. The Ministry of Education (MINEDU) used data from the 2013 *Censo Escolar* (school census) to identify schools that met all these requirements and found 1,360 nationwide, and added 52 schools recognized as "emblematic".[13]

The list containing 1,412 schools was then sent to local coordinators for validation. This process added and removed some schools, resulting in a total of 1,343 schools. MINEDU then hired evaluators to gather further information about these schools and selected 1,000 of them. This list was included in the directive creating the program in September of 2014 (RM Nº 451-2014-MINEDU). The list was amended one more time in February 2015, replacing six schools from the original list (RM Nº 062-2015-MINEDU). Appendix Table A1 presents the full selection process and timeline.

Note that the process of selecting the final 1,000 schools from the original 1,412 was likely driven by unobservable factors, possibly reflecting the bargaining between the central administration and the local coordinators as well as the school districts. Thus, to causally identify the impact of the program, we avoid comparing the left-out schools with the final list. Rather, our identification strategy is based on the first set of eligibility rules, which use clear guidelines derived from observables, i.e., public schools with eight or more sections (and the morning shift only). The distribution of the number of sections across public secondary schools with a morning shift only is reported in Figure 1 (full distribution in Figure A1 in the Appendix).[14] Most eligible schools had five sections, one per grade. Larger schools had two or more sections per grade, making 10 or 15 sections the most common among larger schools. However, the exact number of sections observed in any given year depends on other factors, including local demand.[15] Thus, the choice of eight sections as a key eligibility criterion for *JEC* was quite arbitrary, reflecting the budgetary limits of *JEC* in its first year. As further discussed in Section 4, our main source of identification comes from comparing schools around this threshold, having verified that the probability of participating in *JEC* changes discontinuously at eight sections. Unsurprisingly, we find low density at 7/8 sections, supporting our identification assumption of smoothness around the threshold. Recall also that the selection of schools into *JEC* in 2015 was made in 2014 using information from the 2013 school census, which makes any sort of manipulation around the '7/8' threshold infeasible.

The *JEC* reform continued beyond 2015, with more schools incorporated in 2016 and 2017. In each of those subsequent years, a different eligibility rule was used. In this study, given our identification strategy, we focus only on those schools that benefited from the reform in 2015. The data sources and the methodology to identify causal effects are described in the next section.

3. Data

Central to our analysis are two administrative datasets: at the student level, the 2015 *Evaluación Censal de Estudiantes de Secundaria* (*ECE-S*), our primary source of data for measuring academic achievement; and, at the school level, the 2015-16 *Semáforo Escuela*, a large school survey monitoring educational services provided by schools and a useful source of information to explore mechanisms at the school level. More details on both datasets are provided below.

ECE-S is the nationally standardized census assessment of math and reading, administered to all eighth graders in public and private schools nationwide.[16] The 2015 *ECE-S*, the first of its kind, covers 94.4 percent of all students attending 99.5 percent of all secondary schools in the country. Administered at the end of the school year (in 2015, on the 17th and 18th of November), it consists of 60 multiple-choice questions in math and reading. The z-score transformations of the math and reading tests are our primary variables of interest, calculated using the national mean and standard deviation, which includes both public and private schools.[17] Students are classified into four groups based on their test performance: at grade level, in process, at the beginner level, and below the beginner level. The best-performing students, "at grade level", are ready to face the challenges of the next grade. Nationwide, only 14.9 percent of students achieve this level for reading and 9.6 percent in math (Table A2 in the Appendix)—among public schools, these proportions reduce to 9.9 and 6.3 percent, respectively. Students in the second group, "in process", partially achieved their grade goals but retained knowledge from the previous grade (22.8 percent in reading, 12.8 percent in math). "Beginner-level" students (40 percent of 8th graders in both subjects) did not demonstrate mastery of the previous grade. Finally, 37.2 percent in math and 23.3 percent in reading are below the beginner level. This dismal performance is consistent with the poor scores Peruvian students have obtained in international tests such as PISA and the Third Regional Comparative and Explanatory Study (TERCE). Thus, in our analysis, we also explore whether *JEC* affected the distribution of students by estimating its impact on the probability of scoring at grade level. Besides students' assessment data, *ECE-S* includes a short questionnaire administered to students and including information about their demographic and socioeconomic characteristics (e.g., age, gender, parents' education, native language), and relevant to our study, their own perceptions about parental involvement (regarding, for example, homework and book recommendations) and about their abilities in math and reading. We use these variables to capture possible mechanisms.[18]

Semáforo Escuela is a large-scale school survey used by the MINEDU to monitor the delivery of educational services in public schools.[19] Data are collected monthly from a representative sample on various topics, including characteristics of administrative staff and teachers, access to support for teachers, other government-run programs, access to IT, teachers' level of specialization, and more. Information is provided by the principal (school module) and by up to three randomly selected teachers per school (teachers module). To maximize sample size, we aggregate data collected in 2015 and 2016 (if a school is visited twice, we choose the first observation); thus, we denote the dataset as 2015-16 *Semáforo Escuela*.

We also use data from the 2013 *Censo Escolar* (school census) to determine the number of sections, along with additional information from the schools (used to run the smoothness tests and/or as control variables in our regressions). All these data sources come from MINEDU, which also provided the list of schools that became *JEC* in 2015. A key advantage of *ECE-S* and *Semáforo Escuela* is their large number of observations, which allows us to implement several data-intensive strategies to measure and check the robustness of the impact of *JEC*.

We supplement these with analysis of data from the 2016 *Encuesta Nacional a Docentes* (2016 *ENDO*) and the 2016 round of the Young Lives (YL) longitudinal study. *ENDO* is a biannual survey administered to a representative sample of teachers (up to three per school) from public and private schools (Ministry of Education, 2016). The survey collects information on teachers' professional trajectories, income, time use (dedicated to school activities and other income activities), access to training and IT, attitudes, and motivation, among others. The YL dataset includes information on a national sample of adolescents, born in 2001-2. A detailed description of the YL sampling design can be found in Escobal and Flores (2008).[20] The fifth round of the YL study took place in 2016, a year after the *JEC* reform started. The reform benefited a sizable proportion of the cohort, aged 14-15 at the time of data collection. YL contains information on time use (during a typical school day), socio-emotional competencies (measurements of self-concepts and aspirations), and technical skills (self-reported knowledge of English and digital skills).[21] Time use covers activities such as sleeping, household and caregiving tasks, paid work, time in school, studying outside school, and leisure. Notably, YL also administered its own math and reading tests to all respondents at home, regardless of whether they were enrolled in school. The YL data enable us to extend the analysis of the *JEC* program's effects on a set of child-level outcomes and to investigate mechanisms explaining the reform's impact on academic achievement, alongside school- and teacher-level mechanisms explored through *Semáforo Escuela* and *ENDO* data.

We match each of our four datasets to the 2013 school census to replicate the *JEC* eligibility rule and test the smoothness assumptions, thereby validating our identification strategy.[22] The school census is an annual administrative dataset that records school-level characteristics, including infrastructure, personnel, enrolment, pass rates, and notably, the number of sections (Ministry of Education, 2013). Table A3 in the Appendix provides details on the data sources, including variables, the unit of observation, and the year of data collection.

4. Methodology

Based on the selection rule of the *JEC* program, we estimate its impact using two related methodologies: global polynomial regressions and local randomization.

4.1. Global polynomial regression approach

As described above, having eight or more sections qualifies for *JEC* eligibility but does not fully explain participation, suggesting a discontinuous increase in the probability of becoming a *JEC* school. Thus, we use the discontinuity at eight sections as the instrument for participating in *JEC* in a conventional two-stage least squares (2SLS) model. This is shown by Equation (1) below:

$$JEC_{ij} = \alpha_1 + \pi\mathbb{1}(S_j \geq 8) + h_{11}(S_j \geq 8) + h_{12}(S_j < 8) + \theta X_{ij} + e_{ij} \qquad (1)$$

where JEC_{ij} takes the value of one for if a student (i) is attending a *JEC* school indexed by j and zero otherwise. The indicator function $\mathbb{1}(.)$ returns a one only for schools with eight or more sections ($S_j \geq 8$). Functions $h_1(.)$ and $h_2(.)$ are flexible polynomials in the discrete assignment (running) variable S (number of sections), X is a vector of students' characteristics (age, sex, mother's language, and school attainment) and schools' characteristics (whether the school is located in an urban/rural area, whether the school has a number of sections in a multiple of five, and school district fixed effects), and e_{ij} is the error term. The second stage, estimating the impact of *JEC* (as predicted by the discontinuity) on outcome Y_{ij} is given by Equation (2) and captured by parameter β:

$$Y_{ij} = \alpha_2 + \beta\widehat{JEC}_{ij} + h_{21}(S \geq 8) + h_{22}(S < 8) + \lambda X_{ij} + u_{ij} \qquad (2)$$

where, and u_{ij} is the error term. Our two main outcomes (Y_{ij}) are performance in reading and math. Additional outcomes analyzed to understand potential mechanisms are presented later. Standard errors are heteroskedasticity-robust and clustered at each of the local school districts (known as UGEL). Throughout the different datasets (*ECE-S*, *Semáforo Escuela*, *ENDO*, YL), we keep controls constant when possible.[23]

For the instrument to be valid, the discontinuity at eight sections should strongly predict participation in *JEC*. As discussed in section 2, the Ministry used 2013 data to identify schools with eight or more sections. In Figure 2, we present visual evidence to support this rule. We focus exclusively on public schools with only a morning shift, as only these schools were eligible for *JEC*. As expected, the probability that a public secondary school is part of *JEC* is zero for those with fewer than seven sections and near zero for those with seven sections. At eight sections, the probability discontinuously jumps to nearly 0.60 and remains high before decreasing for very large schools.[24] This feature implies a fuzzy discontinuity, as the probability of being part of *JEC* is less than certain at the threshold, and justifies an instrumental variables approach. The regression counterpart of this evidence is presented in Table 2. In column (1), with linear splines for the running variable, we find that at the threshold of eight sections, a school is 59 percentage points (p.p.) more likely to be part of *JEC*. Using quadratic splines (column (2)) does not change our conclusions. The rule of selecting schools based on having eight sections or more is a strong predictor of *JEC* participation and provides a discontinuous jump that can be used as an instrument, as long as the exclusion restriction is satisfied.

To satisfy the exclusion restriction, the instrument should affect the outcomes only through its impact on participation in *JEC*. Thus, for all other variables, especially for those measured in 2013, there should be no discontinuity at the threshold. This evidence is presented graphically in Figures A2 and A3, and formally tested in Tables A4 and A5, and shows a smooth transition around the threshold for a set of school-level and student-level characteristics observed using 2013 data from the school census, two years prior to the intervention.[25] Furthermore, taking advantage of the longitudinal nature of the Young Lives data, we run placebo tests using outcomes measured two years before the *JEC* reform started –for reading and math in Table A6, and for time use in Table A7 , covering schooling, study, work, household, and leisure activities during a typical school day; in all cases, the placebo tests show no significant effects.

As a final check, following best practice in the literature (see, for instance, What Works Clearinghouse, 2022), we note that our setup satisfies the basic requirements for the regression discontinuity design to be applicable. First, due to the institutional setup, it was not possible for schools to manipulate the running variable. Second, for our main results (math and reading), we use census data, which avoids any potential attrition problems associated with the *JEC* reform.[26] Third, while the running variable has a certain amount of heaping due to the nature of the Peruvian education system (schools with sections in multiples of five are more common), we control for this in our main specification and discuss potential implications for external validity. Fourth, we report results considering different functional forms (2SLS with linear and quadratic splines) and bandwidths (as discussed in Section 5.2). Fifth, as previously explained, our strategy fulfils the basic criteria for a fuzzy regression discontinuity design.

Local polynomial regressions are commonly used in regression discontinuity designs to focus on observations near the threshold, where causal impact is identified. A key challenge arises when the running variable cannot be assumed to be continuous, as is the case with the number of sections variable. When the running variable is discrete, the researcher might be forced to choose a window that is too wide for the approximation bias of the average treatment effect (ATE) estimator to be negligible, leading to unreliable inference using standard confidence intervals (Lee and Card, 2008). Our key strategy to deal with the discrete nature of the running variable is described in section 4.2 (local randomization approach). However, as part of the robustness checks, we report local polynomial regression results using the bias-aware Bounded Second Derivative (BSD) methodology proposed by Kolesár and Rothe (2018). This approach allows the construction of "honest" confidence intervals that explicitly account for potential approximation bias and exhibit good coverage properties when the running variable is discrete. Implementing this method requires placing bounds on the curvature of the outcome–running variable relationship (the second derivative), captured by a parameter denoted as κ. Following Kolesár and Rothe (2018), we report results for different values of κ, with larger values corresponding to more conservative assumptions and wider confidence intervals. For each value of κ and outcome, the procedure determines an optimal bandwidth. Accordingly, we report multiple estimates for each outcome of interest.

4.1.1 Isolating the impact of increasing instructional time

The *JEC* reform expanded the time for math by two hours (90 minutes) compared to only one additional instructional hour (45 minutes) for reading, whereas in regular schools, both courses have the same instructional time. To isolate the impact of the additional instructional time for math, we use a *difference-in-discontinuity* (diff-in-disc) approach. This approach allows us to estimate the impact of the additional 45 minutes for math by exploiting within-student variation to estimate the impact on test scores based on the variation in instructional time by subject, while continuing to use the discontinuity at eight sections as before. The (reduced form) specification is as follows:

$$Y_{ijs} = \alpha + \beta_1 \mathbb{1}(S_j \geq 8) + h_1(S \geq 8) + h_2(S < 8) + \beta_2 Math + \beta_3 Math\, \mathbb{1}(S_j \geq 8) + h_3 Math\, (S \geq 8) + h_4 Math\, (S < 8) + \lambda_i + v_{ijs} \qquad (3)$$

where Y_{ijs} is the score of child *i* from school *j* in subject *s*, *Math* is a dummy variable that takes the value of 1 for the math test score, and 0 for the reading test score, λ_i are individual fixed effects, and v_{ijs} is the error term. The parameter of interest is β_3. We assume that in the absence of the extra instructional hour for math, the impact of *JEC* would have been the same on both subjects, this is equivalent to the common trends assumption. This assumption has three components. First, the effect of instructional time is the same across math and reading. This is the assumption used in the literature to estimate the impact of instructional time exploiting differences across subjects and using student fixed effects—see Lavy (2015, 2020), and Rivkin and Schiman (2015). Second, the effect of the other features of *JEC* is also the same for both subjects. Both assumptions imply a similar technology for math and reading education. Third, there are no spillovers in instructional time (i.e., an extra hour in reading does not improve math test scores), and if there are, the effect found is net of these spillovers.

4.2. Local randomization approach:

The local randomization approach formalizes the idea that the regression discontinuity design behaves like a randomized experiment near the cut-off and imposes randomization-type assumptions that can be adapted to the sharp and fuzzy cases (Cattaneo, Idrobo and Titiunik, 2018).

The standard continuity-based regression discontinuity design used in local polynomial methods assumes that the running variable is a continuous random variable (i.e., that the number of mass points is large), which is hard to justify in our case (Cattaneo, Idrobo and Titiunik, 2018). However, when the running variable is discrete, the local randomization approach has the advantage that the window selection procedure is no longer needed, as the smallest window is well defined (between 7 and 8 sections in our case), and can be used appropriately (Cattaneo, Idrobo and Titiunik, 2018). Put differently, if local randomization holds, then it must hold for the smallest window. We therefore report additional results using the local randomization approach for the fuzzy regression discontinuity case. In Table 2, we report results for the first stage using this approach with *ECE-S* data (column (3)). The results are virtually identical to those obtained using conventional 2SLS.

The discrete nature of the assignment variable requires that observations around the threshold of eight sections behave as a binomial distribution with a probability of success equal to 0.5, i.e., as a "coin flip" (Cattaneo, Idrobo and Titiunik, 2018). We report this test in Table A8 in the Appendix using *ECE-S* data, restricting the sample to schools with either seven or eight sections. The estimated probability is 0.428. There are 267 schools with 7 sections, 200 with 8. This is statistically different from 0.5, though note that there are more schools in the non-eligible category, which does not suggest manipulation towards receiving the program. Our sample includes schools in both urban and rural areas, as JEC schools are present in both. Whilst being located in an urban area is not a selection criterion, we did note that the majority of JEC schools are in urban areas. When restricting the analysis to urban schools, the probability is much closer to 0.5, and in that case, we cannot reject the null hypothesis that each is equal to 0.5. This further reinforces the validity of our analysis. We use all public schools with a morning shift for our estimations (and control for whether the school is located in an urban/rural area), in order to mirror as closely as possible the official selection criteria, but provide a robustness check including only urban schools in Table A9 in the Appendix, which shows similar results (discussed below).

5. Results

In Sections 5.1 to 5.3, we estimate our main results and robustness checks using data from 2015 *ECE-S* and 2015-16 *Semáforo Escuela*, respectively. For these two datasets, we report conventional 2SLS results with linear and quadratic splines only (Gelman and Imbens, 2019), and as a robustness check, we estimate local polynomial regressions using the BSD method (Kolesár and Rothe, 2018). In all cases, we focus on eligible schools only (i.e., public schools with the morning shift only). Considering eligible schools, *ECE-S* and *Semáforo Escuela* report information for 6,121 and 3,288 schools, respectively—the former virtually represents the universe of eligible schools. For the main outcomes (math and reading), we report additional robustness checks by: running a bandwidth sensitivity analysis; collapsing the data at the school level; comparing eligible schools (morning shift only) with ineligible schools; running placebo tests with private schools; and, replicating our identification strategy with the YL data. Given the discrete nature of the running variable, for *ECE-S* and *Semáforo Escuela,* we also report results using the local randomization approach, and consider them qualitatively as important as those obtained with conventional 2SLS. When data collected in 2016 is used (*Semáforo Escuela*), students attending schools that became *JEC* in 2016 are dropped from the sample.

Todd and Wolpin (2003) show that to understand the full effect of education policies, the behavioral changes of all actors involved (students, teachers, and parents) should be incorporated as inputs, as they could result from the policy change itself.[27] In sections 5.4 to 5.6, we use data from 2016 YL, 2016 *ENDO*, and a self-reported module of the 2015 *ECE-S* to incorporate this. For YL and *ENDO*, we refrain from reporting local randomization results because the effective sample sizes are small (information comes from students and teachers observed in 186 and 404 eligible schools, respectively), and there are few schools near the threshold (see Figure A4).[28] For similar reasons, for YL and *ENDO* it is not possible to assess the role of higher polynomials, as there are too few observations to the left of the running variable distribution in these two datasets.[29] However, for these datasets, we report results using the BSD method for outcomes that are statistically significant when using 2SLS (with linear splines).

When using data from *Semáforo Escuela*, *ENDO*, YL, and the *ECE-S* self-reported module, multiple outcomes are observed. When relevant, we create indices to summarize our results as in Kling et al (2007).[30] As for the multiple outcomes, adjusted p-values are corrected for multiple hypotheses based on Benjamini and Hochberg (1995), and we apply the concept of a false discovery rate to allow inference when conducting many tests.[31]

We first focus on results for the main student administrative dataset, the *ECE-S*, which contains information for reading and math. Performance in these subjects is reported as (standardized) test scores and as the probability that a student scores in the highest group (=1 if performed at the expected level for the grade). The latter is defined on criteria established by MINEDU (see Section 3).

Figure 2 shows the reduced-form graphs for the impact of *JEC* on reading and math test scores. A discontinuity is observed for both math and reading outcomes. These findings, which are only suggestive, are formally analyzed in Table 3, which presents the main results for academic achievement. In column (1), we report results for the 2SLS using *ECE-S* data, with linear splines. These results show that *JEC* increased test scores in math and reading by 0.185 and 0.103SD, respectively. Similarly, *JEC* increased the probability of students scoring at grade level in math and reading by 4.2 and 2.9 p.p., respectively. These are relatively large improvements as only 5.8 percent and 8 percent of students in eligible schools perform at grade-level, representing percentage increases of 72.4% and 36.2%, respectively. In column (2), we consider quadratic splines as an alternative specification. When doing this, our conclusions for math remain unaffected; we estimate an effect of *JEC* on math test scores of 0.137SD, and an increase in the probability of being at grade level by 2.5 p.p., which represents a percentage increase of 43%. When using quadratic splines, the coefficients associated with reading decrease in magnitude and become statistically insignificant.

As anticipated in Section 4, since the running variable is discrete, an alternative framework to study the impact of *JEC* is a local randomization approach. The identifying assumption is that in the vicinity of the cutoff, assignment is as good as random (Cattaneo, Idrobo and Titiunik, 2018). Column (3) in Table 3 shows estimates of the impact of *JEC* using the local randomization approach for a fuzzy regression discontinuity with the smallest possible bandwidth: 7 and 8 sections. These new set of results confirm our previous findings: a strong effect for math scores and a weaker impact (but still statistically significant, in this case) for reading—effect of *JEC* on test scores of 0.212 and 0.142 SD (respectively), an increase in the probability of being at grade level by 3.1 and 1.3 p.p. Similarly, when comparing the cumulative distribution functions of the test scores for math and reading of schools with seven sections against those in eight sections (Figure A5 in the Appendix), results also show a clear improvement for math but not necessarily for reading. Previously, we noted that the binomial test for the local randomization holds specifically in urban areas. In Table A9 in the Appendix, we note that our results for local randomization are similar when restricting the sample to eligible schools in urban areas.

The larger (and more robust) effect found for math test score (relative to reading test scores) is confirmed with additional robustness checks reported in the next sub section, and it is consistent with the fact that in the *JEC* program, students received two extra instructional hours for math, and only one extra for reading (the marginal impact of the extra instructional hour for math is formally analyzed in section 5.3). Compared to the broader literature on the impact of school extended-day reforms, our lowest estimate for math is bigger than the effects reported by Bellei (2009) for Chilean secondary schools (0-0.12 SD in math) and by more recent papers using data from PISA who tend to find an impact <0.04 SD for developing countries (Lavy, 2015; Rivkin and Schiman, 2015).[32]

In terms of generalizability, our result is mainly identified among students near the threshold. The main difference observed between the compliers at the threshold with the average student (full sample) is that the former are more likely to study in urban areas (Table A10 in the Appendix). This result is not necessarily surprising, as *JEC* schools are mainly located in urban areas—most rural schools are smaller and thus tend to have 5 or fewer sections. We reflect on the implications of this finding in our conclusions.

5.1.1 Cost-benefit analysis

To contextualize our results, we calculate the LAYS (Angrist et al., 2025) defined as $\frac{\beta_{JEC}}{\delta_b}t$, where β_{JEC} is the impact of JEC on learning, δ_b represents how much a student learns in a year of schooling in a high-quality environment—set to 0.8 SD following Angrist et al. (2025)—and t is the number of years the impact is expected to last. We define β_{JEC} as the simple average of the impact of *JEC* on math and reading in the main specification and set $t = 1$ because the impact is measured after one year. With these assumptions, we estimate an impact of 0.18 LAYS. Considering a list of 200 interventions from 52 countries reviewed by Angrist et al, this places *JEC* in the upper half of the distribution of effective interventions in terms of LAYS, more effective than the typical effect size for general teacher training, cash transfers, and input-only interventions (textbooks, class size, laptops/tablets, grants, libraries, etc.) but below the typical top-end effects seen in pilot studies of targeted tutoring, teaching at the right level, or teacher accountability and incentives (e.g., merit based pay).

While the intervention is reasonably effective, it is relatively costly. According to the MINEDU, in 2015, the *JEC* reform had a total cost of 2015 USD 571 per student per year,[33] which means the intervention has an impact of 0.032 LAYS per USD 100 spent. This places *JEC* in the lower tail of the distribution of cost-effective interventions in the Angrist et al. review, above the impact of a cash conditional transfer and with a magnitude similar to that of a program that provides full scholarships for secondary-level students (Duflo et al. 2021). However, it is important to highlight that our results are likely to underestimate the long-term impact of *JEC*–in our calculations, we assume the impacts will last just one year. Furthermore, this analysis only considers the sub-sample of interventions reported in Angrist et al that have cost data and significant, positive impacts.

5.2 Further robustness checks

In this section, we present additional robustness checks that highlight the strength of the results linked to math. First, we explore an alternative way to estimate the impact of JEC using nonparametric methods that allow us to choose an optimal bandwidth.[34] As described in Section 4.1, we use the BSD approach proposed by Kolesár and Rothe (2018). To apply this strategy, it is necessary to make assumptions about the value of the second derivative of the relationship between the outcome variable and the running variable (denoted as κ). Considering a number of scenarios similar to those used by Kolesár and Rothe,[35] the most common optimal bandwidth (w) observed across our four outcomes is two (i.e., schools with 6, 7, 8 and 9 sections). In Appendix Table A11, we report results fixing the bandwidth to w=2 for different scenarios of κ. Then, in Appendix Table A12, results are reported allowing the algorithm to determine the optimal bandwidth in each case for different scenarios of κ. Keeping the bandwidth fixed (Table A11), results confirm our previous findings. We obtain an effect of *JEC* on math and reading test scores of 0.167 and 0.062 SD (respectively), and on the probability of being at grade level on each of these subjects of 2.6 and 0.6 p.p. (respectively). The coefficients linked to math are statistically significant in all but one scenario—the most conservative, which assumes a very large second derivative. In contrast, the coefficients for reading are statistically insignificant across all κ scenarios. This confirms that the math results are the most robust. Results from Table A12 in the Appendix (optimal bandwidth) show similar evidence, with results for math being statistically significant in all but one scenario, whereas results for reading tend to be statistically insignificant, with estimated parameters for reading having the opposite sign in some scenarios of κ.

Second, we replicate the results presented in Table 3 by collapsing the data at the school level for the analogous group of secondary schools (public, with only the morning shift). These results are reported in Table A13 in the Appendix. Using the collapsed sample, the results for math test scores with linear splines (column (1)) remain very similar, capturing an effect of 0.178 SD and an increase of 3.1 p.p. in the probability of performing at grade level. The point estimates for reading decrease in magnitude, and the coefficients associated with the effect on reading test scores lose statistical significance. Using quadratic splines, we find no effect of *JEC* on math in this specification. However, using the local randomization approach (column (3)), we still find a significant impact of *JEC* on both math and reading (with an increase in test scores equivalent to 0.344 and 0.274 SD, respectively).

We conduct a series of placebo tests as robustness checks. First, we consider public schools with shifts other than morning-only, which are not eligible for *JEC*. For them, there should be no discontinuity at the eight-section threshold. Appendix Figure A7 validates that conjecture. Thus, we can use a difference-in-discontinuity approach and compare the discontinuity at the threshold in eligible and ineligible schools. These results are reported in Table A14 in the Appendix. In this case, the impact of *JEC* on math test scores using linear splines (column (1)), is 0.244 SD. The analogous point estimate using quadratic splines is 0.231 SD (column (2)). Similar results are obtained using the local randomization approach. Second, we consider an alternative group of ineligible schools: private secondary schools with a morning shift. As reported in Appendix Figure A8, there are no discontinuities in test scores at the eight-section threshold for these schools. If anything, there appears to be a reduction in test scores. Appendix Table A15 shows the regression counterpart, using a reduced form specification to link the relationship between the eight-section threshold and test scores. The reduced form results show contradictory results depending on the order of the polynomial used, whereas the results using the local randomization approach (sharp) show no evidence of a discontinuity around the threshold for math scores in those schools.

Finally, we use YL data to validate the main results. Although the YL sample is not fully nationally representative, the fact that data were collected before and after the *JEC* reform began allows us to replicate the ECE results on academic achievement while also implementing additional robustness tests. In 2013 and 2016, YL administered a reading comprehension and a math test designed by members of the YL team (Cueto et al., 2009; Cueto and León, 2012). As described earlier and as expected, *JEC* does not predict changes in test scores in the pre-*JEC* period (Table A6, column (1)). However, once the reform was implemented, YL provides very similar estimates of the *JEC* impact on math test scores to ECE (by 0.287 SD) (Table A6, column (2)). The coefficient for reading, while large in magnitude, is lower than that for math and is statistically insignificant. Taken together, these results indicate a strong and robust effect of *JEC* on math test scores, with reading effects that are robust across many of this extensive set of checks. In the next section, we discuss the mechanisms that might explain these results.

5.3 Mechanisms: the impact of JEC on instructional time and school resources

Isolating the effect of instructional time: We proceed to analyze the potential mechanisms underlying the improvement in learning outcomes, considering school-level factors that were expected to change as a result of the reform. First, we use *ECE-S* data to isolate the impact of additional instructional time in math to determine whether it can explain the larger effect in this subject. To assess the impact of extending the school day *net* of all the other components of *JEC*, we take advantage of the fact that in non-*JEC* schools, math and reading had the same number of instructional hours (four), but *JEC* expanded the instructional time for math by two hours compared to only one additional instructional hour for reading (Table 1). As far as we are aware, there is no other component of the reform that differs between math and reading other than this one. This allows us to use a *difference-in-discontinuity* (diff-in-disc) approach described in section 5.2 to estimate the impact of the additional instructional time for math by exploiting within-student variation, while continuing to use the discontinuity at eight sections as before. Table 4 presents the results of this analysis, focusing on the reduced-form specification. In this case, the parameter of interest is the interaction *Math*1*(*Section≥8*) after controlling for the subject and splines of the running variable for each subject and the student fixed effects. We find that the additional instructional hour in math in *JEC* schools increases test scores by around 0.036 (linear splines) and 0.072 SD (quadratic splines). This is a relatively large effect. Compared to the reduced forms results shown at the bottom of the table, for the specification with linear splines, the effect of one additional hour would explain around 32% (=0.036/0.111) of the overall gains in math. Since instructional time in math increased by two hours in *JEC* schools, the total impact is around 64% (=2*0.036/0.111). This analysis suggests that a large part of the improvement in test scores from *JEC* is due solely to the increase in instructional time, under the strict assumptions described above. However, this result also suggests that the other features of *JEC* are important for explaining the overall effect of the reform on math, as, assuming the effect of instructional time is linear, one-third of the overall effect remains unexplained.

The role of school resources: We use data from *Semáforo Escuela* to document changes in school resources that should have occurred as part of the *JEC* reform. *Semáforo Escuela* collected data in 2015-16. In 2016, additional schools untouched by the reform in 2015 were incorporated into the program, and as a result, schools with fewer than 8 sections were now incorporated. For this reason, schools that became *JEC* in 2016 are dropped from the sample. We note that our conclusions derived from the analysis of the 2015 *ECE-S* remain unaltered when excluding the

2016 *JEC* schools from the *ECE-S* sample—see Appendix Table A16 that fully replicates our main results from Table 3 using conventional 2SLS and a local randomization approach.

As part of its design, *JEC* was expected to increase the availability of school resources: IT infrastructure, staff, and pedagogical resources for both teachers and students. These changes are reflected in the results presented in Table 5. Focusing first on the specification with linear splines (column (1)), improvements in these areas are observed by 0.725, 3.004, and 1.442SD, respectively. The increase in IT infrastructure (Panel A) is explained by more classrooms with computers and laptops, as well as access to Kit *Robotica*, a set of materials used to build and program robots, intended for educational purposes. These are statistically significant, accounting for multiple hypothesis testing in the p-values. In relation to the school staff available (Panel B), the probability of having a complete teaching staff decreases. This could be an unintended consequence of the *JEC* school program, as it imposes an additional burden on school personnel, including teachers. However, this decrease in teaching staff is compensated by an increase in the number of non-teaching staff available, such as security guards and maintenance staff.

The increase in the availability of pedagogical resources (Panel C), arguably one of the most important features of the program together with the expansion of the school day, is explained by a significant increase in the likelihood of having a psychologist at the school, having teachers that provide support to parents, and the school participation in the MINEDU program "*Acompañamiento pedagógico*". These results remain statistically significant after adjusting for multiple hypothesis testing. "*Acompañamiento pedagógico*" is a component of the *JEC* program, though not unique to it. Schools that receive this program are visited by specialists who work with teachers to improve their pedagogical strategies. In turn, psychologists work both with teachers and (directly) with students. These additional resources explain not only the positive impact of

JEC on learning, but also why this intervention may have a greater impact (at least in the short run) than other expansions in the length of the school day observed in developing countries. This speaks to the importance of the complementarity of school inputs as discussed by Mbiti et al (2019) and also echoes the review findings of Ganimian and Murnane (2016), who argue that increasing time at school is only effective when the student experience improves.

In Table 5, we report additional results using quadratic splines for the 2SLS specification (column (2)), and the local randomization approach (column (3)). Across all specifications, our results confirm improvements in school staff and pedagogical resources, particularly among non-teaching staff, and the probability of having "*Acompañamiento pedagógico*" and a psychologist. The results for the IT index appear less robust (in the specification with quadratic splines); however, the improvement in the number of classrooms with laptops is robust across all specifications. In addition, in Table A17 in the Appendix, we report estimates using the BSD method (for multiple values of κ) for the three main indexes. These results confirm the importance of IT, staff, and pedagogical resources.[36]

5.4 Other mechanisms: the impact of JEC on students

We use YL data to examine the effects of *JEC* on students' time use and skills. Students might choose to exert less effort (especially outside school), which could make the net impact on time dedicated to learning activities ambiguous. Although pure effort is unobservable, using YL data, we confirm that *JEC* has a positive overall impact on students' time dedicated to studying. In Table 6, we find that *JEC* increases time at school by 2.1 hours per day, which is slightly larger than the 'official' increase in the length of the school day at *JEC* schools. This increase comes with a reduction in time spent sleeping, doing household chores and other domestic tasks, studying at home, and leisure time; all results except those for leisure time remain robust to multiple hypothesis testing. The reduction in time spent studying at home (approximately 30 minutes per day) represents only a fraction of the realignment in how students spent time outside school and confirms a net increase in total time spent studying of approximately 90 minutes. These results suggest that the gains in test scores come, in part, from increased time spent studying, at the expense of time allocated to home production and leisure. In Table A18 in the Appendix, we show that the results linked to time use measured with the YL data are robust when using the BSD method.

We also look at the potential impact of *JEC* on aspirations, socio-emotional, and technical skills using YL data (Table 7). Socio-emotional skills might improve due to cross-productivity with cognitive skills, as well as through the increased number of psychologists available at the school (who are meant to interact directly with students), while technical skills might improve through greater access to IT resources.[37] Aspirations for higher education might improve through several channels, including improved academic performance and increased self-efficacy. For higher education aspirations, we distinguish between university and technical education. For socio-emotional skills, we consider self-esteem and self-efficacy as outcomes, which have been found to predict life outcomes.[38] For technical skills, we consider the ability to speak English and three scales developed by the YL team to measure digital skills (Cueto et al., 2018).[39] In all cases, information is self-reported. Our 2SLS results suggest a positive impact of the program on technical skills, summarized in our constructed index with effects of 0.131 SD, 0.213 SD and 0.407 SD, respectively (Table 7). However, only the technical skills result is statistically significant, driven by increases in self-reported English knowledge and computer skills (by 0.177 SD and 0.237 SD, respectively). Among these, the effect on English knowledge remains robust to multiple hypothesis adjustment and continues to be statistically significant when using the BSD method (Table A18 in the Appendix).

We also investigate potential changes in students' self-perceptions (about their ability to perform well in math and reading) in response to *JEC*. For this, we go back to the *ECE-S* data. The 2SLS estimates reported in Table 8 confirm negative responses in virtually all dimensions of reading (Panel A). In contrast, for math (Panel B), such negative effects are not found. Students, however, are less likely to help their peers, suggesting a more competitive attitude. Again, the differential impact of *JEC* on math and reading could be explained by how the reform has altered students' perceptions, as we observe that students in reading are more likely to feel marginalized and have lower self-perception.

5.5. Other mechanisms: the impact of JEC on teachers

Using data from *ENDO*, the teacher-level survey, we explore possible changes to teachers' behaviors (time allocation, attitudes, and satisfaction) and training opportunities as well as on predetermined characteristics, which might be informative of teachers' sorting into (or out of) *JEC* schools. Results from this subsection must be interpreted with caution, as the sample size of *ENDO* is relatively small.

In Table 9, Panel A, we show that the total hours spent on all school activities have not changed. This would be consistent with a substitution effect away from in-class activities, such as preparing lessons, grading students' homework, talking with parents, and interacting with other teachers, due to the increase in instructional time. In Panel B, we detect some evidence of a decline in teachers' attitudes and satisfaction, particularly with his/her job (by 0.391 SD). It is important to note that teachers in JEC are paid more than those in regular schools. An increase in dissatisfaction post-reform, despite no increase in overall hours and the compensatory increase in earnings, could suggest potential overall dissatisfaction with the reform itself. These unintended consequences could attenuate the effects of *JEC* and potentially threaten sustainability. A caveat to keep in mind is that the impact of *JEC* on teacher satisfaction becomes marginally insignificant when using the BSD method (see Table A19 in the Appendix).

In Panel C, we find no evidence of changes in teachers' training opportunities. It is important to note that while *Semáforo Escuela* shows evidence that "*Acompañamiento pedagógico*" (support to teachers) has higher coverage in *JEC* schools, this result is not validated by the ENDO dataset, which directly asks teachers. The latter result must be interpreted with caution due to *ENDO*'s small sample size. If the result is taken at face value, it might suggest that not all teachers necessarily benefited from the improved pedagogical resources. Finally, in Panel D, we show that teachers' characteristics (age, sex, experience, credentials, and contract type) are similar across JEC and non-JEC schools, suggesting no sorting effects.

5.6 Other channels: parents

Finally, we explore potential changes in parental behaviors. First, we use data from *ECE-S* to explore whether, in the student's perception, *JEC* led to changes in parental involvement, as measured by: (i) whether students talk to parents about homework, (ii) parents help with homework, (iii) parents help explain topics, (iv) parents care about grades, and (v) parents recommend books.[40] Regression results (using conventional 2SLS) in Appendix Table A20 suggest a negative effect of *JEC* on whether parents explain topics, as well as a marginal decline in the probability that parents recommend books to their children (column (6)). If anything, the latter negative effect is consistent with *JEC*'s weaker impact on reading test scores. Both results remain statistically significant when adjusting by multiple hypothesis testing. We should note, however, that these results on reduced parental involvement are not robust to the application of the BSD method (see Table A21 in the Appendix).

Finally, we investigate potential sorting effects at the student level. More involved parents might choose to transfer their children to *JEC* schools, and these children are likely to have better learning outcomes, generating positive externalities for the rest of the students.[41] If that were the case, our estimates of *JEC*'s impact on learning outcomes could be biased upward. Padillo-Romo (2022) finds a significant impact of sorting on estimates of increased instructional time in Mexico. We do not find such evidence of students' sorting. We compare several features of students in *JEC* based on their prior locations. Whilst the main *ECE-S* dataset does not contain information on student movement between schools, we can, however, exploit the longitudinal nature of the YL data. This is shown in Table A22 in the Appendix. Specifically, we find that of 379 YL participants attending *JEC* schools in 2015, 327 were already attending these schools before the reform began ('stayers'), while 52 students moved from another school to a *JEC* school in 2015 ('movers'). Using pre-*JEC* outcomes (measured in 2013), we show that *JEC* did not attract better students. If anything, the table shows that for math and reading test scores as well as for socio-emotional skills (captured by indices of self-efficacy, self-esteem, agency and pride), movers were already at a disadvantage compared to stayers before the reform began. This could imply possible attenuation of the effects on learning outcomes if *JEC* schools attracted low-performing students.[42] To partially address this, in Table A6 (column (3)) of the Appendix, we report a value-added specification of the impact of *JEC* on math and reading using YL data. The value-added specification controls for performance in math and reading (respectively) as observed before the reform, in 2013. We find that the results are qualitatively similar to those originally found for YL (Table A6, column (2)).

6. Conclusions

This paper evaluates a policy that seeks to improve education quality and learning outcomes within the public school system in an LMIC by increasing the number of pedagogical hours and simultaneously increasing investments in school resources. The *JEC* program expanded the school day from 35 to 45 pedagogical hours a week in Peruvian public schools. We exploit an arbitrary rule used to select schools into the program to identify the effect on math, reading, and other outcomes.

We find that the *JEC* program leads to improved learning as measured by standardized test scores in math and reading. The effects are robust and larger for math. They are smaller, though still positive for reading, and are robust to most of our extensive robustness checks. We note that students in schools with exactly 8 sections are more likely to be urban, suggesting that our results may be less generalizable to rural areas. Exploring several other datasets allows us to investigate key mechanisms of impact, including changes in instructional time, improvements in school resources expected from the *JEC* reform design, and potential changes at the student, teacher, and parent levels. At the school level, we document gains from increased instructional time in math and improvements in school resources, especially in non-teaching staff and in pedagogical resources (including "*Acompañamiento pedagógico*"). At the student level, we find that students do not substitute their time away from studying; rather, they dedicate fewer hours to domestic activities such as housework and decrease their leisure time. We also document evidence of improvements in student self-reported knowledge of English. Despite these positive findings, all of which contribute to the impact of *JEC* on academic achievement, we also find evidence of two unintended consequences of the reform: lower parental involvement at home due to the reform, and a reduced self-perception of students' abilities in reading (though not in math).

Overall, our results suggest that targeted investment can significantly improve the quality of public education. However, the program cost is high, and experimenting with the program's components and how they complement each other (or not) may improve cost-effectiveness.

References

Agüero, J. M., and Beleche, T. (2013). Test-Mex: Estimating the effects of school year length on student performance in Mexico. *Journal of Development Economics*, *103*, 353-361.

Alcázar, L. (2016). *Evaluación del diseño y proceso de implementación del modelo de Jornada Escolar Completa (JEC) para educación secundaria a nivel nacional*. FORGE: Peru.

Angrist, N., Evans, D., Filmer, D., Glennerster, R., Rogers, H., Sabarwal, S. (2025). How to Improve Education Outcomes Most Efficiently? A Review of the Evidence using a unified Metric *Journal of Development Economics* 172 (January 2025, 103382)

Bellei, C. (2009). Does lengthening the school day increase students' academic achievement? Results from a natural experiment in Chile. *Economics of Education Review*, *28*(5), 629-640.

Benjamini, Y. & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B* (Methodological), 57(1), pp.289-300.

Cabrera-Hernández, F. (2020). Does Lengthening the School Day Increase School Value-Added? Evidence from a Mid-Income Country. *The Journal of Development Studies, 56(2)*, 314-335.

Cattaneo, M. D., Idrobo, N., & Titiunik, R. (2018). *A Practical Introduction to Regression Discontinuity Designs: Volume II*. Cambridge Elements: Quantitative and Computational Methods for Social Science, II, 113.

Cerdan-Infantes, P., and Vermeersch, C. (2007). More Time Is Better: An Evaluation of the Full-Time School Program in Uruguay. *World Bank Policy Research Working Paper* 4167.

Cilliers, J., Fleisch, B., Prinsloo, C., Taylor, S. (2020). How to improve teaching practice? An experimental comparison of centralized training and in-classroom coaching. Journal of Human Resources, *55*(3), 926-962.

Cueto, S. & León, J. (2012). Psychometric Characteristics of Cognitive Development and Achievement Instruments in Round 3 of Young Lives. Young Lives Technical Note 25.

Cueto, S., Felipe, C. & León, J. (2018). Digital Access, Use and Skills Across Four Countries: Construction of Scales and Preliminary Results from the Young Lives Round 5 Survey. Young Lives Technical Note 46.

Cueto, S., León, J., Guerrero, G. & Muñoz, I. (2009). Psychometric characteristics of cognitive development and achievement instruments in Round 2 of Young Lives. Young Lives Technical Note 15.

Das, J., Dercon, S., Habyarimana, J., Krishnan, P., Muralidharan, K., & Sundararaman, V. (2013). School inputs, household substitution, and test scores. *American Economic Journal: Applied Economics*, 5(2), 29-57.

Dominguez, P., & Ruffini, K. (2023). Long-Term Gains from Longer School Days. *Journal of Human Resources*, *58*(4), 1385-1427.

Duflo, E., Dupas, P., & Kremer, M. (2021). The Impact of Free Secondary Education. In NBER Working Paper Series (Issue 1254167).

Escobal, J., & Flores, E. (2008). An assessment of the Young Lives sampling approach in Peru. Technical Note 3. Oxford: Young Lives.

Evans, D., & Popova, A. (2015). What Really Works to Improve Learning in Developing Countries? Policy Research Working Paper 7203. World Bank: Washington DC.

Evans, D. K., & Yuan, F. (2022). How Big Are Effect Sizes in International Education Studies? *Educational Evaluation and Policy Analysis*, *44*(3).

Favara, M. & Sánchez, A. (2017). Psychosocial competencies and risky behaviors in Peru. *IZA Journal of Labor & Development, 6(3),* 1-40.

Favara, M., Lavado, P. & Sánchez, A. (2020). Understanding teenage fertility in Peru: An analysis using longitudinal data. *Review of Development Economics, 24(4),* 1217-1236.

Favara, M., Crivello, G., Penny, M., Porter, C., Revathi, E., Sánchez, A., Scott, D. Le Thuc, D., Woldehanna, T. & McKay, A. (2021). Cohort profile update: the Young Lives Study. *International Journal of Epidemiology*, 50(6), 1784-1785e.

Fryer, R. (2014). Injecting Charter School Best Practices into Traditional Public Schools: Evidence from Field Experiments. The Quarterly Journal of Economics, 129(3), 1355–1407, https://doi.org/10.1093/qje/qju011

Gelman, A. and Imbens, G. (2019) Why High-Order Polynomials Should Not Be Used in Regression Discontinuity Designs. *Journal of Business & Economic Statistics*, 37(3), 447-456.

Glewwe, P. and Muralidharan, K. (2016) Improving Education Outcomes in Developing Countries: Evidence, Knowledge Gaps, and Policy Implications. *Handbook of the Economics of Education* 5, pp. 653-743.

Hincapie, D. (2016). Do longer school days improve student achievement? Evidence from Colombia. IDB Working Paper Series No. IDB-WP-679.

Holland, P., Evans, D. and Alfaro, P. (2014). Extending the School Day in Latin America and the Caribbean: Evidence, Experience, Challenges. World Bank Policy Research Working Paper, 7309.

Kerwin, J. and R. L. Thornton; Making the Grade: The Sensitivity of Education Program Effectiveness to Input Choices and Outcome Measures. *The Review of Economics and Statistics* 2021; 103 (2): 251–264. doi: https://doi.org/10.1162/rest_a_00911

Kling, J. R., Liebman, J. B., & Katz, L. F. (2007). Experimental analysis of neighborhood effects. *Econometrica*, 75(1), 83-119.

Kolesár, Michal, and Christoph Rothe. 2018. "Inference in Regression Discontinuity Designs with a Discrete Running Variable." American Economic Review, 108 (8): 2277–2304.

Lavy, V. (2015). Do Differences in Schools' Instruction Time Explain International Achievement Gaps? Evidence from Developed and Developing Countries. *The Economic Journal*, 125(588), F397-F424.

Lavy, V. (2020). Expanding school resources and increasing time on task: Effects on students' academic and noncognitive outcomes. *Journal of the European Economic Association*, 18(1), 232-265.

Lee, D. S. and D. Card (2008): "Regression discontinuity inference with specification error," *Journal of Econometrics*, 142, 655–674.

Levin, H. M., & Tsang, M. C. (1987). The economics of student time. *Economics of education review*, *6*(4), 357-364.

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

Ministry of Education (2014) "Modelo de servicio educativo: jornada escolar completa para las instituciones educativas públicas del nivel de educación secundaria", Setiembre.

Ministry of Education (2013). Censo Educativo, 2013. *Peru, Ministry of Education*, Estadística de la Calidad Educativa, http://escale.minedu.gob.pe/uee (accessed on 6/15/2019)

Ministry of Education (2016). Encuesta nacional a docentes de instituciones educativas y privadas, 2016. *Peru, Ministry of Education*, Estadística de la Calidad Educativa, http://escale.minedu.gob.pe/uee (accessed on 6/15/2019).

Padilla-Romo, M. (2022). Full-time schools, policy-induced school switching, and academic performance. *Journal of Economic Behavior & Organization*, *196*, 79-103.

Pop-Eleches, C., and Urquiola, M. (2013). Going to a Better School: Effects and Behavioral Responses. *The American Economic Review*, *103*(4), 1289-1324.

Rivkin, S. G., and Schiman, J. C. (2015). Instruction time, classroom quality, and academic achievement. *The Economic Journal*, 125(588), F425-F448.

Saavedra, J., & Gutierrez, M. (2020). Peru: A wholesale reform fueled by an obsession with learning and equity. In *Audacious education purposes: How governments transform the goals of education systems* (pp. 153-180). Cham: Springer International Publishing.

Sánchez, A. & Singh, A. (2018). Accessing higher education in developing countries: Panel data analysis from India, Peru, and Vietnam. *World Development, 109,* 261-278.

Todd, P. E., and Wolpin, K. I. (2003). On the specification and estimation of the production function for cognitive achievement. *The Economic Journal*, *113*(485), F3-F33.

What Works Clearinghouse. (2022). What Works Clearinghouse procedures and standards handbook, version 5.0. U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance (NCEE).

Wössmann, L. (2003). Schooling resources, educational institutions and student performance: the international evidence. *Oxford Bulletin of Economics and Statistics*, 65(2), 117-170.

Yorke, L. & Ogando Portela, M.J. (2018). Psychosocial Scales in the Young Lives Round 4 Survey Selection, Adaptation and Validation. Young Lives Technical Note 45. Oxford: Young Lives.

[1] Cross-country studies tend to find no relationship between time spent in school and learning or labor market outcomes (Wössmann, 2003).
[2] Lavy (2020) is a notable exception for the case of Israel, who does find a significant and positive causal relationship between instruction time and test scores.
[3] In 2012, Peru was ranked bottom out of 65 countries in the PISA (Programme for International Student Assessment).
[4] The minimum number of sections is the number of grades in a school. Students within a grade may be split into multiple sections, equivalent to homerooms in the USA, or form classes in the UK.
[5] A further expansion of the *JEC* reform took place in 2016 and 2017, using other eligibility rules. For our main analysis, we use data from 2015—when data beyond 2015 is used, schools that became *JEC* after 2015 are excluded from the sample.
[6] When the running variable is discrete, it is not possible to focus only on observations arbitrarily close to the cutoff, which makes inference from local polynomial regressions more sensitive to functional form assumptions. Kolesár and Rothe (2018) propose confidence intervals that explicitly allow for this limitation. We use this approach as a robustness check for inference. See Section 4 for details.
[7] Note that we include all fixed costs, including e.g. computer labs and English Language labs which may have had greater impacts on subjects that are not measured in our dataset.
[8] This crowding-out behavior has been observed in other scenarios, for example, in response to attending better schools (Pop-Eleches and Urquiola, 2013) and receiving anticipated school grants (Das et al, 2013).
[9] Private schools have a longer calendar and tend to start two weeks earlier.
[10] In 2015, teachers in *JEC* schools were required to work 30 hours per week, compared to 24 hours per week in regular schools. Teachers were compensated for the additional hours, earning the same hourly wage regardless of school type (Alcázar, 2016). A teacher's hourly wage varies according to career performance under the Teacher Career Reform (Reforma Magisterial).
[11] JEC had no significant provisions of meals as part of its design. Peru's public schools do not provide meals to students, except for breakfast in extremely poor districts and only in pre-K to 6 grade schools as part of the *Qali Warma* program. Only secondary schools in indigenous communities in the Amazon are part of this program, which are not part of JEC.
[12] However, note that schools in both the regular and the new system have unassigned hours (6 and 5 per week, respectively). In principle, nothing stops principals in schools that did not participate in *JEC* to allocate those hours to math, reading or any other subject. If that were the case and assuming *JEC* schools do not use unassigned hours in a similar fashion, our results would represent a lower bound of the impact of the increase in school hours, due to the possible "contamination" of the control group.
[13] These are large urban schools created in the 1950s and were labeled as emblematic in recent years. These schools tend to be very large (often with more than 30 sections) and have little impact on our analysis.
[14] In Figure 1, the distribution is truncated at 30 sections for illustrative purposes. Schools with 1 to 30 sections account for 99% of the total number of public schools with morning shift. The full distribution is reported in Figure A1 in the Online Appendix.
[15] A secondary school with eight sections would have at least one section per grade and three grades with two sections each. The assignment of students to sections is not regulated by the Ministry of Education (MINEDU) and varies by school. In some cases, it would reflect tracking of students but in others depends on alphabetical order or other rules. Relevant to our empirical strategy, we do not expect the assignment of students to sections to *discontinuously* change in secondary schools with eight sections. We cannot test for this hypothesis because the MINEDU does not collect this information. Furthermore, personnel at the MINEDU do not think such a discontinuity exists at the threshold of eight sections.
[16] The data can be obtained from the website of the Ministry of Education after submitting a formal request: http://umc.minedu.gob.pe/consultas/. Accessed on 6/20/2019.
[17] This choice gives us conservative estimates of the impact of *JEC*. For example, if we used only the mean and standard deviation from the pool of public schools, our results would be larger. Nonetheless, we opted for a conservative estimate in our paper.
[18] This questionnaire also asked students about their perceptions regarding teachers' behavior. These include 16 questions around clarity of the purpose of the class, explanations, speed of topics, participation in class and feedback. Unfortunately, these questions requested students to combine their views for their math and reading teachers. This limits the possibility to explore changes in teaching practices as results of JEC, separately by subject.

[19] The data can be obtained from the MINEDU website, after submitting a formal request: https://www.minedu.gob.pe/semaforo-escuela/repositorio.php. Accessed on 6/20/2019.
[20] YL children were selected at the age of one from a random sample of 20 districts, from the universe of district, excluding the 5% wealthiest districts. The original sample was composed of 2,052 children, with 100 children per district. This cohort was first visited in 2002 (at the age of 1) and revisited at ages 5, 8, 12 and 15 in 2006, 2009, 2013 and 2016 (respectively).
[21] See Favara et al (2021) who summarize the YL study and provide information about data access.
[22] This is done using the school ID ("*codigo modular*"), which is included in the *ECE-S*, *Semáforo Escuela* and *ENDO* datasets. This information is also captured in the Education History module of the YL study. The authors obtained permission from the Young Lives Study to clean and match the school attended by the YL participants in 2016 to the 2013 school census.
[23] For ECE, the full set of controls is considered. The *Semáforo Escuela* and *ENDO* data are collected at the school level and, thus, the vector of students' characteristics is omitted. For YL, we used the same set of controls as in ECE and add household wealth index in 2002 (the first visit to the YL families) as an additional socio-economic control. For YL, school district fixed effects are replaced by YL cluster fixed effects to take into account the YL sampling design.
[24] Larger schools tend to have a morning and an afternoon shift and are less likely to be eligible. See Appendix Figure A7 for a breakup by shift-type.
[25] Figure A2 plots the average values of a set of predetermined variables grouping the sample of all public secondary schools with morning shift by number of sections: start and end time of the school schedule, length of the school day, access to welfare programs, share of girl students enrolled, passing rates (all and by gender), use and teaching of indigenous language and whether the school has a morning shift only. The data come from the 2013 *Censo Escolar*, which reports information at the school level. It is easy to observe that for all the variables, there is a smooth transition around the threshold of eight sections. As discussed in section 3, the initial selection of schools used the 2013 *Censo Escolar*. The rules for selection into *JEC* were devised in 2014 and made public in October of that year. Thus, schools were not able to alter the number of sections back at the beginning of the school year in 2013. Together, these results confirm that the timing of the reform eliminates the possibility of manipulation of the assignment variable. Appendix Figure A3 similarly shows student-level predetermined characteristics by number of sections: age, gender, whether s/he attended kindergarten, whether s/he repeated a grade, and her/his mother's education and language. We show the same smooth transitions around the threshold for this set of variables. Appendix Tables A4 and A5 provide the regression results, reinforcing the validity of our identification strategy.
[26] The administration of the *ECE-S* math and reading test is compulsory in public and private schools. In practice, the administration of *ECE-S* might fail due to logistical reasons (e.g. due to a blocked road). In 2015, according to MINEDU (http://umc.minedu.gob.pe/evaluacion-censal-de-estudiantes-ece-2015/), ECE-S was administered in 99.5% of schools. Based on our analysis, when the 2015 *ECE-S* and the 2013 school census are matched, we find 6,121 eligible schools in both datasets and 492 eligible schools observed in the school census but not in *ECE-S*. One possibility is that schools were operating in 2013 but not in 2015. Importantly, 94% of the attrited schools have 5 sections or fewer, compared to 64% in non-attrited schools, which suggests attrition is only a problem for schools that are far away from the 7/8 threshold. In particular, only 3 eligible schools with 7 sections failed to take the ECE-S, and 0 eligible schools with 8 sections failed to take the *ECE-S*. Therefore, attrition is not a concern for our analysis.
[27] Pop-Eleches and Urquiola (2013) have shown that behavioral responses are indeed possible in the context of education. They find in Romania that when children attend a better school, they feel marginalized, their parents reduce their efforts when helping them with homework, and teachers sort themselves within the schools. Furthermore, they show that these negative behavioral changes tend to occur early on and after a few years, the effects reduce or even vanish. Thus, one could expect similar behavioral changes with the expansion of the school day. For example, Levin and Tsang (1987) introduce a model of effort and show that if the previous length of the school day represented an equilibrium, extending the number of hours could bring no effect on test scores because students and teachers could reduce their effort levels per hour of instruction.
[28] The number of schools with 7(8) sections observed in our four datasets is as follows: ECE-S: 267(200); *Semáforo Escuela*: 123(134); *ENDO*: 12(9); YL: 5(10).
[29] Figure A4 in the Appendix shows the distribution of the number-of-sections variable across our four datasets. In particular, this figure shows that compared to the universe of public schools with only morning shift (as informed by ECE-S), for *ENDO* and YL, there is a relatively small number of schools to the left of the distribution.
[30] Each index is the weighted average of a group of selected variables. Each of these variables is standardized with a mean of zero and a variance of one. Prior to standardization, the order of the variables for which higher values reflect

non-desirable results is reversed. When there are missing values, and this is not due to filtering, we impute the average to the missing observation (two averages are considered, depending on whether the observation is from a school that has 8 sections or more or fewer than 8 sections).

[31] The intuition behind the FDR approach is to allow researcher to tolerate a certain number of tests to be incorrectly discovered. For instance, an FDR *adjusted* p-value of 0.05 implies that 5 percent of significant tests result in false positives compared with an *unadjusted* p-value of 0.05 that implies 5 percent of all tests result in false positives.

[32] Hincapie (2016) suggests that the expansion in Colombia led to an increase of at most 0.10 SD for 9th graders. In Mexico, Padilla-Romo (2022) reports near-zero effects in the first year of the implementation and up to just under 0.14 SD four years after but for students in primary school (third to sixth graders).

[33] Information provided by MINEDU based on a request submitted by researchers from GRADE. The cost of 571 USD (calculated as the ratio between a cost of 192.3 million USD from 2015 and 337,051 students) corresponds to the total cost of the *JEC* reform in 2015 and is explained mainly by the cost of additional hours for teachers, the cost of additional personnel (pedagogical and administrative), the cost of the English teaching module, and the cost of the new infrastructure.

[34] As a reference, in Figure A6 in the Appendix we show the results of running a bandwidth sensitivity analysis for the 2SLS results for math and reading, considering the largest symmetrical bandwidth (from 1 to 14 sections) and moving gradually to smaller bandwidths (5 to 10, 6 to 9, and 7 to 8 sections). Results for math test scores remain robust to several bandwidth choices. Point estimates for reading test scored are consistently smaller relative to math and become statistically insignificant as bandwidth reduces.

[35] In particular, parameter κ is calculated assuming a change in the linear assumption by 0.25%, 1.25%, 5%, 10%, 25%, and 125%, respectively

[36] As a reference, in Figure A9 in the Appendix we show the results of running a bandwidth sensitivity analysis for the *Semáforo Escuela* results with 2SLS, considering the largest symmetrical bandwidth (from 1 to 14 sections) and moving gradually to smaller bandwidths (5 to 10, 6 to 9, and 7 to 8 sections). Results for *Semáforo Escuela* remain statistically significant in all of the scenarios considered.

[37] According to information obtained from interviewing MINEDU staff in charge of executing *JEC*, computer labs at JEC schools were used to teach students how to use software such as Microsoft Word and Microsoft Excel.

[38] Including access to higher education, risk behaviors, and teenage pregnancy (Sánchez and Singh, 2018; Favara and Sánchez, 2017; Favara, et al., 2020).

[38] The first scale measures access to digital devices (computers, laptops, tablets, smartphones) and to the Internet (72% of the sample reports having used digital devices recently). The second and third scales, which are only applied if the child has access to digital devices and to the Internet (respectively), measure the skills that child has using computer and browsing the Internet.

[40] Each question had a multiple-choice response (i.e., never, rarely, very often, always). Answers selecting “very often” or “always” were coded as one and zero otherwise.

[41] Peru has no legal restrictions on the enrolment of children who do not live in the district where the school is located.

[42] It is important to note that in the YL sample there were 217 students who left *JEC* schools in 2014, before the reform began. Movement in and out of schools might not necessarily be due to reform. The comparison of pre-*JEC* math test scores between incoming (-0.40) and outgoing (-0.26) students (compared to 0.05 for stayers) shows that the net sorting effect, if anything, is negative, although it might be compensated by the fact that there are more outgoing than incoming students.

Tables and figures

Table 1. Distribution of weekly hours by type of public high school

Subject	Regular (1)	JEC (2)
Mathematics	**4**	**6**
Reading	**4**	**5**
English	**2**	**5**
Science	**3**	**5**
History	3	3
Work education	**2**	**3**
Civics	**2**	**3**
Person, family & community	2	2
Physical education	2	2
Art	2	2
Religion	2	2
Tutoring	**1**	**2**
Free	6	5
Total	**35**	**45**

Note: A pedagogical hour is 45 minutes long. Source: Peru Ministry of Education (MINEDU).

Table 2. First stage: Participation in *JEC*

	Two stage least squares (2SLS)		**Local randomization approach**
	Linear splines	**Quadratic splines**	
	(1)	(2)	(3)
Coef.	0.590***	0.599***	0.596
s.e	[0.024]	[0.034]	-
P-value	-	-	(0.000)
N	6121	6121	467

Note: Data come from 2015 ECE-S. The unit of observation is the school. In columns 1 and 2, the sample corresponds to all secondary public schools with only morning shift (eligible schools). The results reported in columns 1 and 2 correspond to Equation (1) (2SLS, first-stage) and were estimated using a linear probability model with linear and quadratic splines (respectively) for the running variable, controlling for a dummy for urban location, a dummy for sections in multiples of five and fixed effects by school district. Robust standard errors clustered at the school district are shown in brackets. For the local randomization approach (column 3), the sample is restricted to those eligible schools with 7 or 8 sections (the smallest bandwidth). * p<0.10, ** p<0.05, *** p<0.01.

Table 3. Impact of *JEC* on academic achievement

	Two Stage Least Square (2SLS)		Local randomization approach
	Linear splines	Quadratic splines	
	(1)	(2)	(3)
Reading Test score			
Coef.	0.102***	0.014	0.142
s.e	[0.038]	[0.068]	-
P-value	-	-	(0.000)
N	187439	187439	15297
F-stat	516.9	178.3	-
Adj. R-sq	0.324	0.324	-
Reading: student test performance is at grade-level			
Coef.	0.029***	0.002	0.013
s.e	[0.008]	[0.013]	-
P-value	-	-	(0.021)
N	187439	187439	15297
F-stat	516.9	178.3	-
Adj. R-sq	0.124	0.124	-
Math Test score			
Coef.	0.185***	0.137*	0.212
s.e	[0.045]	[0.075]	-
P-value	-	-	(0.000)
N	187420	187420	15296
F-stat	516.4	178.0	-
Adj. R-sq	0.245	0.246	-
Math: student test performance is at grade-level			
Coef.	0.042***	0.025*	0.031
s.e	[0.009]	[0.013]	-
P-value	-	-	(0.000)
N	187420	187420	15296
F-stat	516.4	178.0	-
Adj. R-sq	0.075	0.076	-

Note: Data come from 2015 ECE-S. The unit of observation is the student. In columns 1 and 2, the sample corresponds to all students attending secondary public schools with only morning shift (eligible schools). Columns 1 and 2 report results using the discontinuity at 8 sections to measure the impact of JEC as in Equation (2) (2SLS, second-stage), with linear and quadratic splines, respectively, controlling for age and gender of the student, a dummy for mother tongue, as well as fixed effects for their mothers' educational attainment and language spoken, together with a dummy for urban location, a dummy for sections in multiples of five and fixed effects by school district. F-stat refers to the instrument in the first stage. Robust standard clustered at the school district are shown in brackets (* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.). Column 3 reports results using the local randomization approach for fuzzy RD, and corresponding asymptotic p-values. In this case, the sample is the subset of students attending eligible schools with 7 or 8 sections (the smallest bandwidth).

Table 4. Impact of *JEC* on math test scores
Differences-in-discontinuity approach (reduced form)

	Linear splines (1)	Quadratic splines (2)
Math (=1)	0.081***	0.053**
s.e.	[0.017]	[0.027]
Math*1(S≥8)	0.036**	0.072***
s.e.	[0.016]	[0.026]
N	374859	374859
adj. R^2	0.029	0.029
Mean	-0.295	-0.295
Reduced form gains in math	0.111	0.086

Note: Data comes from 2015 ECE-S. The unit of observation is student performance, by subject. The sample corresponds to all students attending secondary public schools with only morning shift (eligible schools). Each column reports the results of a regression with student fixed effects, controlling for linear/quadratic splines that vary by subject (see Equation (3)). The binary variable for the threshold is dropped due to the inclusion of student fixed effects. The last row (reduced form gains in math) is the coefficient for 1(S≥8) in a reduced form regression for math including the controls as in Table 3. Robust standard errors clustered at the school district are shown in brackets (* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$).

Table 5. Impact of *JEC* on school infrastructure, staff resources, and pedagogy

	Two Stage Least Square (2SLS)							**Local Randomization**		
	Linear splines			**Quadratic splines**						
	Coef	**s.e.**	**Adjusted p-value**	**Coef**	**s.e.**	**Adjusted p-value**	**N**	**Coef**	**P-value**	**N**
Panel A: School access to IT (index)	**0.725*****	**(0.121)**		**0.448**	(0.600)		**3,774**	**1.042**	**0.000**	**257**
Number of classrooms with computer	1.009***	(0.361)	0.008	1.827	(1.797)	0.115	3,822	1.303	0.007	257
Number of classrooms with laptop	4.174***	(0.808)	0.000	4.233	(3.035)	0.000	3,822	6.982	0.000	257
Technical equipment receives maintenance	0.061	(0.045)	0.200	-0.158	(0.480)	0.358	3,774	0.144	0.013	257
Kit *Robotica*	0.166***	(0.053)	0.003	-0.595	(0.471)	0.052	3,794	0.062	0.465	256
Panel B: School staff (index)	**3.004*****	**(0.171)**		**4.208*****	(1.402)		**2,363**	**3.278**	**0.000**	**152**
Complete teaching staff	-0.225***	(0.057)	0.000	1.118**	(0.494)	0.199	3,822	0.008	0.922	257
Number of security staff	2.542***	(0.085)	0.000	2.196***	(0.414)	0.000	2,970	2.378	0.000	198
Number of maintenance staff	0.524***	(0.138)	0.000	2.698*	(1.424)	0.298	2,363	1.388	0.000	152
Panel C: School pedagogical support (index)	**1.442*****	**(0.088)**		**1.605****	(0.669)		**3,822**	**1.388**	**0.000**	**257**
School has a psychologist	0.837***	(0.038)	0.000	0.778**	(0.332)	0.000	3,822	0.809	0.000	257
Schools receives program "*Acompañamiento pedagógico*"	0.637***	(0.053)	0.000	0.374	(0.461)	0.000	3,822	0.623	0.000	257
Teachers provide support to parents	0.150***	(0.041)	0.000	0.470	(0.315)	0.623	3,822	0.135	0.050	257

Note: Data from 2015-16 *Semaforo Escuela*. Each coefficient comes from a different estimation. The unit of observation is the school. In columns 1 and 2, the sample corresponds to all secondary public schools with only morning shift (eligible schools), excluding schools that became *JEC* in 2016. Columns 1 and 2 report results using the discontinuity at 8 sections to measure the impact of *JEC* as in Equation (2) (2SLS, second-stage), with linear and quadratic splines, respectively, controlling for a dummy for urban location, a dummy for sections in multiples of five, as well as fixed effects by school district, and month of interview. Robust standard errors clustered at the school district level (* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$) and adjusted p-values that correct for multiple hypothesis testing are reported. Column 3 reports results using the local randomization approach for fuzzy RD, with corresponding asymptotic p-values. In this case, the sample is the subset of students attending eligible schools (excluding schools that became *JEC* in 2016) with 7 or 8 sections (the smallest bandwidth).

Table 6. Impact of *JEC* on time use of students (in hours)

	Two Stage Least Square (2SLS)			
	Linear splines			
	Coef	s.e.	Adjusted p-value	N
Student time use during a typical school day				
Sleeping	-0.604***	(0.108)	0.000	685
Caring for household members	-0.286	(0.220)	0.307	685
Household chores	-0.667***	(0.209)	0.004	685
Domestic tasks	-0.130	(0.252)	0.752	685
Paid activity	0.025	(0.054)	0.784	685
In school	2.083***	(0.246)	0.000	685
Studying outside school	-0.466**	(0.189)	0.034	685
Leisure activities	-0.585*	(0.314)	0.122	685

Note: Data from 2016 Young Lives. Each coefficient comes from a different estimation. The unit of observation is the student. Sample includes all Young Lives participants attending secondary public schools with only morning shift (eligible schools), excluding schools that became *JEC* in 2016. Regressions as in Equation (2) (2SLS, second-stage), with linear splines, controlling for child's age, sex and language, household wealth index, urban location, sections in multiples of five, community fixed effects. Robust standard errors clustered at the community level (* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$). Adjusted p-values that correct for multiple hypothesis testing also reported.

Table 7. Impact of *JEC* on socio-emotional and technical skills

	Two Stage Least Square (2SLS)			
	Linear splines			
	Coef	s.e.	Adjusted p-value	N
Aspirations (index)	0.131	(0.123)		685
Aspirations for higher education	0.022	(0.019)	0.337	685
Aspirations for university	0.062	(0.081)	0.552	685
Socio-emotional skills (index)	0.213	(0.163)		685
Self-efficacy	0.086	(0.129)	0.605	685
Self-esteem	0.066	(0.099)	0.601	685
Pride	0.192	(0.133)	0.222	685
Agency	0.187	(0.116)	0.167	685
Technical skills (index)	0.407**	(0.177)		685
Ability to speak English	0.177**	(0.087)	0.078	685
Access to digital devices	0.055	(0.130)	0.734	685
Computer skills	0.237*	(0.130)	0.116	685
Internet skills	0.090	(0.087)	0.413	685

Notes: Data from 2016 Young Lives. Each coefficient comes from a different estimation. The unit of observation is the student. The same sample and methodology reported in Table 6 used here. Robust standard errors clustered at the community level (* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$). Adjusted p-values that correct for multiple hypothesis testing also reported.

Table 8. Impact of *JEC* on students' self-perceptions

	Perception index (1)	Understands any topic (2)	Learns new material without difficulty (3)	Understand difficult topics (4)	Confident during tests (5)	Helps peers (6)	Does homework without help (7)	Confident about passing the course (8)	Good at solving tasks (9)	Feels more capable while learning (10)	Feels good at the subject (11)
Panel A. Perceptions about reading											
Coef.	-0.077***	-0.035***	-0.034***	-0.047***	-0.035**	-0.025*	-0.036***	-0.038***	-0.046***	-0.037***	-0.025***
s.e.	[0.020]	[0.012]	[0.012]	[0.013]	[0.014]	[0.014]	[0.012]	[0.013]	[0.013]	[0.012]	[0.010]
Adjusted p-value	0.000	0.005	0.010	0.001	0.018	0.108	0.006	0.008	0.001	0.004	0.014
N	165975	165975	165975	165975	165975	165975	165975	165975	165975	165975	165975
F-stat	526,0	526,0	526,0	526,0	526,0	526,0	526,0	526,0	526,0	526,0	526,0
Adj. R-sq	0.050	0.035	0.029	0.030	0.024	0.022	0.026	0.024	0.028	0.018	0.016
Panel B. Perceptions about math											
Coef.	-0.023	-0.022	-0.009	-0.004	0.013	-0.042**	-0.016	-0.003	-0.002	-0.020	-0.008
s.e.	[0.022]	[0.015]	[0.015]	[0.012]	[0.015]	[0.017]	[0.014]	[0.013]	[0.015]	[0.014]	[0.014]
Adjusted p-value	0.317	0.165	0.581	0.734	0.395	0.014	0.276	0.859	0.895	0.179	0.581
N	165304	165304	165304	165304	165304	165304	165304	165304	165304	165304	165304
F-stat	530.4	530.4	530.4	530.4	530.4	530.4	530.4	530.4	530.4	530.4	530.4
Adj. R-sq	0.047	0.033	0.027	0.020	0.024	0.020	0.029	0.025	0.039	0.016	0.031

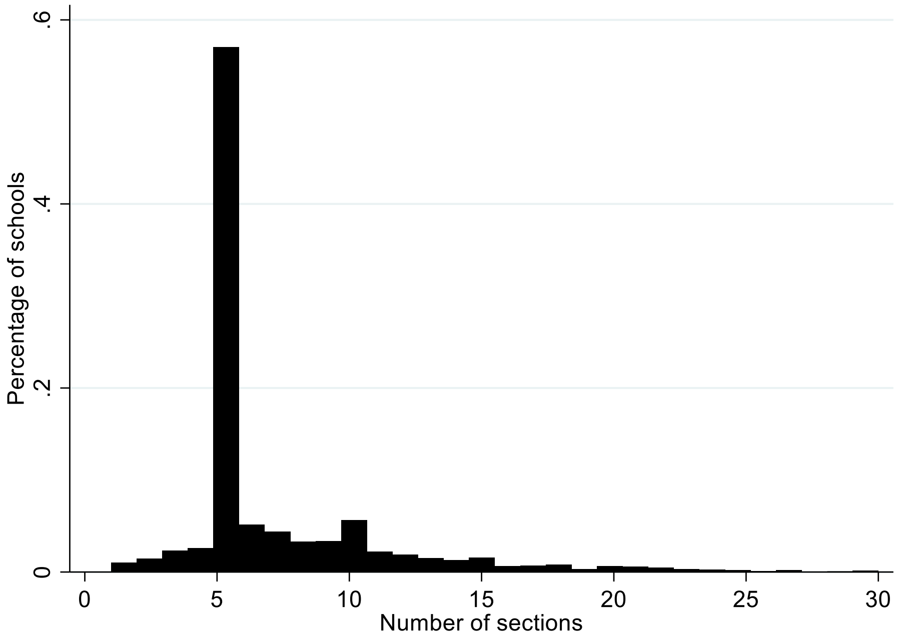
Note: Data come from 2015 ECE-S. Each coefficient comes from a different estimation. The unit of observation is the student. The same sample and methodology used for 2SLS estimations in Table 3 are used here (linear splines). F-stat refers to the instrument in the first stage. Robust standard clustered at the school district and associated significance reported (* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$). Adjusted p-values that correct for multiple hypothesis testing also reported.

Table 9. Impact of *JEC* on teacher characteristics, training, and behaviors

	Two Stage Least Square (2SLS)			
	Linear splines			
	Coef	**s.e.**	**Adjusted p-value**	**N**
Panel A: Teacher's time use				
Teacher time dedicated to school activities (weekdays) index	-0.080	(0.272)	0.831	1,569
Teacher time dedicated to school activities (weekend) index	0.066	(0.207)	0.817	1,569
Teacher time dedicated to activities outside school (weekdays) index	-0.172	(0.262)	0.607	1,569
Teacher time dedicated to activities outside school (weekend) index	0.025	(0.148)	0.908	1,569
Spent time in other income activities	0.065	(0.138)	0.719	1,569
Other income activity: work in other school (public)	-0.003	(0.034)	0.958	1,569
Other income activity: work in other school (private)	0.083	(0.069)	0.316	1,569
Other income activity: own business	0.145	(0.133)	0.372	1,569
Teach in more than 1 high-school	0.011	(0.053)	0.883	1,548
Panel B: Teacher attitudes and satisfaction (index)	**-0.277***	**(0.156)**		**1,564**
Total score on items indicating teacher's satisfaction	-1.606	(1.142)	0.236	1,569
Teacher's satisfaction with his/her job	-0.391**	(0.190)	0.070	1,569
Positive perception of teacher profession	0.269	(0.185)	0.218	1,569
Would choose again to be a teacher	-0.119	(0.128)	0.452	1,569
Teacher is happy with current work	-0.133	(0.098)	0.258	1,568
Decided to be a teacher by choice	-0.185*	(0.099)	0.102	1,564
Panel C: Teacher training & pedagogy (index)	**-0.107**	**(0.172)**		**1,560**
Number of subjects currently teaching	0.149	(0.291)	0.697	1,560
Developed innovative practices	-0.102	(0.142)	0.572	1,569
Top quintile in good teaching practice score	-0.092	(0.107)	0.494	1,564
Received "Acompañamiento pedagógico"	0.062	(0.099)	0.624	1,569
Received ICT training	-0.014	(0.170)	0.955	1,568
Panel D: Teacher predetermined characteristics (index)*	**0.229**	**(0.160)**		**837**
Age	-0.374	(3.733)	0.945	1,565
Female	0.017	(0.161)	0.939	1,569
Completed studies at university	0.209	(0.180)	0.338	1,442
Completed studies at private institution	0.044	(0.095)	0.726	1,440
Completed any post-graduate studies	0.048	(0.145)	0.811	1,569
Number of years teaching in current secondary school	6.295	(5.600)	0.355	837
Top Levels at Escala Magisterial	-0.103	(0.100)	0.401	1,567
Teacher has a permanent contract ("Nombrado/a")	-0.219	(0.237)	0.457	1,569

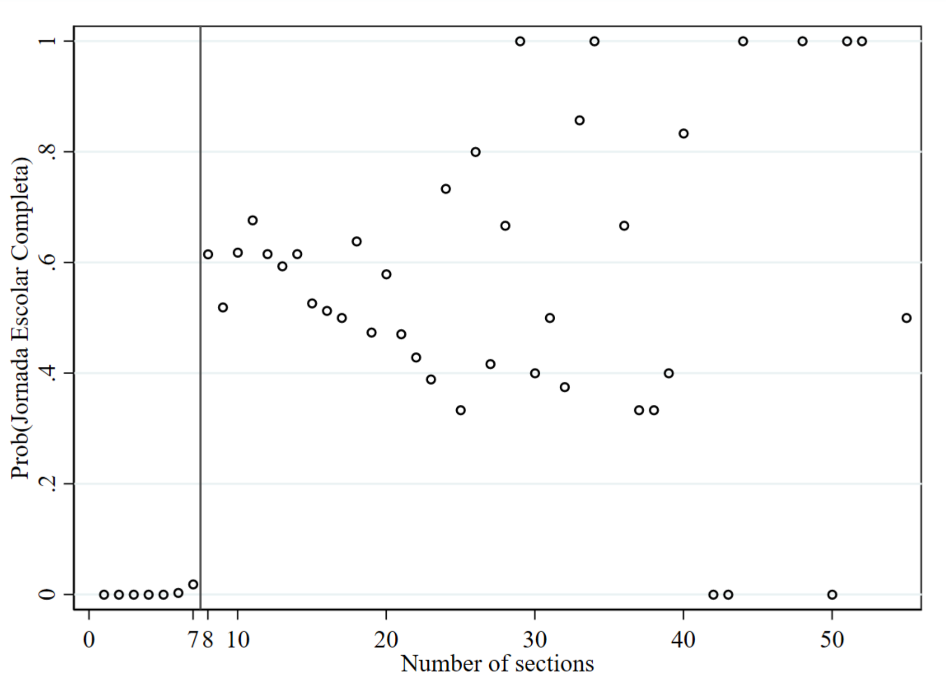
Note: Data come from 2016 *ENDO*. Each coefficient comes from a different estimation. The unit of observation is the teacher. The sample includes all teachers from secondary public schools with only morning shift (eligible schools), excluding schools that became *JEC* in 2016. Regressions use the discontinuity at 8 sections to measure the impact of *JEC* as in Equation (2) (2SLS, second-stage), with linear splines, and controlling for a dummy for urban location, a dummy for sections in multiples of five, as well as fixed effects by school district. Robust standard clustered at the school district and associated significance reported (* p < 0.10, ** p < 0.05, *** p < 0.01). Adjusted p-values that correct for multiple hypothesis testing also reported.

Figure 1. Number of sections in secondary public schools with morning shift only



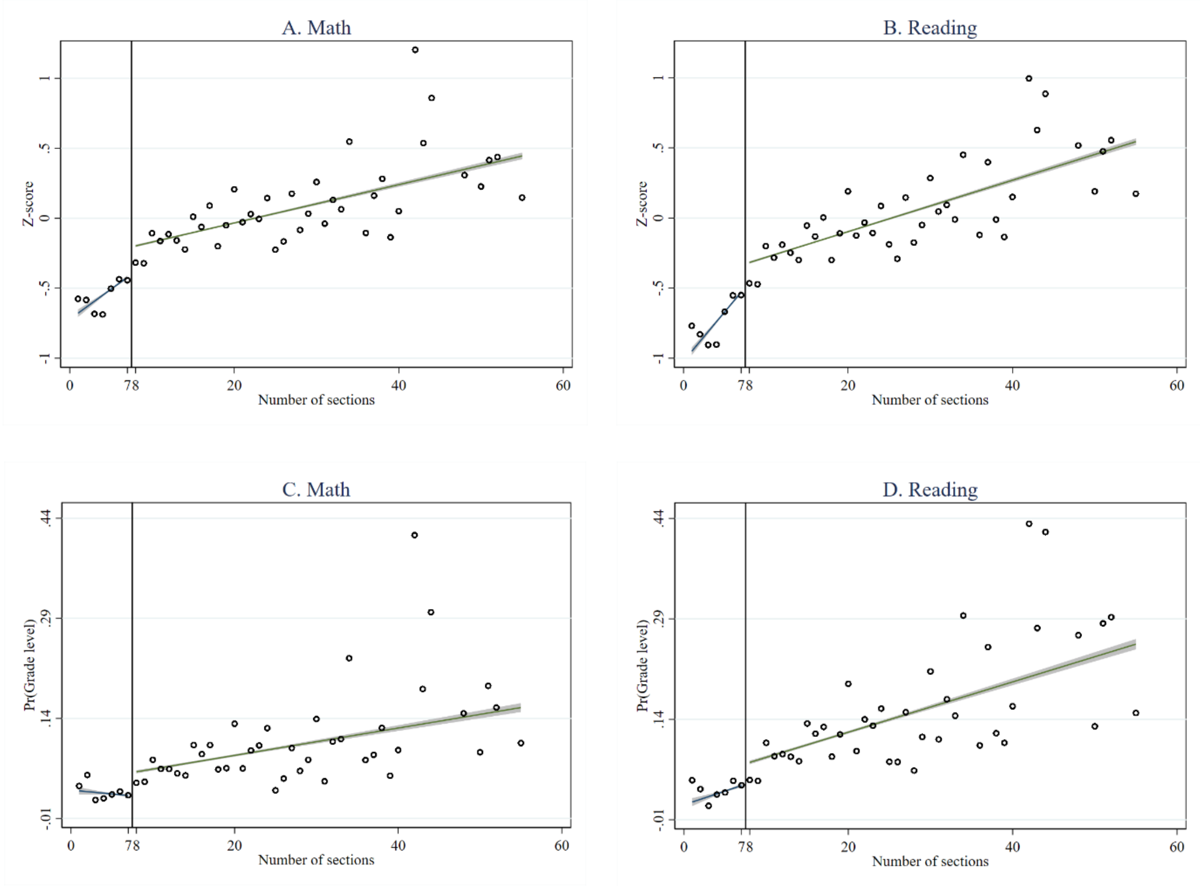
Source: 2013 school census. Distribution truncated at 30 sections for illustrative purposes (full distribution reported in Figure A1 in the Online Appendix).

Figure 2. First Stage: participation in *JEC* by number of sections



Note: Each dot represents the share of schools that belong to JEC, by number of sections. Sample restricted to public secondary schools with morning shift only. One school with 78 sections omitted. *Source*: Authors' calculation based on *2013 Censo Escolar*.

Figure 3. Impact of *JEC* on test scores (reduced form)



Note: Each circle represents the sample average by section. Sample restricted to secondary public schools with morning shift only. One school with 78 sections omitted. Data source: 2015 ECE-S.