Sibling Spillover Effects in Education: Evidence From an Extended School-day Reform in Peru

Gerald McQuade*†   Catherine Porter †   Alan Sanchez ‡§

April 28, 2025

Abstract

We estimate sibling spillover effects on attainment from a national reform that extended the length of the school day and improved other inputs within public secondary schools in Peru. Using regression discontinuity design based on school eligibility criteria, we estimate a positive effect of older siblings' increased schooling on younger sibling attainment, increasing test scores by 0.12 S.D. and 0.14 S.D. in reading and mathematics. At the extensive margin, siblings are also more likely to achieve an "in-progress" grade or above (5.4 p.p.). Positive spillover effects are driven by the effect on girls, with the largest effect amongst sister-sister pairs, compared to null effects for younger brothers. Our results indicate that evaluations which consider only the benefits for the targeted child could systematically understate the benefit-cost ratio of educational reforms.

JEL codes: I26, I28, J24, O15
Keywords: Human capital, Cognitive skills, Spillovers, Regression discontinuity design

---

*Corresponding author: Gerald McQuade, g.mcquade1@lancaster.ac.uk. This paper uses confidential data prepared by the Ministry of Education in Peru (MINEDU) to the authors' specification. This data can be obtained by filing a request online directly with MINEDU at: https://enlinea.minedu.gob.pe/.

†Department of Economics, Lancaster University, Lancaster, UK

‡Oxford Department of International Development, University of Oxford, Oxford, UK

§Grupo de Análisis para el Desarrollo (GRADE), Lima, Peru

1 Introduction

Educational reforms generally target specific individuals, with subsequent policy evaluations commonly focused only on the direct effect on those individuals. However this practice does not capture any potential spillover effects the policy has on others in their family, including siblings. Siblings grow up in the same home, belong to many of the same social groups, experience the same social interactions, and strongly influence each others' choices and outcomes (Black et al., 2021). Importantly, if policies targeting one child have an unanticipated and unmeasured impact on their siblings' outcomes, such evaluations will systematically under- or over-estimate the cost-benefit ratio (Altmejd et al., 2021; Figlio et al., 2023). It is therefore salient for policymakers to quantify potential spillovers, particularly when they affect those not eligible for the reform. However, it can be difficult to causally identify the influence of siblings on each other, given the "reflection problem" which is prevalent in the study of peer effects (Manski, 1993). This is particularly true outside of data-rich high-income contexts, where it remains unclear how sibling spillovers propagate through households which may be subject to credit constraints, and in which the cost of decisions, such as additional schooling and work, may be greater.

This chapter examines the potential for sibling spillovers on younger siblings' outcomes in Peru, exploiting exogenous variation in older siblings' schooling as a result of *Jornada Escolar Completa* (or JEC), a nationwide extended school-day reform which added two pedagogical hours per day and improved school resources in 1,000 public secondary schools. This reform aimed to bring the weekly number of instructional hours and the quality of inputs and resources in treated public secondary schools into line with those offered by high-quality private schools in the country, which often have better learning outcomes than their public counterparts (Agüero et al., 2021). The reform was designed to be comprehensive, with increased pedagogical resources and pay for teachers to reflect the additional responsibilities and contact time, as well as providing students with access to psychological support and improved IT infrastructure. Assignment of schools to the program was not random, however we are able to exploit initial arbitrary eligibility criteria which were based on the number of home-rooms in the school (*secciones*, or sections). The final selection of schools was likely based on several unobservable decisions, therefore we use a fuzzy regression discontinuity design (FRDD) motivated under the assumptions of the local randomisation framework (Cattaneo et al., 2015, 2017, 2024).

Previous work has identified robust moderate to large positive effects of JEC on the learning outcomes of the targeted child (Agüero, 2016; Agüero et al., 2021). This chapter expands this evidence by considering for the first time the potential for sibling spillover effects of the JEC reform. Specifically, we assess the spillover effect of an exogenous change in older siblings' schooling on the learning outcomes of their younger, primary school-aged siblings, who are not exposed to the treatment.

Our findings indicate that primary school-aged younger siblings of a child attending a JEC high school experience a positive spillover effect on their educational attainment, increasing test scores by 0.12 S.D. and 0.14 S.D. in reading and mathematics respectively. While this positive effect on test scores does not significantly impact the likelihood of younger siblings attaining the top grade classification ("at grade", in line with expected outcomes for their grade), it does have a positive and statistically significant effect on the likelihood of being classed as "in progress" (the second highest classification) or above in mathematics (5.4 percentage points (p.p.)). This suggests that spillovers have a greater effect for younger siblings who are lower in the grade distribution, partially reducing the attainment gap. Our findings are consistent with other work which addresses sibling spillovers related to the quality or amount of schooling experienced by one sibling (Figlio et al., 2023; Qureshi, 2018a).

Considering heterogeneities, we find that spillover effects are isolated to younger sisters, regardless of older sibling gender, although the largest effects are found for sister-sister pairs. In contrast, no statistically significant effect is found for boys, regardless of the gender of their older sibling. This finding informs the limited evidence on how spillovers may propagate differently across sibling pair gender mix (Dahl et al., 2023; Nicoletti & Rabe, 2019; Qureshi, 2018a), and complements the literature on the gendered responses to household inputs in education (Autor et al., 2019). Finally, we show that our results are robust to a range of validation and falsification tests.

This study contributes to the growing literature assessing the potential for sibling spillovers in education (Altmejd et al., 2021; Dustan, 2018; Goodman et al., 2015; Gurantz et al., 2020; Joensen & Nielsen, 2018; Karbownik & Özek, 2023; Nicoletti & Rabe, 2019; Oettinger, 2000; Qureshi, 2018a, 2018b), in particular expanding the very limited evidence outside of high-income, data rich contexts. In doing so we provide, to the best of our knowledge, the first evidence of sibling spillovers resulting from a schoolday extension policy, widening the scope of benefits beyond just the targeted child. Our results are salient for policy evaluation, highlighting the importance of accounting for within-family externalities when determining the benefits and costs of educational reforms.

The rest of this chapter is set out as follows: section 2 provides a summary of the existing literature on sibling spillovers in education, and on school day extension reforms (subsection 2.1), as well as providing a conceptual discussion of how sibling spillovers may propagate (subsection 2.4 and subsection 2.5), and summarising the JEC reform and its selection criteria (subsection 2.2 and subsection 2.3). Section 3 describes the datasets and data matching process, including descriptive statistics. Our identification and estimation strategy, as well as pre-analysis checks are discussed in section 4. All results and discussion follow in section 5, and section 6 concludes.

2 Context

2.1 Literature Review

Our work relates to two strands of literature. First, it contributes to the evidence on the effects of school-day extension programs, which aim to address poor student performance with a focus on the importance of instructional time in learning outcomes (Ben-Porath, 1967; Carroll, 1963; Figlio et al., 2018). However, the evidence for its effectiveness is mixed. A US-focused review of literature suggests positive effects of school-day extension programs, in particular amongst students at risk of failing, however the weak research design of many studies makes it difficult to disentangle the effect of increased instructional time from other inputs, limiting the strength of any causal inference (Patall et al., 2010). Outside of a high-income setting, there may be a greater appeal for a transition from part-day to full-day schools, providing increased childcare and even school lunch programs (if offered), acting as a safety net for families (Pablo et al., 2015). However, there is also a higher opportunity cost, given smaller educational budgets and the lower quality of other complementary inputs (Agüero et al., 2021), although the design of JEC as a comprehensive reform with increased school resources and improved infrastructure may mitigate this issue. Furthermore, increased school time comes with a trade-off, reducing students' ability to work in family businesses, help with chores, or care for relatives, which may prove important in credit constrained and low-income contexts.

Within Latin America a number of countries have implemented school-day extension programs, including Mexico (Cabrera-Hernández, 2020), Argentina (Edo & Nistal, 2022; Llach et al., 2009), Dominican Republic (Garganta et al., 2022), Chile (Barrios-Fernández & Bovini, 2021; Bellei, 2009), Uruguay (Cerdan-Infantes & Vermeersch, 2007), Brazil (Almeida et al., 2016), and Peru. While the evidence suggests a mostly positive impact on learning, amongst other outcomes, there is a great deal of heterogeneity between studies, and a recent cost-benefit exercise suggests there are likely other more cost-effective policies as alternatives (see Pablo et al., 2015, for a review and in-depth discussion).

Second, it contributes to the literature on within-family spillovers. Although there is a developed literature addressing intergenerational spillovers (e.g. parent-to-child; see Black & Devereux, 2011), less studied is the potential for spillovers between siblings on each other's educational choices and human capital outcomes. As siblings grow up together and make choices concurrently, estimates of cross-sibling correlations in outcomes likely suffer from the same "reflection problem" identified in peer effects by Manski (1993). That is, it is difficult to infer from observed outcomes whether peer-group behaviour affects individual behaviour, or if group behaviour is simply the average of individual behaviour – even after controlling for shared characteristics. A further practical issue is that studying sibling spillovers generally requires access to high quality administrative data on students' enrolment and attainment, which can also be successfully linked to those

of their siblings and/or wider household (Dahl et al., 2023). As such, the evidence on sibling spillovers has been very limited until recently. In an early contribution, Oettinger (2000) uses two-stage least squares to estimate the impact of having an older sibling graduate high school on younger siblings' graduation rates in the US using the National Longitudinal Survey of Youth, finding a positive effect on younger siblings' graduation probability. However, in the absence of exogenous variation (e.g. a policy reform, or natural experiment) their instrument validity (sibling-specific background characteristics) relies on strong assumptions.

Several papers take advantage of country/state-wide school and health administrative data, matching healthcare records of siblings to an individual's attainment data, as well as data on school-peer abilities, to assess the impact on learning outcomes of having a sibling with a physical, mental, or learning disability (Black et al., 2021; Breining, 2014; Persson et al., 2021). More closely related to this chapter are studies which focus on spillovers across siblings resulting from educational choices or inputs. Generally, these exploit quasi-random variation created by a policy reform, providing exogenous variation in the treatment status of one sibling. A number of these studies focus on college or school course choice, exploiting pre-determined entry cut-off scores to assess the impact of an older sibling being accepted to a selective school, college, or major on the choices of younger siblings. Goodman et al. (2015) estimate the relationship between the college choices of siblings, finding younger siblings are more likely than their peers to enroll in 4-year and highly-competitive colleges when their older siblings do so first. Altmejd et al. (2021) assess spillovers in college and major choices in Chile, Croatia, Sweden and the US, using a regression discontinuity design with multiple college-specific thresholds for admissions, finding younger siblings are more likely to enrol in the same college and college/major combination if their older sibling was marginally accepted and attended, especially amongst social groups with lower college enrolment rates.[1]

Looking at high school level decisions, Dustan (2018) explores how an older sibling scoring above the threshold on a placement exam for high-quality schools in Mexico city positively influences the probability of younger siblings applying for those schools, while Gurantz et al. (2020) find that younger siblings in the US are more likely to take Advanced Placement courses if their older sibling marginally passes an exam. Similarly, Joensen and Nielsen (2018) exploit a pilot program in Denmark which provides variation in the requirements for older siblings to take advanced courses in secondary school, but not for younger siblings, to identify positive spillovers on the probability that younger siblings choose the same courses later on. They find younger siblings are more likely to choose advanced mathematics and science classes, especially amongst brothers who are close in

---

[1]Additionally, Aguirre and Matta (2021) also assess spillovers in college and major choices in Chile, finding a large effect of older siblings' college choices on younger siblings choices, however do not find any effect on the choice of major.

age, suggesting the sex-mix of the sibling dyad and age-spacing matter for spillovers. This finding is echoed by Dahl et al. (2023), who exploit admission thresholds based on GPA in Sweden for oversubscribed college majors. Their results suggest the magnitude and direction of sibling spillovers are dependent on the sex-mix of the dyad, with same-sex younger siblings more likely to apply for the same major as their older siblings'. Younger brothers are particularly likely to follow their brother if there is a larger than 3 year gap in age (therefore they were not enrolled in high school at the same time). In contrast, younger brothers are less likely to go for the same course as an older sister, more so when there is a less than 3-year gap in age.

Finally, the strand of work closest to our analysis is that which assesses spillover effects on younger sibling attainment, rather than course or college choice, by exploiting variation in the quality or amount of schooling experienced by older siblings. Nicoletti and Rabe (2019) estimate the spillover effect of having a high-achieving older sibling on younger siblings' achievement, using a fixed effects value-added model, and by instrumenting the test scores of older siblings with the mean scores of their peers, finding a significant positive spillover from older to younger siblings. Similarly, using student level data which links siblings via birth records, and further linking to school level data in North Carolina, Qureshi (2018a) estimates the spillover effect of a child being taught by a more experienced teacher on their older or younger siblings, finding a positive spillover to the younger child. While these solutions are useful as they don't rely on a specific policy reform for variation, leading to greater generalisability of results, they may require strict and potentially unrealistic assumptions (Sacerdote, 2014; Todd & Wolpin, 2003), and access to rich panel data which can be linked across schools and siblings, limiting their viability outside of data-rich high income contexts.

Another identification strategy exploits differences in the school starting age of older siblings, exploiting either sharp or fuzzy cutoffs in the date for school eligibility. Age at first entry to school has previously been shown to have a significant impact on the focal child, with children who are older relative to their peers showing better attainment at the same stages (Bedard & Dhuey, 2006; McEwan & Shapiro, 2008). Karbownik and Özek (2023) identify spillovers from having an older sibling who is born after the cutoff (hence one of the oldest in their class) on younger siblings in Florida, with effects concentrated in low socio-economic status households, who score higher on standardised tests. These results are similar to those found by Zang et al. (2023) in North Carolina.

The closest work to ours is Figlio et al. (2023), who assess the impact of an older sibling marginally missing a minimum reading score threshold in 3rd grade, leading to grade retention and being provided additional targeted support, on younger sibling outcomes in Florida. They find a large spillover effect of an older sibling's increased schooling and extra support on younger sibling scores in the same reading test.

The majority of the literature focuses on high-income contexts, where the mechanisms

through which spillovers transmit likely differ from those in low- and middle-income contexts, where older siblings can play a large role in the care of younger siblings, including helping with homework or tutoring. Qureshi (2018b) estimates how a more educated eldest sister impacts younger brothers' educational attainment in rural Pakistan. They use an instrumental variables approach, exploiting cultural norms for chaperones which create significant disparities in girls' access to schooling based on their distance to the nearest school, finding that a more educated older sister is associated with higher literacy, numeracy, and years of schooling for younger brothers.

This study will expand the literature on sibling spillovers in low- and middle-income countries by investigating how variation in the amount and quality of public schooling experienced by an older sibling may have a causal spillover effect on the outcomes of their younger siblings. Specifically we will assess the potential for externalities on the academic achievement of the younger sibling of a child who attends a Jornada Escolar Completa (JEC) public school in Peru. Our approach differs from other similar studies as assignment to a JEC school is not based on measures of the older siblings' academic ability, but rather on school level selection criteria driven primarily by budget constraints, discussed in detail below. Additionally, our study is the first to consider the potential for sibling spillovers as a result of a school-day extension reform, described in detail in the following section.

2.2 Jornada Escolar Completa

There has been a significant expansion in access to education in Peru since the beginning of the 21st century, with high rates of enrolment in both primary and secondary school, and low rates of children out-of-school, however there is significant heterogeneity in the quality of education and learning (Saavedra & Gutierrez, 2020). This is typified by Peru's performance in PISA (Programme for International Student Assessment), where it has consistently ranked near the bottom in mathematics, reading, and science, and ranked last among the 65 participating countries in 2012.

In light of this, Peru has enacted several reforms to the education system, including introducing the *Jornada Escolar Completa* (JEC) program in 2015. It was designed as a comprehensive reform, planned to improve the quality of public secondary education and close learning gaps (Escobar & Sanchez Castro, 2021). As a result of increased school enrolment in the 1970s, many Peruvian secondary schools operated with separate morning and afternoon shifts to allow for a greater capacity (Saavedra & Gutierrez, 2020). A major component of JEC was to introduce a full-day model, increasing the number of pedagogical hours from 35 to 45 hours per week (equating to 2 hours extra per school day) in 1,000 public secondary schools nationwide.[2] A breakdown of differences with regular shift-based

---

[2]A pedagogical hour is 45 minutes long.

public schools by subject is provided in Table A1. However, the reform was designed to be comprehensive, therefore the program aimed to also improve complementary inputs and resources, focusing on 3 components: 1) improved pedagogical support; 2) improved school management and organisational practices; and 3) improved physical infrastructure and increased IT resources. The pedagogical component includes a support programme for teachers (*Acompañamiento pedagogico*), as well the provision of psychologists to meet with students at least twice annually. Online support and training for school management was offered, and the salaries of teachers and principals were increased in line with the additional workload required. Finally, the number of computers and laptops available to classrooms was increased, with additional IT maintenance support being provided. These changes were designed to mimic the contact time and resources provided by high-quality private schools (Alcázar, 2016; Escobar & Sanchez Castro, 2021). Given valid concerns about the potential for low quality complementary inputs to limit the potential positive impacts of schooling reforms in low- and middle-income contexts (Kerwin & Thornton, 2021; Mbiti et al., 2019; Pablo et al., 2015), it is likely that improving these resources will lead to a more effective impact of increased instructional time (Agüero et al., 2021).

Previous research has identified large to moderate positive effects on a range of outcomes for the targeted child. Using a fuzzy regression discontinuity design, Agüero et al. (2021) find significant increases in scores for mathematics (0.23 S.D.) and reading (0.19 S.D.) for children who attend a school which was part of the first round of JEC implementation in 2015, as well as improvements in socio-emotional competencies and technical/digital skills. Using the Young Lives Peruvian survey, Sanchez and Favara (2019) also find effects on non-school based tests of mathematics and reading comprehension (0.13 and 0.19 S.D., respectively) and higher self-reported self-esteem and self-efficacy. Rodrigues and Campos Flores (2021) use a mixed research design to evaluate the impacts of JEC beyond the initial sample of schools. The expansion of the program in 2016-2017 followed different eligibility criteria than in 2015, widening the types of schools eligible, therefore the outcomes for students in schools made eligible in these years were assessed using propensity score matching and a difference-in-differences approach. While Fuzzy RDD estimates for schools made eligible in 2015 were similarly as large as those of Agüero et al. (2021), the effect for students in schools made eligible in 2016-2017 were smaller in magnitude (0.04-0.07 S.D.). This may reflect that as the list of eligible schools widens, additional schools may not benefit as much from changes, or may simply reflect potential omitted variable bias in estimates. Looking at the long-term impacts of JEC on higher education outcomes using the Young Lives COVID-19 phone surveys, Hidalgo Arestegui (2021) finds students in JEC schools are more likely to completely secondary school (11.3 p.p.). While the estimated effect for students' likelihood of accessing university was not statistically significant at conventional levels, students of JEC schools were more likely to enter into a STEM major/stream at university (16.5 p.p.). Sanchez and Favara (2019)

7

provide some support for these findings, seeing a 7.3 p.p. increase in students' aspirations to complete university education.

There is so far limited evidence of the wider impacts of JEC outside of the educational outcomes of the targeted child. Ortega (2018) and Sanchez and Favara (2019) assess how extended school-days may impact teenage fertility and sexual behaviour. Ortega (2018) estimate a moderate reduction in adolescent pregnancy (0.6 p.p.) using an instrumental variables approach. The findings of Sanchez and Favara (2019) provide some support for this, with male students having improved sexual health and contraceptive knowledge and female students showing increased pride, self-esteem, and sense of agency over their lives. Finally, and most closely related to our work, Ersoy and Forshaw (2023) assess the potential spillover effects of a child attending a JEC school on the labour market outcomes of parents, using a novel approach to match respondent households to schools via a geo-spatial algorithm. While they find weak evidence of a positive effect on fathers' individual income, this finding is not robust. Their dataset does not record the school of children in the household, relying instead on matching publicly-schooled children of secondary school age to the nearest public school by walking distance. Therefore their results by design are subject to measurement error, which may introduce significant noise or bias in estimates.

2.3 Selection Criteria

Eligibility for JEC was based on a number of criteria, summarised in Table A2. It applies only to public schools which operate a morning shift (allowing for the expansion of school hours without impacting afternoon shift students). Schools were also required to be large enough to accommodate for additional resources and infrastructure. Specifically, schools were required to have a minimum of 8 *secciones* (or sections, equivalent to home-rooms or form classes in the US or UK). This choice of 8 sections is arbitrary and primarily driven by budgetary constraints (Agüero, 2016). Given secondary education consists of five grades, schools most commonly have five sections, one per grade, and the density of schools remains relatively smooth around 8 sections as shown in Figure 1 (this is tested empirically in subsection 4.3).

Prior to implementing JEC, personnel from the Peruvian Ministry of Education (MINEDU) used data on school characteristics, infrastructure and enrolment from the 2013 school census (*Censo escolar*) to identify 1,360 eligible secondary schools which match these criteria. A further 52 "emblematic" schools, which may not have meet all requirements but were believed to benefit greatly from inclusion, were added to this list. The list of 1,412 potential schools was sent to local coordinators to validate that schools met the requirements, with several schools being removed and added, to select a total of 1,343 schools. MINEDU then hired evaluators to reduce this list down to 1,000 schools, which were then

Figure 1: Distribution of Public Morning-Shift Schools by Number of Sections
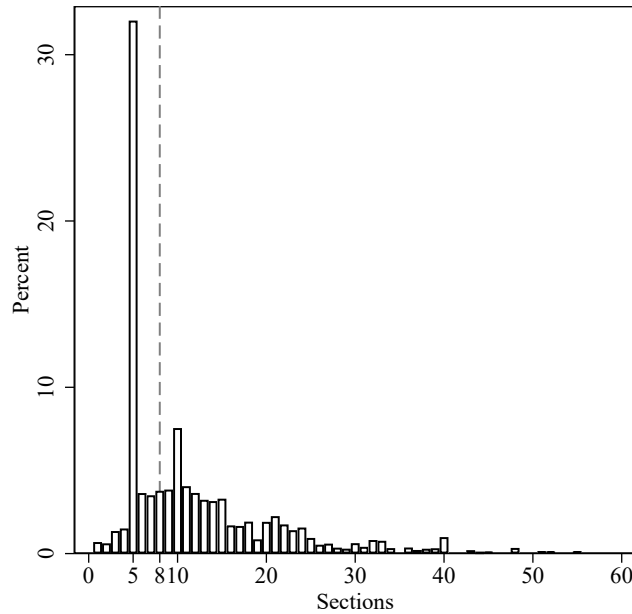
included in the September 2014 government directive announcing the JEC reform (RM Nº451-2014-MINEDU). Finally, this list was amended once more in February 2015 by replacing 6 schools (RM Nº062-2015-MINEDU), prior to the implementation of JEC for the school year beginning March 2015.

The process of moving from the original 1,360 eligible schools to the final selection of 1,000 schools is driven by unobserved characteristics, reflecting the potential influence of local coordinators bargaining with administrators (Agüero, 2016), and endogenous selection of "emblematic schools" and replacement schools. As such a simple comparison of sibling pairs attending included schools with those attending excluded schools would not be suitable. We instead exploit the initial criteria used to identify the first list of 1,360 schools. Specifically, we restrict our sample to schools which are publicly administrated and have only a morning shift registered, and exploit the discontinuous jump in participation in JEC at 8 or more sections.

Notably, while the JEC reform was expanded in 2016 and 2017 to include further schools, the eligibility rules changed to allow schools with different characteristics to join, therefore extending our identification strategy would not be feasible. As such we focus only on those schools eligible under the 2015 criteria, excluding from our analysis schools which joined JEC in subsequent years.

2.4 Spillover Transmission Mechanisms

Sibling spillovers can arise through two channels. First, spillover effects could occur through a within-family peer effects channel (Manski, 1993; Sacerdote, 2014) based on the

interactions between siblings. This could be due to direct interaction, with older siblings sharing knowledge and influencing behaviours by helping with homework or teaching skills. This could be of particular importance in low- and middle-income contexts, where often older siblings play an important role in caring for their younger siblings and may have more formal education than their parents (Qureshi, 2018b).

However, in this context the direction of effect is unclear *a priori*. In attending a JEC school, older siblings may have less time to spend in the household, directly impacting interaction by reducing the time they can spend interacting with their sibling, or indirectly by passing on these household responsibilities (e.g. household chores, unpaid labour or caring for others) to the younger sibling, negatively impacting younger sibling time use for schoolwork (substitution effect). Alternatively, given robust findings of the positive effects of increased instructional time and improved resources resulting from JEC on students' educational outcomes (Agüero et al., 2021; Rodrigues & Campos Flores, 2021), older siblings attending a JEC school may have greater knowledge, mastery of topics, and ability to provide more effective help and academic mentoring to younger siblings, leading to a greater return from time spent learning together (productivity/ability effect).

Furthermore this channel may lead to indirect spillovers. For example an older sibling attending a JEC school and benefitting from the reform may represent a role model, or encourage sibling rivalry and competitiveness, influencing the behaviours, aspirations, and choices of their younger siblings. This indirect channel is prevalent in the literature addressing how older sibling choices of school, college, or major impacts younger sibling educational choices (Aguirre & Matta, 2021; Altmejd et al., 2021; Dustan, 2018). Generally, spillovers occurring through this indirect channel are expected to be positive, although it could be possible that younger siblings respond negatively to older sibling achievement.

Second, an indirect consequence of a policy may be that parents reallocate limited household resources to take advantage of any policy benefits. Given the positive impact of increased instructional time and better school inputs, they may shift resources towards the targeted sibling (reinforcing behaviour; Becker & Tomes, 1986; Grätz & Torche, 2016), shifting resources away from other siblings and potentially negatively impacting their outcomes. Or they may shift resources toward the untreated younger sibling (compensatory behaviour; Fan & Porter, 2020; Pitt et al., 1990) leading to positive spillover effects due to benefitting from increased household inputs. Alternatively, parents could respond in a more complex manner, focusing on equalising inputs across siblings (potentially positive spillover effects; Berry et al., 2020), or by investing differently across differing dimensions of human capital (leading to an ambiguous overall effect; Yi et al., 2015).

Finally, sibling spillovers may operate differently across sibling pair characteristics (Black et al., 2021; Karbownik & Özek, 2023), in particular by sibling pair gender mix, with potentially stronger effects for same gender pairs (Karbownik & Özek, 2023; Nicoletti

& Rabe, 2019; Qureshi, 2018a; Zang et al., 2023) compared with mixed-gender sibling pairs, although such effects can vary across country contexts (Altmejd et al., 2021), and may not be as clear in low- and middle-income contexts, where for example older sisters may have disproportionately more responsibility for caring for and educating younger siblings (Qureshi, 2018b).

Unfortunately a limitation of this study is that we cannot directly address the potential underlying mechanism due to data limitations, however we do provide some context for how spillovers are transmitted by exploring how effects may propagate differently across sibling characteristics.

2.5 The Reflection Problem

As noted above, simple estimates of correlations between siblings' educational outcomes likely suffer from the "reflection problem". This problem is pervasive in studies of peer effects, where it is hard to disentangle from observed outcomes whether peer-group behaviour affects an individual's behaviour, or if group behaviour is the aggregate of individual behaviour. Manski (1993) identifies three component reasons for why peers may exhibit similar outcomes: 1) correlated effects, where individuals tend to behave similarly due to having similar individual characteristics or environmental influences; 2) exogenous effects, where individual behaviour varies with the exogenous characteristics of the group; and 3) endogenous effects, where the behaviour of an individual is influenced by the behaviour of others in the group. The endogenous effect between group members (e.g. siblings in a household) is the effect of interest in this study, however even if one can adequately control for exogenous aspects at the group level (e.g. socioeconomic status, urban/rural location, etc.), simple estimates of group-member correlations will likely still be contaminated with unobserved correlated effects. For siblings, the issue of correlated effects is even greater than with larger peer groups, where it is possible that random assignment to the group can be exploited. We therefore attempt to isolate the endogenous effect of siblings' schooling on each other by exploiting exogenous variation in the amount of schooling experienced by one sibling, but not the other in the pair, due to the Jornada Escolar Completa reform.

3 Data

Data for this analysis comes from three sources: i) the *Evaluación Censal de Estudiantes* (ECE), a national assessment of student learning; ii) *Sistema de Información de Apoyo a la Gestión de la Institución Educativa* (SIAGIE), a digital platform used by schools to manage administrative student data, and iii) the *Censo Escolar*, an annual school level census. Student-level data from ECE was matched to corresponding records in SIAGIE.

As this data includes confidential, identifying data, this matching was carried out by MINEDU, based on the request and specification of the authors. Further School level data was matched from the *Censo Escolar* by authors. Details on each dataset and the matching process are discussed below.

First, data on attainment for primary school children was collected from the 2015 and 2016 waves of the *Evaluación Censal de Estudiantes* (ECE), a national assessment of student learning administered in every region of the country to children in all primary schools, both public and private, with at least 5 students registered in the grade evaluated. The assessments are carried out in either Spanish or the native language for students whose first language is not Spanish. Primary school students are assessed in reading and mathematics.[3] For each topic, students are assigned a score based on the dichotomous Rasch model with mean 500 and standard deviation 100, and using cut-off points are categorised in one of three classifications: *satisfactorio*, indicating the child displays the expected ability for this grade and is likely prepared to face the next grade; *en proceso*, indicating only partial achievement of learning outcomes for this grade, but that they are likely still on track to achieve this; and *en inicio*, indicating the child has not displayed expected ability, managing only tasks below what is required for this grade.[4] In this study we refer to these as "at grade", "in progress" and "initial", with scores standardised by grade. Descriptives are provided in Table 1.

For the 2015 wave, only students in the 2nd grade of primary were assessed, while in 2016 students in both the 2nd and 4th grade were assessed. As such our dataset encompasses three consecutive cohorts of students (who, following normal progression, would have been in grades 1, 2, and 3 when JEC was first implemented in 2015), although assessment is conducted at two different stages. Tests were conducted in primary schools towards the end of the school year (March-December), on November 10th-11th in 2015, and in 2016 on November 29th-30th and December 1st-2nd for the 2nd and 4th grades respectively. Coverage of schools and students for surveys is provided in Table A3, with surveys administered to between 94-96.9% of students countrywide in their relevant grades.

As well as test scores, ECE provides basic information on school (school ID code, location, public/privately administered) and child characteristics (gender, age, first language, and parent education level). Summary statistics for the pooled sample are provided in column (1) of Table 2.

Second, Student-level data from ECE is matched to corresponding records in the *Sistema de Información de Apoyo a la Gestión de la Institución Educativa* (SIAGIE), a digital platform developed by the Peruvian ministry of education (MINEDU) which is used by schools to manage student registration, log enrolment and attendance, and record

---

[3]Assessments are also carried out on a wider range of subjects for students in the 2nd grade of secondary school, but for primary students only reading and mathematics are assessed.

[4]for 4th grade students there is an additional category *previo al inicio*, for students below even *en inicio*. To remain comparable across waves and grades, these two categories are merged.

Table 1: Outcome Descriptives

	(1)	(2)	(3)	(4)
		Matched	+ Public	+ 7/8
	All	sibling	+ morning	Sections
Mathematics				
Z-score	-0.00	-0.01	-0.13	-0.18
	(1.00)	(1.00)	(1.01)	(0.98)
Initial or lower	0.29	0.30	0.35	0.36
	(0.46)	(0.46)	(0.48)	(0.48)
In-progress	0.40	0.40	0.39	0.40
	(0.49)	(0.49)	(0.49)	(0.49)
At-grade	0.31	0.30	0.27	0.25
	(0.46)	(0.46)	(0.44)	(0.43)
Reading				
Z-score	-0.00	-0.07	-0.29	-0.35
	(1.00)	(0.98)	(0.96)	(0.92)
Initial or lower	0.15	0.16	0.21	0.20
	(0.36)	(0.36)	(0.40)	(0.40)
In-progress	0.41	0.43	0.47	0.50
	(0.49)	(0.50)	(0.50)	(0.50)
At-grade	0.44	0.41	0.33	0.30
	(0.50)	(0.49)	(0.47)	(0.46)
Observations	1533067	152761	59966	4368

Notes: Column (1) presents the mean scores and proportion of students per score classification for the pooled ECE sample of all students. column (2) provides statistics for those students with at least one matched sibling attending a secondary school in 2015, with column (3) restricted further to only those students with a sibling attending a public morning shift-only school. Column (4) includes only those matched sibling pairs within the window $W = [x_{-1}, c] = [7, 8]$ used for local randomisation analyses, with all previous sample restrictions.

13

Table 2: Summary Statistics

	(1) All	(2) Matched sibling	(3) + Public + morning	(4) + 7/8 Sections
Household characteristics				
Mother completed primary	0.85	0.81	0.70	0.66
	(0.36)	(0.39)	(0.46)	(0.47)
Indigenous language	0.06	0.08	0.16	0.22
	(0.24)	(0.28)	(0.36)	(0.41)
Macro region				
Costa	0.56	0.52	0.37	0.32
	(0.50)	(0.50)	(0.48)	(0.46)
Sierra	0.32	0.37	0.49	0.55
	(0.47)	(0.48)	(0.50)	(0.50)
Selva	0.12	0.11	0.14	0.13
	(0.33)	(0.32)	(0.35)	(0.34)
Younger sibling characteristics				
Public school	0.72	0.82	0.96	0.97
	(0.45)	(0.39)	(0.20)	(0.16)
School is urban	0.87	0.83	0.68	0.69
	(0.34)	(0.38)	(0.47)	(0.46)
Child is female	0.49	0.49	0.49	0.49
	(0.50)	(0.50)	(0.50)	(0.50)
Older sibling characteristics				
Attends JEC school		0.16	0.34	0.39
		(0.37)	(0.47)	(0.49)
Public school		0.77	1.00	1.00
		(0.42)	(0.00)	(0.00)
School is urban		0.83	0.66	0.67
		(0.37)	(0.47)	(0.47)
Child is female		0.50	0.49	0.48
		(0.50)	(0.50)	(0.50)
Observations	1533067	152761	59966	4368

Notes: Column (1) presents summary statistics for the pooled ECE sample of all students. column (2) provides statistics for those students with at least one matched sibling attending a secondary school in 2015, with column (3) restricted further to only those students with a sibling attending a public morning shift-only school. Column (4) includes only those matched sibling pairs within the window $W = [x_{-1}, c] = [7, 8]$ used for local randomisation analyses, with all previous sample restrictions.

student class performance. A common student ID is shared across ECE and SIAGIE, allowing respondents in ECE to be matched to their record. When enrolling students, a parent or guardian must provide their national identity card number (*Documento Nacional de Identidad* or DNI). Using this unique identifier for a parent or guardian, potential siblings can be identified as students who are enrolled under the same parent DNI number. All records of children registered with the same DNI number and who were enrolled in a secondary school in 2015, the year JEC was implemented, were identified and matched by MINEDU.

This method of matching is likely imperfect, for example two siblings would not be matched if an older child was registered by the mother, while the younger child is registered by the father, leading to different associated DNI numbers.[5] We cannot account for this in our dataset, however we expect that those missing due to this limitation are missing at random. Notably, many other studies in this literature also suffer from imperfect matching, and are often based on string-based matching using home address and shared last names (for example: Dustan, 2018; Goodman et al., 2015; Gurantz et al., 2020; Karbownik & Özek, 2023; Nicoletti & Rabe, 2019), which could potentially lead to a greater number of false-positive matches, or based on birth records, which do not allow for matching of siblings who co-reside together but were born to different mothers, or born out of state/country (Dahl et al., 2023; Figlio et al., 2023; Qureshi, 2018a; Zang et al., 2023).

Finally, using publicly available data from the 2013 *Censo Escolar*, a school census which was used by MINEDU personnel to define the original eligibility criteria for JEC, we match school level data on school type (public or private), shift pattern, and the number of sections, which is used to identify and determine treatment status, as well as other predetermined characteristics for testing covariate balance across the discontinuity (see below). These are matched to older siblings using the school code provided within SIAGIE for the institution they were registered to in 2015. Summary statistics for the sub-sample of matched siblings are provided in column (2) of Table 2.

We exclude matches to siblings who are not school age, or who are listed as the same age or younger than the younger sibling. This matching method allows for multiple older siblings attending secondary school to be identified, with 15.8% of matches having more than one older sibling identified. Where applicable, we choose over other matched siblings a sibling who: 1) attends a JEC school (as this indicates the younger sibling is exposed to our treatment, i.e. having at least one older sibling attending a JEC school); or 2) otherwise attends a public, morning shift only school that is not a JEC school. For multiples that remain, we select the first matched sibling. As discussed in subsection 2.3, the JEC

---

[5]Communication with MINEDU staff suggest that this is likely uncommon, as generally mothers register children at school. Our data reflects this, with 78.9% of representatives who register a child being female.

reform was expanded in 2016 and 2017. As inclusion criteria changed for these schools, they are no longer comparable with our initial selection criteria. To avoid potential bias in the estimated effect of treatment, we exclude sibling pairs for which the older sibling attended a school that subsequently joined JEC after 2015. Summary statistics for our analytical sample are provided in column (3) of Table 2.

4 Empirical Strategy

4.1 Local Randomisation

We exploit an initial eligibility criteria, which required schools to have eight or more classes to be eligible for JEC, to assess the potential for spillover effects arising due to exogenous variation in the schooling of an older sibling, on the educational outcomes of a younger sibling. This rule provides a clearly defined threshold at which the conditional probability of being assigned to JEC changes discontinuously (Lee & Lemieux, 2010). As the final selection of schools for the JEC treatment was driven by other unobservable decisions, the discontinuous jump in treatment probability is less than one, resulting in imperfect compliance (Imbens & Lemieux, 2008). We therefore implement a fuzzy regression discontinuity design (RDD), instrumenting participation in JEC with this initial eligibility rule.

All RDDs consist of three fundamental aspects: a score, which all units receive, a cut-off point, and a treatment rule that assigns units to treatment at values above the cutoff point. Conventional inference relies on the assumption that the conditional expectations of potential outcomes given a score are continuous (and differentiable) at the cutoff point, as formalised by Hahn et al. (2001), wherein the difference between the treatment and control average observed outcomes is equal to the average treatment effect in the limit (Cattaneo et al., 2020). This justifies the fitting of local polynomial regressions to approximate the unknown regression functions above and below the cutoff and extrapolating these towards the cutoff point. This framework is referred to as the continuity-based RDD framework. Notably however, this framework assumes that the score variable that determines treatment assignment is a continuous random variable which can take any value. While continuity based approaches are still valid in the presence of discrete values when the number of values is sufficiently large, it can be inadequate when a discrete score variable consists of very few "mass points" (that is that many observations may share the same, relatively few values of a score) or that the mass points are sparse, such that most values are far from the cutoff (Cattaneo et al., 2024).

Our score variable is the number of sections in each school, and represents one such case. A suitable alternative is to consider a local randomisation approach to regression discontinuity design (LRRDD), formalised by Cattaneo et al. (2015) following work by

Lee (2008). Under this framework, the validity of comparisons between treated and control units stems from assuming that, at least for a small window surrounding the cutoff, treatment is "as-if randomly" assigned. While this assumption can be seen as relatively stronger than the continuity assumption (at least when the score variable is continuous), it justifies the use of methods from experimental literature for estimation and inference, interpreting RDD as a local randomised experiment near the cutoff (Cattaneo et al., 2017). This approach is intuitive and closely-aligned with the justifications provided by Thistlethwaite and Campbell (1960) in first introducing RDD. The advantages of this method are that it avoids modelling assumptions, instead relying on assumptions on the assignment mechanism for units near the cutoff, and that inference is robust to the use of discrete score variables with few mass points (Kolesár & Rothe, 2018), indeed as little as one mass point either side of the cutoff (Cattaneo et al., 2024). Furthermore, this approach is easily extendable for the analysis of regression discontinuity designs with imperfect compliance (fuzzy RDD). The basic local randomisation framework for regression discontinuity, extension to fuzzy RD designs, and the required assumptions for identification and estimation, are set out in Appendix B.

Under the relevant assumptions, estimation of the fuzzy LRRDD local average treatment effect can proceed through the standard two-stage least squares procedure. In a first stage, we estimate for a sibling pair i the effect of the cutoff rule, $T_i = \mathbb{1}(X_i \geq 8)$, on the probability receiving treatment as the difference-in-means of those observations just below and just above the cutoff within a small window $W = [c - w, c + w]$ where local randomisation holds:

$$D_i^W = \alpha_0 + \alpha_1 T_i + v_i$$

Where D_i^W is the treatment received by sibling pair i with an older sibling attending a secondary school with a number of sections $X_i \in W$. We restrict our sample to include only pairs with an older sibling attending a morning shift-only public secondary school, aligning with the other initial eligibility criteria as listed in Table A2, which provides better comparison between the treatment and control groups. Given a discrete integer running variable our window is defined as the smallest possible window $W = [x_{-1}, c] = [7, 8]$, consisting of schools with 7 sections and 8, minimising the required extrapolation. Additionally, if local randomisation holds, it must hold within the narrowest window, therefore a window selection procedure, such as that proposed by Cattaneo et al. (2015) is not necessary, however we conduct this as a robustness check in subsection 5.3. Summary statistics for sibling pairs within this window are provided in column (4) of Table 2.

The second stage estimates the local average treatment effect (LATE) for compliers of JEC on younger sibling outcomes Y_i, using the fitted values from the first stage $\widehat{D}_i^W$, again as the difference-in-means between those sibling pairs just below and just above the

cutoff within the same window:

$$Y_i = \beta_0 + \beta_1 \widehat{D}_i^W + \varepsilon_i$$

Where Y_i are our measures of younger sibling educational attainment. Specifically, we assess the grade-standardised score in reading and mathematics, as well as the probability of being classified as "at grade", or as "in progress" or higher. We make the assumption that potential outcomes are independent of the score variable X_i (LR 1 and LR 2 in Appendix B) within the window $W = [7, 8]$, therefore the regression functions are flat and the effects are estimated as the vertical distance between average observed outcomes.[6] Given our relatively large local randomisation sample, inferences rely on standard Gaussian large-sample approximations based on a heteroskedastic-robust covariance estimator of variance (Cattaneo et al., 2024).[7]

4.2   Two-Stage Least Squares

In addition to our local randomisation approach to RDD, we also apply a parametric global polynomial approach, estimating the effect of a sibling pair receiving the JEC treatment, instrumented by the initial eligibility threshold, using conventional two-stage least squares (2SLS). We estimate as a second stage the following:

$$Y_{ij} = \gamma_0 + \gamma_1 \widehat{D_{ij}} + \gamma_2(X_{ij} < 8) + \gamma_3(X_{ij} \geq 8) + \gamma_4 \mathbf{Z_i} + \mu_j$$

Where $\gamma_2(X_{ij} < 8)$ and $\gamma_2(X_{ij} \geq 8)$ are polynomial functions of the running variable X_{ij}, for sibling pair i, for which the older sibling attends school j, interacted with an indicator of being above or below the threshold. Our primary results are estimated using both a linear and quadratic form.[8] $\mathbf{Z_{ij}}$ is a vector of controls: sibling pair characteristics are captured by an indicator of being the same gender and age difference measured in years, as well as mother's level of education; fixed effects for survey year and older sibling's grade are also included. While controls are not required for identification of the RDD effect, they can provide asymptotic efficiency gains in large samples (Cattaneo et al., 2023), however we show in our robustness checks that 2SLS results are robust to estimation without covariates. $\widehat{D_{ij}}$ represents the residuals from the first stage measuring the impact of the cutoff rule $T_{ij} = \mathbb{1}(X_{ij} \geq 8)$ on the probability of receiving treatment, with all other terms included as listed in the second stage. Standard errors μ_{ij} are cluster robust at the

---

[6]While it is possible to relax this assumption (given other less restrictive assumptions (Cattaneo et al., 2017)) to allow for polynomial adjustments as in standard parametric RDD designs, it is neither necessary, nor applicable given our window consists of only one mass point either side of the cutoff.

[7]Alternatively, inference for small samples can be obtained under Fisherian finite-sample methods (Cattaneo et al., 2015).

[8]higher order polynomials are not recommended, as they likely lead to noisy estimates and poor coverage of confidence intervals (Gelman & Imbens, 2019).

level of the older sibling's school j. Cattaneo et al. (2020) note that global polynomial approximations, while providing a good approximation of unknown regression functions overall, likely provide poor approximations of the conditional expectations at the boundary point, and can be influenced by outliers far from the cutoff, leading to unreliable RDD estimates. Therefore, this specification is provided primarily for comparability, and we present the LRRDD results as our primary results.

4.3 Validation

Similar to a conventional IV setting, we require evidence of a non-zero and sufficient first stage. First stage results are reported in Table 3 for our local randomisation RDD effect, as well as for our linear and quadratic 2SLS specifications. A large and statistically significant effect is estimated for all three specifications, with a discontinuous jump in older sibling participation in JEC of between 60.3-68.0%. Graphical evidence of this discontinuity is provided in Figure 2. The relative strength of our first stage relationship can be measured by the effective F-statistic, with Cattaneo et al. (2024) recommending a higher rule-of-thumb threshold for RDD contexts of 20 or more. All three specifications provide evidence of a strong and relevant first stage.

Table 3: First Stage: Participation in JEC

	Local rand.	Two-stage least squares (2SLS)	
	(1)	(2)	(3)
Sections≥8	0.680	0.604	0.661
	[0.000]***	-	-
	-	(0.024)***	(0.036)***
Spline	-	Linear	Quadratic
eff. F-stat.	4307.00	651.32	339.43
N	4368	59192	59192

Notes: * p< 0.10, ** p< 0.05, *** p< 0.01. Local randomisation asymptotic p-values based on heteroskedastic robust standard errors are reported in square brackets. 2SLS Cluster robust standard errors at the older sibling school level are reported in parentheses. Local randomisation results modelled without polynomial adjustment for the smallest possible window. For 2SLS, linear and quadratic splines of the running variable are specified. 2SLS additional covariates include sibling pair age difference in years, an indicator of being the same gender, and mother's educational attainment in years. Fixed effects for older sibling grade and survey year are also included.

Additionally, we assess the validity of our local randomisation RDD estimates by testing for systematic differences between treated and control groups near the cutoff. For the implementation of all validity tests we focus on intention-to-treat effects, as recommended by Cattaneo et al. (2024), as these tests aim to assess how similar observations are just

Figure 2: Discontinuity in Participation in JEC at Cutoff
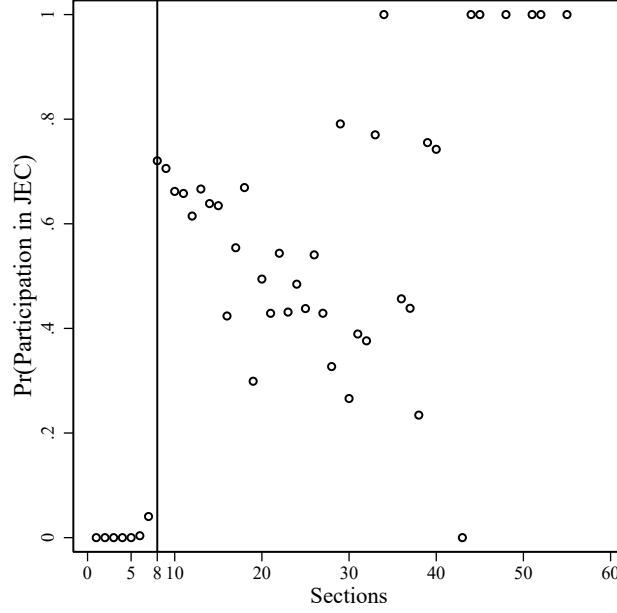
below and above the cutoff, rather than the difference between those who receive treatment and those who do not. Balance tests for a list of predetermined sibling pair and school level characteristics are provided in Table 4. These tests provide evidence of the validity of our LRRDD estimates, as well as suggestive evidence that our assumption of as good as random assignment within the window W holds, although we cannot account for imbalances in unobserved characteristics. Results suggest covariates are well balanced within our window, with non-significant differences estimated for those below and above the cutoff, with exception that sibling pairs are slightly less likely to report speaking an indigenuous language (-3.0% above the threshold) and schools are less likely to be located in districts which eligible for the Crecer welfare program. Balance tests for our 2SLS sample are provided in Table A4, as well as graphical evidence of covariate smoothness in Figure A1 for household/sibling dyad characteristics and school level characteristics in Figure A2.

Additionally, we assess the density of observations. If units lack the ability to control precisely the value of cutoff score, placement of units below and above the cutoff should be as if random in the window W around the cutoff (Cattaneo et al., 2017). Agüero et al. (2021) provide evidence that students did not systematically select in to JEC schools.[9] Table 5 provides the results of a binomial test with null hypothesis $H_0 : Pr(k = 0.5)$. While we reject the null that the success probability is exactly equal to a half, our observed probability (k=0.52) indicates placement above the cutoff is close to that under a simple

---

[9]Additionally, the final list of JEC schools was only published in February 2015 (RM No 062-2015-MINEDU). With the school year running from March to December, it is unlikely that many parents could select in to treatment.

Table 4: Local Randomisation: Predetermined Covariates Balance Test

	School district receives program		Proportion of students		Pass rate
	Juntos	Crecer	Girls	Indigenous	Total
Panel A: School level					
Sections≥8	0.009	-0.053	0.003	-0.004	-0.004
	[0.532]	[0.000]***	[0.260]	[0.754]	[0.359]
Control mean	0.425	0.728	0.471	0.216	0.670
N	4368	4368	4368	4368	4368
	Sibling Dyad			Household	
	Same gender	Age diff.	Grade diff.	Indig. lang.	Parent educ.
Panel B: Sibling dyad level					
Sections≥8	-0.015	0.047	0.041	-0.030	0.006
	[0.314]	[0.415]	[0.366]	[0.017]**	[0.698]
Control mean	0.509	6.533	6.383	0.235	0.662
N	4368	4368	4368	4368	4320

Notes: * p< 0.10, ** p< 0.05, *** p< 0.01. Estimates are the local randomisation intention-to-treat effect for units above the cutoff, within W. Local randomisation asymptotic p-values based on heteroskedastic robust standard errors are reported in square brackets. Results modelled without polynomial adjustment for the smallest possible window.

unbiased coin flip, suggesting sorting is unlikely near the cutoff. This result does not imply that the local randomisation assumptions are violated, and results are consistent with the fact that we would naturally expect slightly more students to be placed in 8 sections schools, given the additional class likely meaning more students per school on average, and with an assignment process using a bernoulli trail with a true probability of success k slightly more than one half (Cattaneo et al., 2024).[10] We carry out further RDD validation and falsification checks as part of our robustness checks in subsection 5.3.

Table 5: Local Randomisation: Binomial Test

	Observations below cutoff	Observations above cutoff	Observed probability	p-value
$H_0 : Pr(k = 0.5)$	2103	2265	0.52	0.015

Notes: Expected probability $Pr = 0.5$. Sample includes sibling pairs for which the focal child attends a public secondary school with morning only shift and either 7 (below cutoff) or 8 sections (above cutoff).

---

[10]The conventional McCrary (2008) density test used in the continuity framework for non-parametric and parametric RDD is not suitable for discrete running variables with few mass points, providing misleading results (Frandsen, 2017). Therefore we do not provide an equivalent test for 2SLS results.

5 Results

5.1 Main Results

LRRDD Results are presented in Table 6. columns (1) and (4) provide estimates of the spillover effect for compliers of JEC (the LATE) on younger sibling scores, with effects of 0.120 S.D. and 0.135 S.D. estimated for reading and mathematics. Estimates for the probability of being graded as "at grade" (the highest group) are reported in columns (2) and (5), with estimated effects on the probability of being "in progress" or above reported in columns (3) and (6), for reading and mathematics respectively. While our estimates are positive, the effect sizes for reading grade probability are close to null, suggesting that although the spillover has a relatively large effect on the intensive margin, the effect on younger sibling reading does not have a significant impact at the extensive margin. The control mean score in reading for our analytical sample is -0.393 S.D. lower than the average calculated for all students in their grade, with siblings in our sample skewing towards the lower end of the distribution of attainment. For mathematics, while the effect estimated for the highest group ("at grade") is not statistically significant at conventional levels, an increase of 5.4 p.p. is estimated for the probability of being classed as "in progress" and above, significant at the 5% level. This effect at the extensive margin is modest, representing an 8% increase from the control mean, however performance in mathematics is generally poorer than reading for all students in ECE (see Table 1), therefore this represents a tangible improvement for treated siblings in our sample.

Table 6: Effects of JEC on Younger Sibling Outcomes: Local Randomisation Inference

	Reading			Mathematics		
	(1) Z-score	(2) At grade	(3) ≥ In-prog.	(4) Z-score	(5) At grade	(6) ≥ In-prog.
Sections≥8	0.120 [0.003]***	0.006 [0.760]	0.011 [0.530]	0.135 [0.002]***	0.030 [0.124]	0.054 [0.011]**
Control mean	-0.393	0.293	0.795	-0.226	0.238	0.626
N	4368	4368	4368	4368	4368	4368

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Estimates are of the local average treatment effect (LATE) above the cutoff, within W. Local randomisation asymptotic p-values based on heteroskedastic robust standard errors are reported in square brackets. Results modelled without polynomial adjustment for the smallest possible window.

That we see a stronger impact on younger sibling mathematics than on reading, in particular at the extensive margin, is consistent with the effects found by (Agüero et al., 2021), who find a larger and more robust effect on targeted child mathematics, while impacts for reading are smaller in magnitude and less robust, likely reflecting that students attending JEC schools benefitted from a 2-hour increase in mathematics, compared with a

1-hour increase in reading. This may indicate that our spillover effects operate through the direct within-family peer effect channel, with older siblings seeing greater improvements in their own mathematics ability, and therefore likely providing higher quality help to younger siblings (e.g. if they help their younger sibling with homework, or with studying generally).

Reduced form intention-to-treat effects are provided in Table A5, with findings consistent with the fuzzy RDD estimates, although smaller in magnitude, reflecting downwards bias as a result of non-compliance. Additionally we re-estimate our main results using a parametric 2SLS specification as discussed above; results based on a linear and quadratic functional form are presented in Table 7, in panels A and B respectively. Findings under the linear specification are consistent with our primary LRRDD results: the older sibling's attendance of a JEC school is associated with a positive spillover effect on younger sibling reading and mathematics scores of 0.159 S.D. and 0.135 S.D.. In contrast with the LRRDD effects, while there is a positive effect estimated for both the probability of being "at grade" and for being "in progress" or higher, the effect for "at grade" is large (7.3 p.p. and 6.1 p.p.) and significant at the 1% level, while the effects for "in progress" and higher are no longer significant at conventional levels, although they remain positive (3.2 p.p. and 4.3 p.p., respectively).[11]

Table 7: Effects of JEC on Younger Sibling Outcomes: 2SLS

	Reading			Mathematics		
	(1) Z-score	(2) At grade	(3) ≥ In-prog.	(4) Z-score	(5) At grade	(6) ≥ In-prog.
Panel A: Linear specification						
Sections≥8	0.159	0.073	0.032	0.135	0.061	0.043
	(0.059)***	(0.024)***	(0.022)	(0.063)**	(0.023)***	(0.028)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
N	59192	59192	59192	59192	59192	59192
Panel B: Quadratic specification						
Sections≥8	-0.056	-0.038	-0.011	-0.057	-0.013	-0.035
	(0.095)	(0.041)	(0.036)	(0.099)	(0.039)	(0.045)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
N	59192	59192	59192	59192	59192	59192

Notes: * p< 0.10, ** p< 0.05, *** p< 0.01. Estimates are of the local average treatment effect (LATE) at the cutoff, including a linear (panel A) and quadratic (panel B) spline of the running variable. Cluster robust standard errors at the older sibling school level are reported in parentheses. Additional covariates include sibling pair age difference in years, an indicator of being the same gender, and mother's educational attainment in years. Fixed effects for older sibling grade and survey year are also included.

Interestingly, estimates under the quadratic form are estimated as slightly negative, although none are statistically significant, with large standard errors greater in magni-

---

[11]graphical evidence of the discontinuity in outcomes at the cutoff, based on a linear fit, is presented in Figure A3.

tude than the effect size. This suggests the global quadratic specification likely provides imprecise estimates of our effect. As previously noted, global parametric specifications tend to provide a poor approximation of the RDD effect at the boundary point (Cattaneo et al., 2020), in particular those of higher order which can provide noisy estimates and poor coverage of confidence intervals (Gelman & Imbens, 2019). Although the estimates, at least under a linear spline, are consistent with our LRRDD results, they may also be subject to the effects of large outliers, particularly far away from the cutoff; therefore we consider our findings under local randomisation to be more robust. An additional concern is the inclusion of so-called "bad controls", which may bias the effect upwards from zero. Given the inclusion of additional covariates is not required for the identification of RDD effects, we re-estimate our 2SLS results without the inclusion of controls in Table A6, and find our estimates remain robust, although with slightly reduced magnitude and increased variance, as expected (Cattaneo et al., 2023; Noack et al., 2023).

5.2 Heterogeneity

In this section, we consider the potential for heterogeneity in effects across sibling pair gender mix. First we assess differences in the spillover effect across the gender of the younger sibling. Results are presented in columns (1) and (4) of Table 8. A large spillover effect is estimated for younger sisters of compliers of JEC, with smaller non-significant effects estimated for younger brothers. Although our results are consistent with the literature documenting gender differences in sensitivity to family inputs (Autor et al., 2019), indicating that girls respond more to their older sibling's increased schooling than boys, and with the average female advantage seen in education (Autor et al., 2016, 2023) generally, our effects are the opposite of those found by Figlio et al. (2023), who find spillover effects of an older sibling being flagged for 3rd grade retention are concentrated amongst younger brothers, rather than sisters.

Similar to other work in this literature, we further assess if there are heterogeneities based on the sibling-pair gender mix, which may reveal the role of older sibling gender or gender matching/difference in the transmission of spillover effects. In general, evidence for gender differences in sibling spillovers is mixed (Dahl et al., 2023; Steelman et al., 2002). However, recent studies find stronger effects for same-gender siblings (Karbownik & Özek, 2023; Nicoletti & Rabe, 2019; Qureshi, 2018a; Zang et al., 2023) compared with mixed-gender sibling pairs, although some studies find mixed gender effects, such as (Qureshi, 2018b) who shows in Pakistan that older sister education matters for younger brothers educational outcomes.[12] Furthermore in a multi-country study, Altmejd et al. (2021) find that the importance of gender composition varies across countries.

---

[12]They do not estimate effects on younger sisters, as they generally attend the same schools as their older sister.

Table 8: Heterogeneous Effects of JEC on Younger Sibling Outcomes: By Gender Mix

	Female			Male		
	(1)	(2) Older sister	(3) Older brother	(4)	(5) Older sister	(6) Older brother
	All			All		
Reading						
Z-score	0.189	0.216	0.163	0.057	0.025	0.076
	[0.002]***	[0.015]**	[0.050]*	[0.297]	[0.754]	[0.320]
At grade	0.023	0.048	0.001	-0.009	-0.024	-0.000
	[0.438]	[0.274]	[0.986]	[0.751]	[0.559]	[0.997]
≥ In prog.	0.053	0.025	0.081	-0.027	-0.041	-0.018
	[0.041]**	[0.525]	[0.023]**	[0.275]	[0.237]	[0.613]
Mathematics						
Z-score	0.227	0.315	0.146	0.049	0.051	0.037
	[0.000]***	[0.000]***	[0.095]*	[0.416]	[0.553]	[0.650]
At grade	0.055	0.077	0.035	0.005	0.012	-0.003
	[0.044]**	[0.050]*	[0.354]	[0.851]	[0.755]	[0.944]
≥ In prog.	0.105	0.130	0.082	0.007	-0.008	0.018
	[0.001]***	[0.005]***	[0.058]*	[0.798]	[0.850]	[0.649]
N	2136	1033	1103	2232	1070	1162

Notes: * p< 0.10, ** p< 0.05, *** p< 0.01. Estimates are of the local average treatment effect (LATE) above the cutoff, within W. Local randomisation asymptotic p-values based on heteroskedastic robust standard errors are reported in square brackets. Results modelled without polynomial adjustment for the smallest possible window.

Breaking down same and mixed gender effects by specific genders, results suggest that brother pairs are more likely to make similar choices in college or majors (Dahl et al., 2023; Joensen & Nielsen, 2018), suggesting boys are more likely to be influenced by their older brother's choices, but effects on younger sibling attainment are less clear (Karbownik & Özek, 2023; Zang et al., 2023). Columns (2) and (5) show the effects of having an older sister attend a JEC school, for girls and boys respectively. Likewise, columns (3) and (6) show the effects for having an older brother. As with our analysis above, results are concentrated amongst female younger siblings. Positive effects are estimated for female younger siblings regardless of having an older sister or brother, however the magnitude of effect sizes estimated for sister pairs are larger than those for the mixed younger sister-older brother pairs, suggesting that while girls generally respond more to spillovers, the response is greatest for same gender spillovers from their older sister. This would be consistent with both the female advantage in sensitivity to family inputs, as well as the evidence that same-gender pairs are more responsive, likely due to greater direct interaction, a higher likelihood to share similar interests, and to experience stronger role model effects (Benin & Johnson, 1984). Results estimated under the linear form for two-stage least squares are reported in Table A7, and are consistent with the findings under our primary specification.

Another potentially important characteristic for heterogeneities is sibling age or grade-spacing, with effects generally larger for siblings who are close in age or grade (Dahl et al., 2023; Figlio et al., 2023; Zang et al., 2023) Closely spaced siblings are more likely to share peer networks and interact more, particularly if they attended the same primary school together. Given our policy reform and grade specific testing, we see limited variation in age-spacing in our sample. Additionally, as our dataset provides a snapshot, we do not observe siblings' educational histories, therefore we are unable to identify if students attended primary school together for any period of time prior to 2015, or if they attended the same primary school.[13] Therefore we are unable to accurately address the potential for heterogeneity by age or grade difference.

5.3 Robustness

This section presents additional analyses to test the robustness and validity of our primary findings. We provide three falsification tests. First, we assess if our estimated effects are spurious by conducting a placebo test with an alternative group of sibling pairs which were not eligible for the JEC reform, where the older sibling attends a private school. It is expected that there should be no systematic differences in outcomes between those units within the window below and above the cutoff. Results are presented in Table A8, with no significant effects estimated across our outcomes.

---

[13]Grade progression in Peru is less clear-cut than in high-income contexts, with almost 30% of Peruvian respondents to the 2009 PISA reporting having repeated at least one grade (OECD, 2011).

Second, we assess the potential for spurious results by selecting placebo cutoffs where the probability of treatment assignment is not expected to change, keeping our window width the same (Imbens & Lemieux, 2008). We select one placebo threshold below the real cutoff at 4 sections, with the window $W = [3, 4]$, and one above with $W = [11, 12]$. These cutoffs are selected to avoid contamination from the actual treatment effect, such that no observations from the original window are present. Additionally, we choose these windows as they include similar numbers of observations either side of the cutoff, whereas windows including either 5 or 10 sections may show systematic differences as these are the most common mass points (see Figure 1). Results are presented for both placebo cutoffs in Table A9, with no significant effects estimated, with exception of the positive effect for "at grade" reading in the $W = [11, 12]$ window (2.9 p.p.), significant at the 5% level, against a null effect in our main results. These tests together provide good evidence that our results at the actual threshold are unlikely to be spurious.

Third, we assess the robustness of our results to the choice of window. Following Cattaneo et al. (2015), we conduct a data-driven window selection procedure, based on the balance of a set of pre-determined covariates, Z_i, which iteratively tests the null hypothesis, $H_0 : Z_i(1) = Z_i(0) \quad \forall i$, within a widening window. We set a conservative threshold significance level for rejection, set at $\alpha \geq 0.15$ as recommended by Cattaneo et al. (2024), tolerating a higher probability of type I error to lower the chance of failing to reject a false null of balanced covariates.[14] Table A10 displays the results for 5 increasingly widening windows, beginning with the original narrowest window, with a minimum step of 1 between windows, given our discrete running variable. The second and third columns provide the name and p-value of the variable with the minimum difference-in-means p-value within that window. the suggested window in this case is $W = [x_{-2}, c_{+1}] = [6, 9]$, after which the minimum p-value falls below our significance level threshold. Table A11 presents results for our main outcomes, re-estimated using this window, showing results consistent with those estimated in our primary specification.

Our analysis uses data collected in the 2015 and 2016 waves of ECE. Unfortunately, due to disruptions and damage caused by flooding and heavy rains related to the 2017 El Niño Southern Oscillation (ENSO), the 2017 ECE was not conducted. Furthermore the 2019 ECE survey was conducted only in secondary schools, with further planned surveys disrupted by the COVID-19 pandemic. An ECE survey was conducted in primary schools for 2018 amongst 4th grade students only. Given our sample includes 2nd grade students tested in 2016, it is likely that the 2018 survey consists of a large overlap of respondents, however in our anonymised datasets we are unable to identify students across surveys. Therefore it is not possible to account for potential repeated observation with a panel data

---

[14]As we have already shown two of our pre-determined covariates are not fully balanced in Table 4, these are excluded from our set Z_i. We therefore use this window only for testing the sensitivity of results, with all other inference being based on the original narrowest window.

structure in our analysis if we were to include the 2018 survey. Additionally, the disruption of the 2017 ENSO, the worst to hit Peru since 1925 (Ramírez & Briones, 2017), may have impacted learning outcomes significantly, overtaking the effect of JEC. Therefore we did not consider the 2018 ECE survey as suitable for our primary analysis. However as a robustness check, we re-run our primary analysis on a pooled sample including the 2018 survey wave in Table A12, with consistent results estimated, although slightly reduced in magnitude.

Finally, we estimate our results separately by survey year and wave, to assess if effects are similar across different groups. Results are presented in Table A13. While generally consistent, relatively larger effect sizes are estimated for the 2015 2nd grade group, as well as for the 2016 4th grade group (who would generally be in 3rd grade in 2015), while smaller effects are estimated for the 2016 2nd grade group (1st grade in 2015), which are not significant at conventional levels. this may indicate that the spillover effect is greater for older children who are, on average, closer in age to their older siblings at the time of the reform in 2015.

6   Discussion and Conclusions

This study documents the potential for sibling spillover effects from a school day extension reform in Peru. Specifically, we exploit an arbitrary cutoff rule based on the number of classes (sections), which was used to define the selection of schools into the JEC program, to provide exogenous variation in the amount and quality of schooling experienced by an older sibling. We measure the impact of having an older sibling attending a school just above this cutoff on the reading and mathematics scores of younger siblings who are not yet eligible for the program. Using a local randomisation approach to regression discontinuity design we estimate the effect for compliers above the cutoff, finding evidence of a positive spillover effect on both standardised reading and mathematics scores of having an older sibling attend a JEC school, compared with those below the cutoff. In addition to impacts on scores, we also find some evidence of a spillover effect on the extensive margin, with a positive effect on the probability of being categorised as "in progress" or above for mathematics. However, we find no significant effect at the extensive margin for reading. Assessing heterogeneity in spillover effects, we find that effects are concentrated amongst girls, with the largest effect found for older sister-younger sister sibling pairs. Our results are robust to a range of falsification tests, indicating that estimates are valid and unlikely to be spurious.

There are however, several limitations that cannot be addressed by this study. While our novel matching strategy allows us to link siblings across schools to assess the impacts on educational outcomes, our dataset does not allow for an in-depth exploration of the specific mechanisms through which these spillovers are transmitted between siblings. We

28

are also unable to conduct detailed analysis of how spillover effects may differ across household background characteristics, and we can only provide limited evidence of how effects differ across sibling pair characteristics. Additionally, given data limitations, we cannot produce robust evidence of how persistent effects are for siblings as they age. Finally, we focus primarily on the potential for sibling spillovers for younger siblings, in particular siblings who were still attending primary school, and therefore too young to be exposed to JEC. However, we are unable to provide any insight into the potential implications for the older siblings of the targeted child, or on other family members, who may be positively or negatively impacted by the focal child's increased time spent in school.

This study contributes to the growing literature for sibling spillovers in education, in particular, it expands the very limited literature assessing educational spillovers outside of high income contexts. Additionally, to our best knowledge, we provide the first evidence of sibling spillovers arising from a school day extension policy. We therefore also contribute to the literature on the efficacy of school day extension policies. Specifically, it provides evidence that there are positive spillover effects of the JEC reform, beyond the previously established impacts for the targeted child. Importantly, our findings of significant spillovers between siblings suggest that policy evaluations which fail to account for such externalities within families will likely provide misleading conclusions. It is therefore salient that policymakers consider the potential for spillovers when determining the benefits and costs of an educational reform.

References

Agüero, J. (2016). *Evaluación de impacto de la jornada escolar completa*. GRADE and FORGE.

Agüero, J., Favara, M., Porter, C., & Sanchez, A. (2021). *Do More School Resources Increase Learning Outcomes? Evidence from an Extended School-Day Reform*. https://doi.org/10.2139/ssrn.3818651

Aguirre, J., & Matta, J. (2021). Walking in your footsteps: Sibling spillovers in higher education choices. *Economics of Education Review*, *80*, 102062. https://doi.org/10.1016/j.econedurev.2020.102062

Alcázar, L. (2016). *Evaluación del diseño y proceso de implementación del modelo de jornada escolar completa (jec) para educación secundaria a nivel nacional. informe final*. GRADE and FORGE.

Almeida, R., Bresolin, A., Borges, B., Mendes, K., & Menezes-Filho, N. (2016). *Assessing the Impacts of Mais Educação on Educational Outcomes: Evidence between 2007 and 2011* (No. 7644). The World Bank. https://doi.org/10.1596/1813-9450-7644

Altmejd, A., Barrios-Fernández, A., Drlje, M., Goodman, J., Hurwitz, M., Kovac, D., Mulhern, C., Neilson, C., & Smith, J. (2021). O Brother, Where Start Thou? Sibling Spillovers on College and Major Choice in Four Countries. *The Quarterly Journal of Economics*, *136*(3), 1831–1886. https://doi.org/10.1093/qje/qjab006

Autor, D., Figlio, D., Karbownik, K., Roth, J., & Wasserman, M. (2016). School Quality and the Gender Gap in Educational Achievement. *American Economic Review*, *106*(5), 289–295. https://doi.org/10.1257/aer.p20161074

Autor, D., Figlio, D., Karbownik, K., Roth, J., & Wasserman, M. (2019). Family Disadvantage and the Gender Gap in Behavioral and Educational Outcomes. *American Economic Journal: Applied Economics*, *11*(3), 338–381. https://doi.org/10.1257/app.20170571

Autor, D., Figlio, D., Karbownik, K., Roth, J., & Wasserman, M. (2023). Males at the Tails: How Socioeconomic Status Shapes the Gender Gap. *The Economic Journal*, *133*(656), 3136–3152. https://doi.org/10.1093/ej/uead069

Barrios-Fernández, A., & Bovini, G. (2021). It's time to learn: School institutions and returns to instruction time. *Economics of Education Review*, *80*, 102068. https://doi.org/10.1016/j.econedurev.2020.102068

Becker, G. S., & Tomes, N. (1986). Human Capital and the Rise and Fall of Families. *Journal of Labor Economics*, *4*(3), S1–S39.

Bedard, K., & Dhuey, E. (2006). The Persistence of Early Childhood Maturity: International Evidence of Long-Run Age Effects. *The Quarterly Journal of Economics*, *121*(4), 1437–1472. https://doi.org/10.1093/qje/121.4.1437

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. https://doi.org/10.1016/j.econedurev.2009.01.008

Benin, M. H., & Johnson, D. R. (1984). Sibling Similarities in Educational Attainment: A Comparison of Like-Sex and Cross-Sex Sibling Pairs. *Sociology of Education*, *57*(1), 11–21. https://doi.org/10.2307/2112464

Ben-Porath, Y. (1967). The Production of Human Capital and the Life Cycle of Earnings. *Journal of Political Economy*, *75*(4), 352–365.

Berry, J., Dizon-Ross, R., & Jagnani, M. (2020). *Not Playing Favorites: An Experiment on Parental Fairness Preferences*. https://doi.org/10.3386/w26732

Black, S. E., Breining, S., Figlio, D. N., Guryan, J., Karbownik, K., Nielsen, H. S., Roth, J., & Simonsen, M. (2021). Sibling Spillovers. *The Economic Journal*, *131*(633), 101–128. https://doi.org/10.1093/ej/ueaa074

Black, S. E., & Devereux, P. J. (2011). Recent Developments in Intergenerational Mobility. In D. Card & O. Ashenfelter (Eds.), *Handbook of Labor Economics* (pp. 1487–1541, Vol. 4). Elsevier. https://doi.org/10.1016/S0169-7218(11)02414-2

Breining, S. N. (2014). The presence of ADHD: Spillovers between siblings. *Economics Letters*, *124*(3), 469–473. https://doi.org/10.1016/j.econlet.2014.07.010

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. https://doi.org/10.1080/00220388.2018.1563680

Carroll, J. B. (1963). A Model of School Learning. *Teachers college record*, *64*(8), 1–9.

Cattaneo, M. D., Frandsen, B. R., & Titiunik, R. (2015). Randomization Inference in the Regression Discontinuity Design: An Application to Party Advantages in the U.S. Senate. *Journal of Causal Inference*, *3*(1), 1–24. https://doi.org/10.1515/jci-2013-0010

Cattaneo, M. D., Idrobo, N., & Titiunik, R. (2024). *A Practical Introduction to Regression Discontinuity Designs: Extensions*. Cambridge University Press.

Cattaneo, M. D., Idrobo, N., & Titiunik, R. (2020). *A Practical Introduction to Regression Discontinuity Designs: Foundations*. Cambridge University Press.

Cattaneo, M. D., Keele, L., & Titiunik, R. (2023). Covariate Adjustment in Regression Discontinuity Designs. In J. R. Zubizarreta, E. A. Stuart, D. S. Small, & P. R. Rosenbaum (Eds.), *Handbook of Matching and Weighting in Causal Inference* (pp. 153–168). Chapman and Hall.

Cattaneo, M. D., Titiunik, R., & Vazquez-Bare, G. (2017). Comparing Inference Approaches for RD Designs: A Reexamination of the Effect of Head Start on Child Mortality. *Journal of Policy Analysis and Management*, *36*(3), 643–681. https://doi.org/10.1002/pam.21985

Cerdan-Infantes, P., & Vermeersch, C. (2007). *More Time Is Better : An Evaluation Of The Full Time School Program In Uruguay* (No. 4167). The World Bank. https://doi.org/10.1596/1813-9450-4167

Dahl, G. B., Rooth, D.-O., & Stenberg, A. (2023). *Intergenerational and Sibling Spillovers in High School Majors*.

Dustan, A. (2018). Family networks and school choice. *Journal of Development Economics*, *134*, 372–391. https://doi.org/10.1016/j.jdeveco.2018.06.004

Edo, M., & Nistal, M. (2022). *More time less time? The effect of lengthening the school day on learning trajectories*.

Ersoy, E., & Forshaw, R. (2023). *The Effects of Longer Secondary School Days on Parents' Labour Market Outcomes: Evidence from the Jornada Escolar Completa Reform in Peru*. https://doi.org/10.2139/ssrn.4559716

Escobar, X., & Sanchez Castro, A. L. (2021). *Evaluación de diseño e implementación de la Jornada Escolar Completa: Lecciones aprendidas del modelo de jornada escolar ampliado en escuelas públicas*. Ministerio de Educación.

Fan, W., & Porter, C. (2020). Reinforcement or compensation? Parental responses to children's revealed human capital levels. *Journal of Population Economics*, *33*(1), 233–270. https://doi.org/10.1007/s00148-019-00752-7

Figlio, D., Holden, K. L., & Ozek, U. (2018). Do students benefit from longer school days? Regression discontinuity evidence from Florida's additional hour of literacy

instruction. *Economics of Education Review*, *67*, 171–183. https://doi.org/10.1016/j.econedurev.2018.06.003

Figlio, D. N., Karbownik, K., & Özek, U. (2023). *Sibling Spillovers May Enhance the Efficacy of Targeted School Policies.* https://doi.org/10.3386/w31406

Frandsen, B. R. (2017). Party Bias in Union Representation Elections: Testing for Manipulation in the Regression Discontinuity Design when the Running Variable is Discrete. In *Regression discontinuity designs: Theory and applications.* (pp. 281–315). Emerald Publishing Limited.

Garganta, S., Pinto, M. F., & Zentner, J. (2022). *Extended School Day and Teenage Fertility in Dominican Republic.* Inter-American Development Bank. https://doi.org/10.18235/0004496

Gelman, A., & Imbens, G. (2019). Why High-Order Polynomials Should Not Be Used in Regression Discontinuity Designs. *Journal of Business & Economic Statistics*, *37*(3), 447–456. https://doi.org/10.1080/07350015.2017.1366909

Goodman, J., Hurwitz, M., Smith, J., & Fox, J. (2015). The relationship between siblings' college choices: Evidence from one million SAT-taking families. *Economics of Education Review*, *48*, 75–85. https://doi.org/10.1016/j.econedurev.2015.05.006

Grätz, M., & Torche, F. (2016). Compensation or Reinforcement? The Stratification of Parental Responses to Children's Early Ability. *Demography*, *53*(6), 1883–1904. https://doi.org/10.1007/s13524-016-0527-1

Gurantz, O., Hurwitz, M., & Smith, J. (2020). Sibling effects on high school exam taking and performance. *Journal of Economic Behavior & Organization*, *178*, 534–549. https://doi.org/10.1016/j.jebo.2020.07.026

Hahn, J., Todd, P., & Van der Klaauw, W. (2001). Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design. *Econometrica*, *69*(1), 201–209.

Hidalgo Arestegui, A. (2021). *Long term effects of an extended school-day programme: Evidence from a longitudinal study in Peru.*

Imbens, G. W., & Angrist, J. D. (1994). Identification and Estimation of Local Average Treatment Effects. *Econometrica*, *62*(2), 467–475. https://doi.org/10.2307/2951620

Imbens, G. W., & Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics*, *142*(2), 615–635. https://doi.org/10.1016/j.jeconom.2007.05.001

Joensen, J. S., & Nielsen, H. S. (2018). Spillovers in education choice. *Journal of Public Economics*, *157*, 158–183. https://doi.org/10.1016/j.jpubeco.2017.10.006

Karbownik, K., & Özek, U. (2023). Setting a Good Example?: Examining Sibling Spillovers in Educational Achievement Using a Regression Discontinuity Design. *Journal of Human Resources*, *58*(5), 1567–1607. https://doi.org/10.3368/jhr.58.5.0220-10740R1

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

Kolesár, M., & Rothe, C. (2018). Inference in Regression Discontinuity Designs with a Discrete Running Variable. *American Economic Review*, *108*(8), 2277–2304. https://doi.org/10.1257/aer.20160945

Lee, D. S. (2008). Randomized experiments from non-random selection in U.S. House elections. *Journal of Econometrics*, *142*(2), 675–697. https://doi.org/10.1016/j.jeconom.2007.05.004

Lee, D. S., & Lemieux, T. (2010). Regression Discontinuity Designs in Economics. *Journal of Economic Literature*, *48*(2), 281–355. https://doi.org/10.1257/jel.48.2.281

Llach, J., Adrogué, C., Gigaglia, M., & Orgales, C. R. (2009). Do Longer School Days Have Enduring Educational, Occupational, or Income Effects? A Natural Experiment in Buenos Aires, Argentina. *Economía*, *10*(1), 1–43.

Manski, C. F. (1993). Identification of Endogenous Social Effects: The Reflection Problem. *The Review of Economic Studies*, *60*(3), 531–542. https://doi.org/10.2307/2298123

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. https://doi.org/10.1093/qje/qjz010

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, *142*(2), 698–714. https://doi.org/10.1016/j.jeconom.2007.05.005

McEwan, P. J., & Shapiro, J. S. (2008). The Benefits of Delayed Primary School Enrollment: Discontinuity Estimates Using Exact Birth Dates. *Journal of Human Resources*, *43*(1), 1–29. https://doi.org/10.3368/jhr.43.1.1

Nicoletti, C., & Rabe, B. (2019). Sibling spillover effects in school achievement. *Journal of Applied Econometrics*, *34*(4), 482–501. https://doi.org/10.1002/jae.2674

Noack, C., Olma, T., & Rothe, C. (2023). *Flexible Covariate Adjustments in Regression Discontinuity Designs*. https://doi.org/10.48550/arXiv.2107.07942

OECD. (2011). When Students Repeat Grades or Are Transferred Out of School.

Oettinger, G. S. (2000). Sibling Similarity in High School Graduation Outcomes: Causal Interdependency or Unobserved Heterogeneity? *Southern Economic Journal*, *66*(3), 631–648. https://doi.org/10.2307/1061429

Ortega, J. (2018). Dentro del colegio y lejos del embarazo: El efecto de la Jornada Escolar Completa (JEC) sobre el embarazo adolescente en el Perú. *Investigaciones*.

Pablo, A., Peter, H., & Evans, D. (2015, June 16). *Extending the School Day in Latin America and the Caribbean*.

Patall, E. A., Cooper, H., & Allen, A. B. (2010). Extending the School Day or School Year: A Systematic Review of Research (1985-2009). *Review of Educational Research*, *80*(3), 401–436.

Persson, P., Qiu, X., & Rossin-Slater, M. (2021). *Family Spillover Effects of Marginal Diagnoses: The Case of ADHD*. https://doi.org/10.3386/w28334

Pitt, M. M., Rosenzweig, M. R., & Hassan, M. N. (1990). Productivity, Health, and Inequality in the Intrahousehold Distribution of Food in Low-Income Countries. *The American Economic Review*, *80*(5), 1139–1156.

Qureshi, J. A. (2018a). Siblings, Teachers, and Spillovers on Academic Achievement. *Journal of Human Resources*, *53*(1), 272–297.

Qureshi, J. A. (2018b). Additional Returns to Investing in Girls' Education: Impact on Younger Sibling Human Capital. *The Economic Journal*, *128*(616), 3285–3319. https://doi.org/10.1111/ecoj.12571

Ramírez, I. J., & Briones, F. (2017). Understanding the El Niño Costero of 2017: The Definition Problem and Challenges of Climate Forecasting and Disaster Responses.

International Journal of Disaster Risk Science, 8(4), 489–492. https://doi.org/10.1007/s13753-017-0151-8

Rodrigues, M., & Campos Flores, A. (2021). *Evaluación de impacto de la Jornada Escolar Completa: Ampliación de la jornada escolar en escuelas públicas.* Ministerio de Educación.

Saavedra, J., & Gutierrez, M. (2020). Peru: A Wholesale Reform Fueled by an Obsession with Learning and Equity. In F. M. Reimers (Ed.), *Audacious Education Purposes: How Governments Transform the Goals of Education Systems* (pp. 153–180). Springer International Publishing. https://doi.org/10.1007/978-3-030-41882-3_6

Sacerdote, B. (2014). Experimental and Quasi-Experimental Analysis of Peer Effects: Two Steps Forward? *Annual Review of Economics, 6*(1), 253–272. https://doi.org/10.1146/annurev-economics-071813-104217

Sanchez, A., & Favara, M. (2019). Consequences of Teenage Childbearing in Peru. *Young Lives Working Paper, 185.*

Sekhon, J. S., & Titiunik, R. (2017). On interpreting the regression discontinuity design as a local experiment. *Advances in Econometrics, 38*, 1–28. https://doi.org/10.1108/S0731-905320170000038001

Steelman, L. C., Powell, B., Werum, R., & Carter, S. (2002). Reconsidering the Effects of Sibling Configuration: Recent Advances and Challenges. *Annual Review of Sociology, 28*, 243–269. https://doi.org/10.1146/annurev.soc.28.111301.093304

Thistlethwaite, D. L., & Campbell, D. T. (1960). Regression-discontinuity analysis: An alternative to the ex post facto experiment. *Journal of Educational Psychology, 51*(6), 309–317. https://doi.org/10.1037/h0044319

Todd, P. E., & Wolpin, K. I. (2003). On the Specification and Estimation of the Production Function for Cognitive Achievement. *The Economic Journal, 113*(485), F3–F33. https://doi.org/10.1111/1468-0297.00097

Yi, J., Heckman, J. J., Zhang, J., & Conti, G. (2015). Early Health Shocks, Intra-Household Resource Allocation and Child Outcomes. *The Economic Journal, 125*(588), F347–F371. https://doi.org/10.1111/ecoj.12291

Zang, E., Tan, P. L., & Cook, P. J. (2023). Sibling Spillovers: Having an Academically Successful Older Sibling May Be More Important for Children in Disadvantaged Families. *American Journal of Sociology, 128*(5), 1529–1571. https://doi.org/10.1086/724723

Appendices

A   Additional tables and figures

Figure A1: Covariate Smoothness Across Cutoff: Sibling-Pair Characteristics

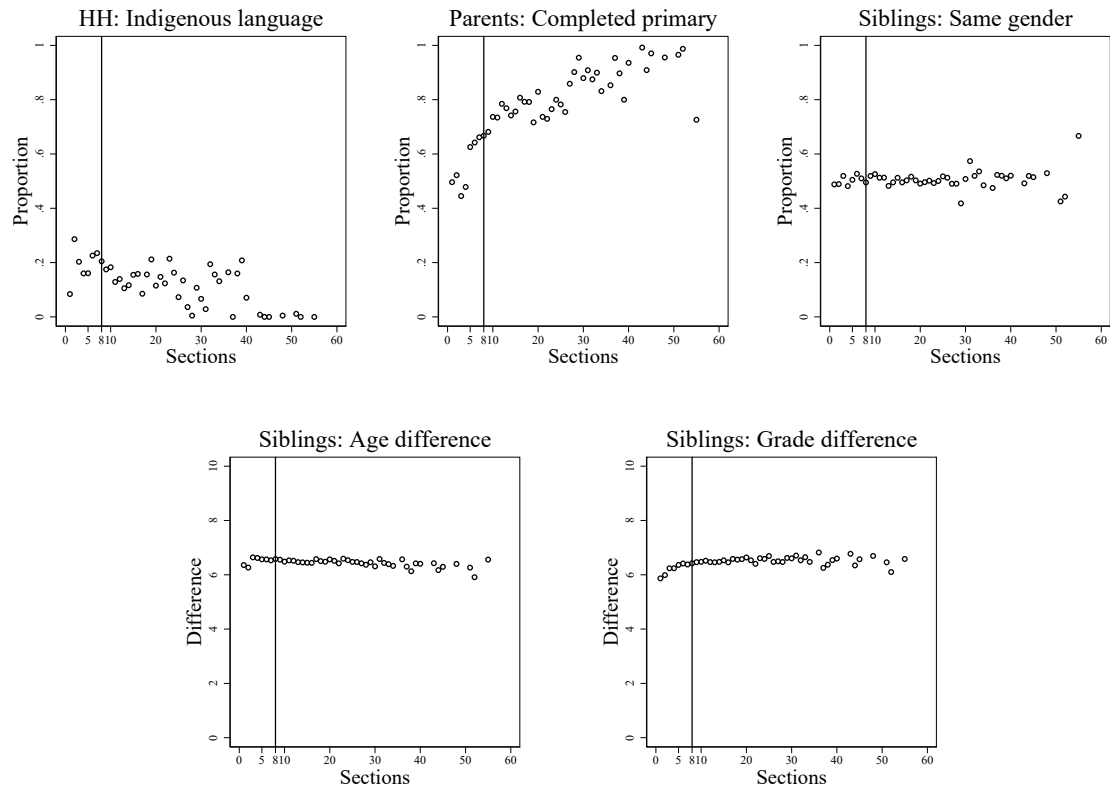**Figure A2:** Covariate Smoothness Across Cutoff: School-Level Characteristics
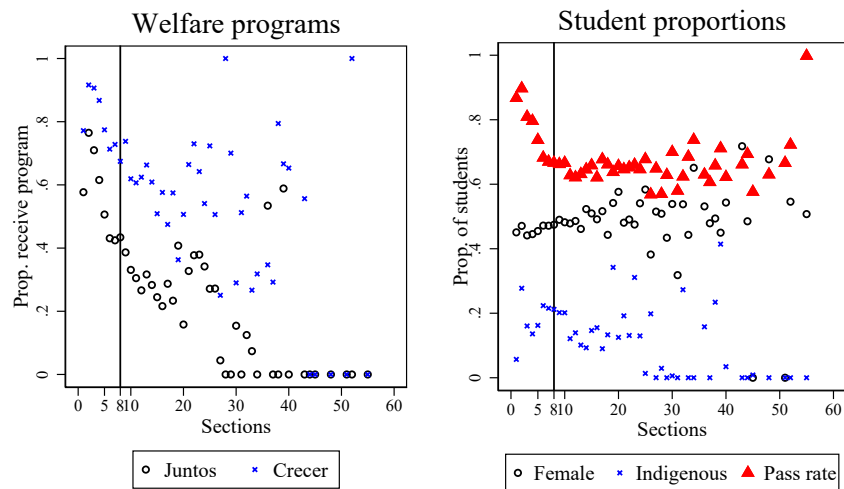
Figure A3: Discontinuity in Outcomes at Cutoff

Reading
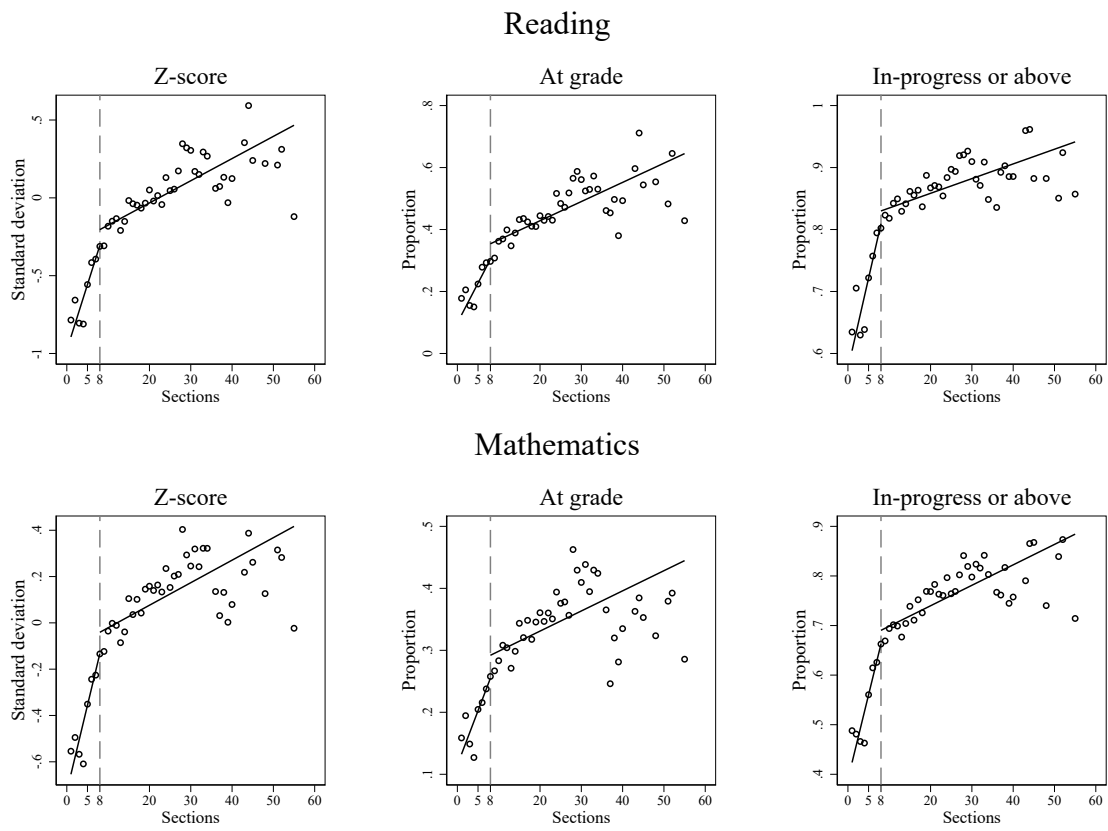
Mathematics

Table A1: Distribution of Weekly Hours by Public School Type

Subject	Basic	JEC
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

Notes: A pedagogical hours is 45 minutes long.
Source: Peru's Ministry of Education.

Table A2: Selection Process for JEC

	Steps	Schools
1.	**Initial eligibility:** – Public secondary schools – Eight or more sections – Registered as morning shift only – School premises only used in the morning – Sufficient space for additional resources	1362
2.	52 "emblematic" schools added.	1412
3.	Local coordinators carry out validation process. Remaining schools required to provide additional information.	1343
4.	External assessors use information to select 1000 schools.	1000
5.	JEC reform announced in September 2014, along with list of schools (RM N° 451-2014-MINEDU)	1000
6.	On February 10th list is modified, with six schools replaced (RM N° 062-2015-MINEDU)	1000

Source: Alcázar (2016) and Agüero et al. (2021).

Table A3: ECE Coverage, by Year and Grade

Year	Grade	% Schools evaluated	% Students evaluated
2015	2nd Grade	99.7	94
2016	2nd Grade	99.8	96.5
	4th Grade	99.8	96.9

Notes: Coverage of primary schools is provided only as total for all grades evaluated. *Source:* MINEDU.

Table A4: TSLS Predetermined Covariate Smoothness Tests

	School district receives program		Proportion of students		Pass rate
	Juntos	Crecer	Girls	Indigenous	Total
Panel A: School level					
Sections≥8	0.007	-0.029	0.014	-0.047	0.030
	(0.036)	(0.035)	(0.011)	(0.027)*	(0.012)**
Control mean	0.508	0.775	0.457	0.171	0.736
N	59338	59338	59338	59338	59338

	Sibling Dyad			Household	
	Same gender	Age diff.	Grade diff.	Indig. lang.	Parent educ.
Panel B: Sibling dyad level					
Sections≥8	-0.010	0.041	-0.029	-0.054	-0.014
	(0.012)	(0.040)	(0.027)	(0.024)**	(0.021)
Control mean	0.506	6.563	5.702	0.174	0.616
N	59338	59338	59338	59338	59338

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Estimates are the reduced form intention-to-treat effect at the cutoff, including a linear spline of the running variable. Cluster robust standard errors at the older sibling school level are reported in parentheses.

Table A5: Effects of JEC on Younger Sibling Outcomes: ITT Results

	Reading			Mathematics		
	(1) Z-score	(2) At grade	(3) ≥ In-prog.	(4) Z-score	(5) At grade	(6) ≥ In-prog.
Sections≥8	0.082	0.004	0.008	0.092	0.020	0.037
	[0.003]***	[0.760]	[0.530]	[0.002]***	[0.124]	[0.011]**
Control mean	-0.393	0.293	0.795	-0.226	0.238	0.626
N	4368	4368	4368	4368	4368	4368

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Estimates are the local randomisation intention-to-treat effect for units above the cutoff, within W. Local randomisation asymptotic p-values based on heteroskedastic robust standard errors are reported in square brackets. Results modelled without polynomial adjustment for the smallest possible window.

Table A6: Effects of JEC on Younger Sibling Outcomes: 2SLS, No Controls

	Reading			Mathematics		
	(1)	(2)	(3)	(4)	(5)	(6)
	Z-score	At grade	≥ In-prog.	Z-score	At grade	≥ In-prog.
Panel A: Linear specification						
Sections≥8	0.133	0.060	0.017	0.128	0.057	0.039
	(0.063)**	(0.026)**	(0.023)	(0.064)**	(0.023)**	(0.029)
Controls	No	No	No	No	No	No
N	59954	59954	59954	59954	59954	59954
Panel B: Quadratic specification						
Sections≥8	-0.085	-0.064	-0.060	-0.045	-0.012	-0.045
	(0.102)	(0.043)	(0.038)	(0.102)	(0.039)	(0.046)
Controls	No	No	No	No	No	No
N	59954	59954	59954	59954	59954	59954

Notes: * p< 0.10, ** p< 0.05, *** p< 0.01. Estimates are of the local average treatment effect (LATE) at the cutoff, including a linear (panel A) and quadratic (panel B) spline of the running variable. Cluster robust standard errors at the older sibling school level are reported in parentheses.

Table A7: Heterogeneous Effects, by Gender, 2SLS

	Female			Male		
	(1)	(2)	(3)	(4)	(5)	(6)
	All	Older sister	Older brother	All	Older sister	Older brother
Reading						
Z-score	0.298	0.235	0.351	0.030	0.019	0.033
	(0.069)***	(0.084)***	(0.086)***	(0.065)	(0.076)	(0.081)
At grade	0.113	0.124	0.101	0.039	0.037	0.038
	(0.030)***	(0.037)***	(0.038)***	(0.028)	(0.035)	(0.034)
≥ In prog.	0.067	0.032	0.097	-0.003	-0.012	0.006
	(0.027)**	(0.035)	(0.034)***	(0.025)	(0.033)	(0.033)
Mathematics						
Z-score	0.277	0.251	0.297	0.003	-0.025	0.026
	(0.071)***	(0.086)***	(0.089)***	(0.071)	(0.084)	(0.088)
At grade	0.102	0.092	0.112	0.023	0.019	0.027
	(0.026)***	(0.033)***	(0.034)***	(0.029)	(0.035)	(0.037)
≥ In prog.	0.118	0.091	0.139	-0.026	-0.034	-0.019
	(0.033)***	(0.043)**	(0.041)***	(0.032)	(0.038)	(0.041)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
N	29237	14546	14691	30101	14598	15503

Notes: * p< 0.10, ** p< 0.05, *** p< 0.01. Estimates are of the local average treatment effect (LATE) at the cutoff, including a linear spline of the running variable. Cluster robust standard errors at the older sibling school level are reported in parentheses. Additional covariates include sibling pair age difference in years, an indicator of being the same gender, and mother's educational attainment in years. Fixed effects for older sibling grade and survey year are also included.

Table A8: Local Randomisation Placebo Test, Private Schools

	Reading			Mathematics		
	(1)	(2)	(3)	(4)	(5)	(6)
	Z-score	At grade	≥ In-prog.	Z-score	At grade	≥ In-prog.
Sections≥8	-0.050	-0.085	-0.034	-0.032	0.020	-0.037
	[0.640]	[0.173]	[0.320]	[0.780]	[0.722]	[0.508]
Mean Dep.	0.130	0.532	0.924	-0.003	0.278	0.722
N	424	424	424	424	424	424

Notes: * p< 0.10, ** p< 0.05, *** p< 0.01. Estimates are the local randomisation intention-to-treat effect for units above the cutoff, within W. Local randomisation asymptotic p-values based on heteroskedastic robust standard errors are reported in square brackets. Results modelled without polynomial adjustment for the smallest possible window.

Table A9: Local Randomisation Placebo Test, Placebo Cutoffs

	Reading			Mathematics		
	(1)	(2)	(3)	(4)	(5)	(6)
	Z-score	At grade	≥ In-prog.	Z-score	At grade	≥ In-prog.
Panel A: Below cutoff						
Sections≥4	-0.006	-0.005	0.009	-0.042	-0.022	-0.003
	[0.894]	[0.787]	[0.702]	[0.371]	[0.192]	[0.895]
Mean Dep.	-0.805	0.155	0.630	-0.568	0.149	0.466
Window	[3,4]	[3,4]	[3,4]	[3,4]	[3,4]	[3,4]
N	1718	1718	1718	1718	1718	1718
Panel B: Above cutoff						
Sections≥12	0.016	0.029	0.007	-0.009	-0.004	-0.003
	[0.547]	[0.041]**	[0.504]	[0.758]	[0.767]	[0.844]
Mean Dep.	-0.150	0.370	0.842	-0.002	0.308	0.702
Window	[11,12]	[11,12]	[11,12]	[11,12]	[11,12]	[11,12]
N	4616	4616	4616	4616	4616	4616

Notes: * p< 0.10, ** p< 0.05, *** p< 0.01. Estimates are the local randomisation intention-to-treat effect for units above the cutoff, within W. Local randomisation asymptotic p-values based on heteroskedastic robust standard errors are reported in square brackets. Results modelled without polynomial adjustment for the smallest possible window.

Table A10: Alternative Window Selection Procedure

Window	Variable	Min. p-value	Obs.<8	Obs.≥8
[7, 8]	Prop. female students	0.251	2103	2265
[6, 9]	Age difference	0.308	4286	4573
[5, 10]	Juntos district	0.000	23503	9099
[4, 11]	Juntos district	0.000	24408	11529
[3, 12]	District Juntos eligible	0.000	25221	13715

Notes: Variable refers to the predetermined covariate with the lowest difference in means p-value across the cutoff within that window. List excludes two variables that are not balanced as shown in Table 4. Significance level for rejecting the null, $H_0 : Z_i(1) = Z_i(0) \forall i$, is set at $\alpha \geq 0.15$, as recommended by Cattaneo et al. (2024).

Table A11: Effects of JEC on Younger Sibling Outcomes: Alternative Window

	Reading			Mathematics		
	(1)	(2)	(3)	(4)	(5)	(6)
	Z-score	At grade	≥ In-prog.	Z-score	At grade	≥ In-prog.
Sections≥8	0.095	0.017	0.037	0.106	0.036	0.046
	[0.000]***	[0.078]*	[0.000]***	[0.000]***	[0.000]***	[0.000]***
Control mean	-0.405	0.286	0.776	-0.235	0.227	0.620
Window	[6,9]	[6,9]	[6,9]	[6,9]	[6,9]	[6,9]
N	8859	8859	8859	8859	8859	8859

Notes: * p< 0.10, ** p< 0.05, *** p< 0.01. Estimates are the local randomisation intention-to-treat effect for units above the cutoff, within W. Local randomisation asymptotic p-values based on heteroskedastic robust standard errors are reported in square brackets. Results modelled without polynomial adjustment for all windows.

Table A12: Effects of JEC on Younger Sibling Outcomes: Including 2018 Survey Wave

	Reading			Mathematics		
	(1)	(2)	(3)	(4)	(5)	(6)
	Z-score	At grade	≥ In-prog.	Z-score	At grade	≥ In-prog.
Sections≥8	0.088	-0.003	-0.020	0.095	0.015	0.046
	[0.014]**	[0.864]	[0.246]	[0.011]**	[0.364]	[0.014]**
Control mean	-0.387	0.281	0.748	-0.226	0.239	0.625
N	5539	5539	5539	5539	5539	5539

Notes: * p< 0.10, ** p< 0.05, *** p< 0.01. Estimates are of the local average treatment effect (LATE) above the cutoff, within W. Local randomisation asymptotic p-values based on heteroskedastic robust standard errors are reported in square brackets. Results modelled without polynomial adjustment for the smallest possible window.

Table A13: Effects of JEC on Younger Sibling Outcomes, by Survey Wave and Grade

	Reading			Mathematics		
	(1)	(2)	(3)	(4)	(5)	(6)
	Z-score	At grade	≥ In-prog.	Z-score	At grade	≥ In-prog.
Panel A: Year 2015 2nd grade						
Sections≥8	0.183	0.060	0.013	0.099	0.009	0.069
	[0.018]**	[0.113]	[0.575]	[0.190]	[0.791]	[0.068]*
Control mean	-0.381	0.328	0.902	-0.269	0.241	0.627
N	1805	1805	1805	1805	1805	1805
Panel B: Year 2016 2nd grade						
Sections≥8	0.032	0.001	0.028	0.041	0.043	-0.019
	[0.631]	[0.979]	[0.175]	[0.599]	[0.234]	[0.588]
Control mean	-0.355	0.346	0.914	-0.049	0.293	0.716
N	1321	1321	1321	1321	1321	1321
Panel C: Year 2016 4th grade						
Sections≥8	0.156	-0.020	0.057	0.222	0.033	0.109
	[0.024]**	[0.481]	[0.136]	[0.003]***	[0.267]	[0.004]***
Control mean	-0.454	0.177	0.484	-0.333	0.176	0.532
N	1242	1242	1242	1242	1242	1242

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Estimates are of the local average treatment effect (LATE) above the cutoff, within W. Local randomisation asymptotic p-values based on heteroskedastic robust standard errors are reported in square brackets. Results modelled without polynomial adjustment for the smallest possible window.

B Local Randomisation Framework

B.1 Sharp Local Randomisation RDD

Given a score, X_i, treatment assignment is given by $T_i = \mathbb{1}(X_i \geq c)$, where c is the cutoff point. Y_i is the *observed* outcome for unit i, with *potential* outcomes $Y_i(0)$ and $Y_i(1)$ under control and treatment. We assume for some scalar $w > 0$ there exists a window $W = [c - w, c + w]$ containing the cutoff c, such that the assumptions described below hold. Let $\mathbf{X}_W$ be the vector of scores and $\mathbf{T}_W$ be the vector of treatment statuses of all units i within that window for which $X_i \in W$, and let their equivalent potential outcomes be given by $\mathbf{Y}_W(0)$ and $\mathbf{Y}_W(1)$. Following Cattaneo et al. (2017, 2024), we define the conditions required for the basic local randomisation framework, assuming random potential outcomes drawn from a super-population. We first assume that there exists some window $W = [c - w, c + w]$, for which the potential outcomes of a unit i are statistically independent of their score X_i:

LR 1 (Independence of Scores from Potential outcomes):

$$(Y_i(0), Y_i(1)) \perp\!\!\!\perp X_i | X_i \in W$$

Alternatively, This can be stated in terms of probability distribution functions:

$$\mathbb{P}[X_i \leq x | Y_i(0), Y_i(1), X_i \in W] = \mathbb{P}[X_i \leq x | X_i \in W] \tag{B1}$$

This ensures that, for all units $X_i \in W = [c - w, c + w]$, placement above and below the cutoff is not related to potential outcomes, and potential outcomes are not related to the score[15]. This implies $\mathbb{E}[Y_i(d)|X_i, X_i \in W] = \mathbb{E}[Y_i(d)|X_i \in W]$ for $d = 0, 1$, meaning the conditional expectations are constant functions of the score inside the window (although they may have non-zero slopes outwith).

We can therefore define the parameter of interest. defining $\mathbb{E}_W[\cdot]$ as the conditional expectation with respect to the probability $\mathbb{P}_W[\cdot]$, computed conditionally for units with $X_i \in W$. Defining N_W as the number of units in window W, the local randomisation sharp RD treatment effect is given by:

$$\theta_{SRD} = \frac{1}{N_W} \sum_{i:X_i \in W} \mathbb{E}_W\left[\frac{T_i Y_i}{\mathbb{P}_W[T_i = 1]}\right] - \frac{1}{N_W} \sum_{i:X_i \in W} \mathbb{E}_W\left[\frac{(1 - T_i)Y_i}{1 - \mathbb{P}_W[T_i = 1]}\right] \tag{B2}$$

Which expresses the RD as the difference in average observed outcomes just below and just above the cutoff within the window. Under LR 1, this is equivalent to our parameter

---

[15]Alternatively a less strict assumption can be made that potential outcomes are independent from treatment assignment. This assumption however does not imply that potential outcomes are unrelated to the score variable (Sekhon & Titiunik, 2017), requiring an additional exclusion restriction assumption that potential outcomes are not a function of the score(Cattaneo et al., 2017). While our assumption is stricter, given that our narrow window this is reasonable, and ensures this exclusion holds implicitly.

of interest:

$$\theta_{SRD} \equiv \frac{1}{N_W} \sum_{i:X_i \in W} \mathbb{E}_W[Y_i(1) - Y_i(0)] \tag{B3}$$

The average treatment effect of T_i on Y_i within the window. The next section will extend the local randomisation framework for cases of imperfect compliance (Fuzzy RDD).

B.2 Fuzzy Local Randomisation RDD

In Fuzzy RDD, with assignment of treatment $T_i = \mathbb{1}(X_i \geq c)$, either some units with $X_i \geq c$ fail to receive the treatment or $X_i < c$ receive the treatment anyway. Therefore, the change in probability of receiving the treatment at the cutoff is not 0 to 1. Following Cattaneo et al. (2024), we define D_i as the indicator of treatment received, which has two potential values: $D_i(1)$ is the treatment received by unit i when assigned to the treatment condition ($T_i = \mathbb{1}(X_i \geq c) = 1$), and $D_i(0)$ treatment received when assigned to control ($T_i = 0$), with $D_i(1), D_i(0) \in \{0, 1\}$. We can write, for example, $D_i(0) = 1$ if a unit receives treatment even though it is assigned to the control condition (non-compliance), or $D_i(0) = 0$ if it is assigned control assignment and does not receive treatment, and similarly with $D_i(1)$. These are defined as our *potential* treatments.

Given noncompliance, we redefine potential outcomes as $Y_i(T_i, D_i(T_i))$, which includes arguments for both treatment assigned and *potential* treatment status. There are now four potential outcomes. The potential outcome for unit i assigned to treatment is given by $Y_i(1, D_i(1)) = D_i(1)Y_i(1, 1) + (1 - D_i(1))Y_i(1, 0)$, resulting in $Y_i(1, 1)$ if $D_i(1) = 1$ and $Y_i(1, 0)$ if $D_i(1) = 0$, with the potential outcomes for unit i assigned to control defined analogously. To interpret the below parameters of interest we must, in addition to assuming potential outcomes are not a function of the score (LR 1), assume that potential treatments $D_i(1)$ and $D_i(0)$ are also unaffected by the score X_i within W:

LR 2 (Independence of Score from Potential Treatments):

$$(D_i(0), D_i(1)) \perp\!\!\!\perp X_i | X_i \in W$$

under LR 2, as the potential treatments $D_i(1), D_i(0)$ are not related to the score, we can say that our augmented potential outcome $Y_i(T_i, D_i(T_i))$ is also not related to the score, fulfilling LR 1. This allows for the definition of an assumption which is analogous to the exclusion restriction evoked in traditional instrumental variable (IV) settings. For a given value of treatment received $D_i = d$, for all units i with $X_i \in W$:

LR 3 (Exclusion Restriction on the Treatment Assignment):

$$Y_i(T_i, d) = Y_i(d) \qquad \forall d$$

This implies that the treatment assignment affects potential outcomes and potential treatments only through the treatment received D_i, not but not directly. Given assignment $T_i = \mathbb{1}(X_i \geq c)$ is a function of X_i, assuming that our augmented outcomes $Y_i(T_i, 0)$ and $Y_i(T_i, 1)$ are not related to the score implies that, given a particular value of treatment received $D_i = d$, potential outcomes do not depend on treatment assignment T_i.

However, LR 3 alone is not sufficient to recover the effects of treatment in fuzzy RDD. This is due to the decision to comply with treatment assignment still being unrestricted. To be able obtain our treatment effects of interest, we must make additional assumptions. Following the standard set up of Imbens and Angrist (1994), we define four different groups based on their compliance decisions: *Compliers*, whose treatment received matches their assigned treatment status (whose potential treatments are such that $D_i(1) = 1$ and $D_i(0) = 0$); *Always-takers*, who always take up treatment whether assigned or not ($D_i(1) = D_i(0) = 1$); *Never-takers*, who always refuse treatment regardless of assignment ($D_i(1) = D_i(0) = 0$; and *defiers*, who would always receive the opposite treatment from the one assigned ($D_i(1) = 0$ and $D_i(0) = 1$. We assume that, within $W = [c - w, c + w]$:

LR 4 (Monotonicity):

$$D_i(1) \geq D_i(0) \qquad \forall i, X_i \in W$$

This condition ensures that assignment to treatment affects the treatment received in a monotone way. That is, if respondents on average are more likely to receive treatment given $D_i(1)$ than $D_i(0)$, then anyone who would receive treatment under $D_i(0)$ must also receive treatment under $D_i(1)$. As such, this condition rules out the potential for defiers.

Let B_i be a binary variable denoting compliance status, $B_i = 1$ for compliance with treatment assignment and $B_i = 0$ for non-compliance. Then under assumptions LR 1 – LR 4 it can be shown that the effect of the treatment received for the sub-population of units that comply with treatment assignment within the window W, commonly referred to as the Local Average Treatment Effect (LATE) is given by:

$$\theta_{FRD} = \frac{1}{N_W} \sum_{i:B_i, X_i \in W} \mathbb{E}\left[Y_{1,1} - Y_{1,0} | B_i = 1\right] \tag{B4}$$

When the above assumptions are met, we can show that the ratio of the effect of the treatment assignment on the outcome, θ_{ITT}, and the effect of the treatment assignment on treatment received, θ_{FS} is equivalent to the LATE for compliers, θ_{FRD}:

$$\theta_{FRD} \equiv \frac{\theta_{ITT}}{\theta_{FS}} \tag{B5}$$

From this we can proceed with identifying the required parameters given our assumptions.

B.3 Intention-to-Treat Effects

Define the effect of being *assigned* to treatment, whether treatment is received or not, on the outcome Y_i as θ_Y. Applying the fuzzy RDD context to Equation B2 we can estimate the following parameter:

$$\theta_Y = \frac{1}{N_W} \sum_{i:X_i \in W} \mathbb{E}_W \left[\frac{T_i Y_i(1, D_i(1))}{\mathbb{P}_W[T_i = 1]} \right] - \frac{1}{N_W} \sum_{i:X_i \in W} \mathbb{E}_W \left[\frac{(1 - T_i)Y_i(0, D_i(0))}{1 - \mathbb{P}_W[T_i = 1]} \right] \tag{B6}$$

The difference in average observed outcomes just below and just above the cutoff within the window. Under LR 1 and LR 2:

$$\theta_Y = \theta_{ITT}, \qquad \theta_{ITT} \equiv \frac{1}{N_W} \sum_{i:X_i \in W} \mathbb{E}_W [Y_i(1, D_i(1)) - Y_i(0, D_i(0))] \tag{B7}$$

such that the estimated difference in average observed outcomes above and below the cutoff is equivalent to the causal effect of T_i on Y_i, commonly referred to as the "intention-to-treat effect" (ITT), θ_{ITT}, within the window W.

B.4 First Stage

We also define the effect of assignment to treatment T_i on the probability of receiving D_i, which reveals information regarding compliance, as θ_D, treating D_i as the outcome. We define the following parameter:

$$\theta_D \equiv \frac{1}{N_W} \sum_{i:X_i \in W} \mathbb{E}_W \left[\frac{T_i D_i}{\mathbb{P}_W[T_i = 1]} \right] - \frac{1}{N_W} \sum_{i:X_i \in W} \mathbb{E}_W \left[\frac{(1 - T_i)D_i}{1 - \mathbb{P}_W[T_i = 1]} \right] \tag{B8}$$

θ_D captures the difference in the probability of *receiving* the treatment between units *assigned* to treatment and units *assigned* to control, within the window W. Under assumptions LR 1 and LR 2, we can interpret this parameter as the causal impact of T_i on D_i. Following the IV literature, we refer to this as the first stage effect, θ_{FS}:

$$\theta_D = \theta_{FS}, \qquad \theta_{FS} \equiv \frac{1}{N_W} \sum_{i:X_i \in W} \mathbb{E}_W [D_i(1) - D_i(0)] \tag{B9}$$

B.5 Estimation of LATE

In the local randomization framework, θ_{ITT} and θ_{FS}, can be estimated by calculating sample difference-in-means between units above (with subscript $W+$) and below (subscript $W-$) the cutoff for units with scores in W:

$$\widehat{\theta}_{ITT} = \bar{Y}_{W+} - \bar{Y}_{W-} \quad \text{and} \quad \widehat{\theta}_{FS} = \bar{D}_{W+} - \bar{D}_{W-}, \quad \widehat{\theta}_{FRD} = \frac{\bar{Y}_{W+} - \bar{Y}_{W-}}{\bar{D}_{W+} - \bar{D}_{W-}} \tag{B10}$$

where

$$\bar{Y}_{W+} = \frac{1}{N_{W+}} \sum_{i:X_i \in W} \omega_i T_i Y_i, \quad \bar{Y}_{W-} = \frac{1}{N_{W-}} \sum_{i:X_i \in W} \omega_i (1 - T_i) Y_i \tag{B11}$$

and

$$\bar{D}_{W+} = \frac{1}{N_{W+}} \sum_{i:X_i \in W} \omega_i T_i D_i, \quad \bar{D}_{W-} = \frac{1}{N_{W-}} \sum_{i:X_i \in W} \omega_i (1 - T_i) D_i, \tag{B12}$$

,

where ω is a weighting scheme for unit i. Inference is based on standard IV large sample approximations using the Delta method, applied to observations within the window W. As θ_{FRD} is a ratio, it will be undefined when the denominator is zero, thus a further assumption must be made that the first stage exists and is non-zero:

LR 5 (Instrument Relevance):

$$\theta_{FS} \neq 0, \quad or \quad \mathbb{P}[D_i = 1 | X_i \geq c, X_i \in W] > \mathbb{P}[D_i = 1 | X_i < c, X_i \in W]$$

Where the second definition is similar to that stated by Cattaneo et al. (2017). Furthermore, similar to IV settings, when the cutoff rule has a non-zero but small effect on the probability of treatment ("weak instruments"), standard Gaussian approximations of the distributions of test statistics are not reliable. The strength of the first stage can be assessed by the size of the F-statistic in the first stage regression, with Cattaneo et al. (2024) recommending a higher rule of thumb threshold of an F-statistic of 20 or more.