The Effect of a Preschool Expansion on Early Learning Outcomes in Peru

Stephanie Majerowicz* and Claude Raisaro†

Version: June 2023. Draft please do not cite without author permission

Abstract

Can public preschool education improve early learning outcomes and narrow socioeconomic gaps in academic performance in developing countries? This paper presents quasi-experimental evidence of the national expansion of public preschools in Peru on learning outcomes. We exploit town-level and within family variation in exposure to preschool due to the gradual expansion of preschools across Peru. We find that having access to a regular preschool improves second-grade standardized test scores for reading comprehension and mathematics by between 0.05 and 0.12 standard deviations. Exploring mechanisms, we look at two different preschool types rolled out in Peru: regular preschools and community preschools (in which local mothers deliver the service with limited supervision). Assignment of the different types of preschool is based on the number of preschool-aged students in each town and we exploit discontinuities in this assignment rule through a regression discontinuity design. We find some evidence that being assigned to a preschool with a trained teacher and proper infrastructure has a positive impact on student learning for students in towns near the cut-off compared to those assigned to community schools. Finally, we find that despite contributions to learning, having access to preschool appears to widen rather than close socioeconomic gaps in early achievement, suggesting that complementary measures targeting the poorest students are necessary for greater educational equity.

*Universidad de los Andes. Email: smajerowicz@uniandes.edu.co

†University of Zurich. Email: claude.raisaro@econ.uzh.ch. This project would not have been possible without the support of María Antonieta Alva, José Carlos Chávez and Minister Jaime Saavedra in the Ministry of Education of Peru, as well as the assistance of Ministry staff, particularly Andrei Bartra, Ricardo Montero, María Pia Basurto, Liliana Miranda, and Zoila Llempen. We are also grateful to participants at the Harvard Economic Development Seminar, briq and IZA, for helpful comments on earlier versions of the paper.

1 Introduction

Many countries have recently embarked on efforts to roll out large-scale preschool programs designed to address poor learning outcomes in primary school. A growing body of evidence suggests that early childhood education can be particularly important for disadvantaged students, helping close achievement gaps that develop early in life (Heckman, 2006; Engle et al., 2011). The evidence, which largely comes from the United States, suggests that comprehensive preschool programs have strong positive effects across a number of short and run-long outcomes (McCoy et al., 2017; Cunha et al., 2006; Yoshikawa et al., 2013), but the complexity and scope of these programs make them difficult to scale-up, particularly in developing countries. Much less is known about the capacity of large-scale national preschool programs to affect early learning outcomes and narrow socioeconomic gaps in developing countries (Richter et al., 2017; Britto et al., 2017). Moreover, there are still large gaps in our understanding of which components of preschool matter the most and therefore how national preschool programs should be structured.

This paper presents quasi-experimental evidence from the expansion of a national preschool program in Peru on learning outcomes over the short and run long. Between 2010 and 2015, the government of Peru launched an ambitious program to ensure universal preschool access. Based on growing evidence of the importance of early childhood interventions, Peru's Access to Preschool program sought to close the achievement gaps between students and improve overall learning outcomes by better-preparing students for primary school. This expansion of preschool successfully increased enrollment by more than 25 percentage points in one decade, closed the urban-rural preschool enrollment gap, and led to national enrollment levels (88% in 2014) five percentage points *above* the average for high-income countries (WDI, 2014). Nevertheless, while the program has successfully bridged gaps in access to preschool, little is known about its impact on learning.

Comparing children that go to preschool with those who do not is likely to yield biased estimates of the impact of preschool for various reasons. First, towns with preschools are likely to differ from towns without them in many other ways that affect educational outcomes, not only because they are likely to be wealthier, but also but also in the quality of other educational inputs. Similarly, families that choose to send children to preschool are likely to be different from families that do not. Even within families, selection problems can arise if families choose to send certain children to preschool and then differentially invest in those children relative to their other children (for example, sending their male child, or the child perceived to be "most gifted"). Comparisons across towns, across families and even within families that do not address these selection problems would yield biased estimates.

To overcome this selection bias and estimate the effect of preschool access on student learning outcomes, we take advantage of the gradual roll-out of preschools across rural areas of Peru, which generated variation in exposure across and within cohorts of students that is exogenous to the decisions of the families. We exploit this variation in two ways. First, we explore cross-town variation through an event study comparing towns that received a preschool during our study period against towns that had not yet received one. This dynamic difference-in-differences estimator allows us to control for all differences across towns that are constant over time, and estimate a causal treatment effect under the assumption that learning outcomes would have evolved similarly in treatment and control towns had the preschools not been built.

Second, we exploit variation in the exposure of siblings to the rollout of preschools using a family fixed effects estimator. We take advantage of detailed administrative data to compare younger siblings who had access to a new preschool built in their town to older siblings who were too old to benefit from the preschool. We use siblings in families who did not experience a change in the availability of preschool to control for time trends, general differences between younger and older siblings, and other national interventions that differentially affect certain cohorts. This empirical strategy is similar to the one employed by Deming (2009) in the evaluation of Head Start, but with the advantage that variation in within-household exposure to preschool comes from the creation of preschools and not from decisions to send one child instead of another.

We find that having access to a preschool improves 2nd grade standardized test scores in both reading comprehension and mathematics. According to the cross-town event study estimates, having a preschool built in the town increased learning by 0.04 standard deviations for reading comprehension and a slightly higher impact of 0.05 SD for mathematics. Turning to the within family estimator, we find that having a regular preschool built in your town increases standardized test scores for the exposed child compared to their older siblings by 0.12 s.d. and 0.07 s.d. for reading comprehension and mathematics, respectively. These are both intent-to-treat estimates, which should be scaled up by the proportion of students who effectively attend preschool. While we do not know this proportion for students in the sample, we can estimate the proportion of students who attend preschool using administrative enrollment data available in recent years, which yields approximate average treatment effect between 0.16 to 0.2 s.d. for reading and 0.09 and 0.1 s.d. for math.

We estimate heterogeneous effects by socioeconomic characteristics to explore whether preschool helped better prepare the most vulnerable students for primary school, effectively narrowing socioeconomic early achievement gaps. We compare the gap between poor and non-poor students within a primary school and its interaction with preschool. We find some evidence that the socioeconomic gap widens for students who attended preschool relative to students who did not. This counter intuitive result could be due to differing quality of the preschools attended by poor and non-poor students, or due to complementary inputs that are provided for the wealthier students. This suggests that additional investments in quality and complementary inputs are necessary to help close early childhood achievement gaps.

While preschool is intended to improve outcomes through developing cognitive and social skills of kids, there is a question as to how much the quality of the teacher and other inputs like classroom infrastructure matter. Many developing countries with dispersed populations, including Peru, have implemented variations of early childhood education that rely on community members, rather than trained teachers, to provide care for students. Are these alternatives enough? Could the benefits of early childhood education be achieved through these more cost-effective means or is a trained teacher necessary for the developmental benefits of preschool? It may be the case, for example, that being exposed to other children is enough to generate the early childhood stimulation that is crucial to healthy cognitive development, or that the real benefit of preschool education is freeing women to work which can lead to improved family incomes and therefore better nutrition that translates into learning improvements.

Peru's dispersed rural population makes building regular preschools prohibitively costly for large segments of the country. For this reason, the Peruvian government developed three types of preschools depending on the number of preschool-age children in a given town. We explore the variation in the types of preschools built in each town in Peru to shed light on the relative importance of the different inputs in early childhood education. The regular preschools provide a trained teacher and formal school infrastructure, while the community preschools provide day care and interactions with peers, but with a much more limited structured learning environment and instead of a trained teacher, a community member (usually) a local mother provides the service with some periodic supervision from a teacher.

We first look at the differential impact of each type of school using our two identification strategies above. In the event-study, only regular preschools appear to improve learning outcomes, while we find no significant effects of the community preschools on learning. On the other hand, the family-estimator does find significant effects for both types of preschools, but the effects are higher for the regular preschool compared to the community preschool. In particular, while having access to a regular preschool improves learning scores by 0.07 and 0.12 s.d. for math and reading comprehension respectively, for community preschools the increase is only of 0.04 and 0.06 s.d. respectively, somewhere between two thirds and half of the impact of the regular preschools.

While this suggests that the type of preschool matters, this differential impact is not causal. It is possible that the underlying conditions of these communities (which are smaller, poorer and more dispersed) make preschool ineffective—perhaps these students are missing necessary complementary inputs (parental education, nutrition, etc.) such that no type of preschool would work for them.

We are able to take advantage of the rule that determines the decision to build one type of school rather than another based on the number of preschool-aged children in the town in order to estimate the causal differential effect of attending a formal preschool versus a community preschool. The decision to build a type of preschool depended on the population of the town, which allows us to estimate the causal effect through a fuzzy regression discontinuity design. We use administrative census data to estimate the preschool age population of all towns at the time the school is being designed and built. This allows us to compare students in towns right above the cut-off who received a regular preschool to those below the cut-off who received a community preschool. We find some evidence that the type of preschool matters: living in a town that received a regular preschool has an effect of about 0.9 standard deviations relative to students in towns that received a community preschool. This contrasts with recent evidence from Colombia where transferring students from community to regular preschools actually had a negative impact on their cognitive development (Bernal et al., 2019).

This paper makes contributions to the evidence base of the effect of preschool on learning outcomes at national scale. While there is abundant evidence on the short and long run effects of preschool, the literature has largely focused on estimating the effects of small, targeted and comprehensive preschool programs in developed countries. Experimental evidence from the Perry Preschool and the Abecederian programs show large and persistent effects of the programs on outcomes ranging from high school graduation and college attendance, to crime and income through age 40 (Schweinhart et al., 2005; Heckman et al., 2010). Quasi-experimental evidence from Head Start, a much larger albeit not universal preschool program in the United States, shows similarly positive short and long run results (Deming, 2009). Nevertheless, both Perry Preschool and Head Start involved a package of interventions for both students and parents, of which preschool was only one. It is therefore difficult to extrapolate the findings to developing country contexts with a much more limited scope of providing preschool education.

Well-identified evidence on early childhood education in developing countries remains scarce, although a growing evidence base suggests positive impacts on a variety of outcomes (see Nores and Barnett (2010) for a review). In Latin America, a few studies from Argentina, Uruguay and Bolivia using quasi-experimental methods and matching estimators find evidence that preschool education improves schooling and learning outcomes, and the labor market participation of mothers (Berlinski et al., 2008, 2009; Berlinski and Galiani, 2007; Berlinski et al., 2011; Behrman et al., 2004).[1] In Colombia, Bernal and Fernandez (2013) use a matching estimator, as well as variation in length of exposure to evaluate a community preschool program, Hogares Comunitarios. They find that participation leads to declines in malnutrition (presumably from the meals component) and improvements in cognitive and social development.

This paper makes three contributions to the existing literature. First, it extends existing studies to low income populations in a developing country in the context of national preschool program.[2] Second, it improves on the empirical strategy used in previous studies

[1] In Argentina, Berlinski et al. (2009) implement a difference-in-difference estimator similar to Duflo (2001) and find that an additional year of preschool education improves test scores in 3rd grade by 8%, and increases maternal employment (Berlinski and Galiani, 2007), later confirmed by a regression discontinuity design (Berlinski et al., 2011). In Uruguay, Berlinski et al. (2008) use a within-family estimator similar to Deming (2009), and find some significant impacts on years of education attained by age 16. However, this latter paper does not have exogenous variation in access to preschool so that, like Deming (2009) there is an open question of why a sibling went to preschool and the other did not. Behrman et al. (2004) use a matching estimator in Bolivia and find that impacts are dependent on age and exposure duration, with robust effects only for those exposed more than 7 months.

[2] A few studies have looked at the impact of preschool in Peru, but they are more than 15 years old and do not have a well-identified empirical strategy (Cueto and Diaz, 1999; Diaz, 2006). Cueto et al. (2016) in a more recent study use instrumental variables to look at the impact on learning outcomes of attending preschool for students with early life stunting, which represents a methodological improvement although the instruments may violate the exclusion restriction.

using exogenous variation in the availability of preschool and administrative student-level data. Third, it sheds light on some mechanisms and unpacks which characteristics of preschools matter for improving learning outcomes by exploring the differential impact of different types of preschools.

The rest of the paper proceeds as follows: Section 2 provides background on the preschool expansion and the different types of preschools in Peru. Section 4 discusses the data and empirical strategy to identify the main treatment effect, and Section 5 presents the results. Section 6 explores socioeconomic preschool gaps. Section 7 then discusses the empirical strategy and results for the regression discontinuity across different preschool types, and Section 8 concludes.

2 Background: Peru's Preschool Expansion

Like its peers across the region, Peru experienced large improvements in primary and secondary school enrollment over the past two decades. However, Peru continues to face strong challenges in educational quality and equity, with most students in the country failing to meet basic learning standards measured by both national and international standardized test scores. In 2016, less than half of the students in Peru met learning standards for second grade in primary school in reading comprehension, while only about one third met them in mathematics. Even more disheartening, only 14% of students who made it to second grade of secondary school achieved learning standards in reading, and only 11.5% in mathematics (Escale, 2017).

While there are low levels of quality overall, there are also large challenges in equity, with huge persisting gaps in educational coverage and quality across urban and rural populations. In 2001, only 23.9% of 25-34 year-olds had completed secondary school in rural areas, while the proportion in urban areas was 70%. These gaps have narrowed in recent years but remain large: by 2017, the proportion of 25-34 year-olds that had completed secondary school had risen to 38.6% in rural areas, still less than half that of their peers in urban areas at 80%. Similarly, students in rural areas lag behind their urban peers in educational achievement: in 2016, only 16% and 17.3% of rural students met learning standards in reading and mathematics respectively, far below their urban peers with 50% and 36.6% respectively (Escale, 2017).

In addition, while there were large improvements in access to primary school, lack of access to early childhood education means that many students across the dispersed rural populations of Peru are not prepared for entering primary school. The Access to Preschool Program was implemented in order to close these large rural-urban gaps in access to preschool education, as well as improve overall enrollment rates which were low: in 2001, net enrollment rates of children in urban areas were 60% compared to 42% in rural areas. From 2005 to 2015, over 15,000 preschools were created all over Peru.

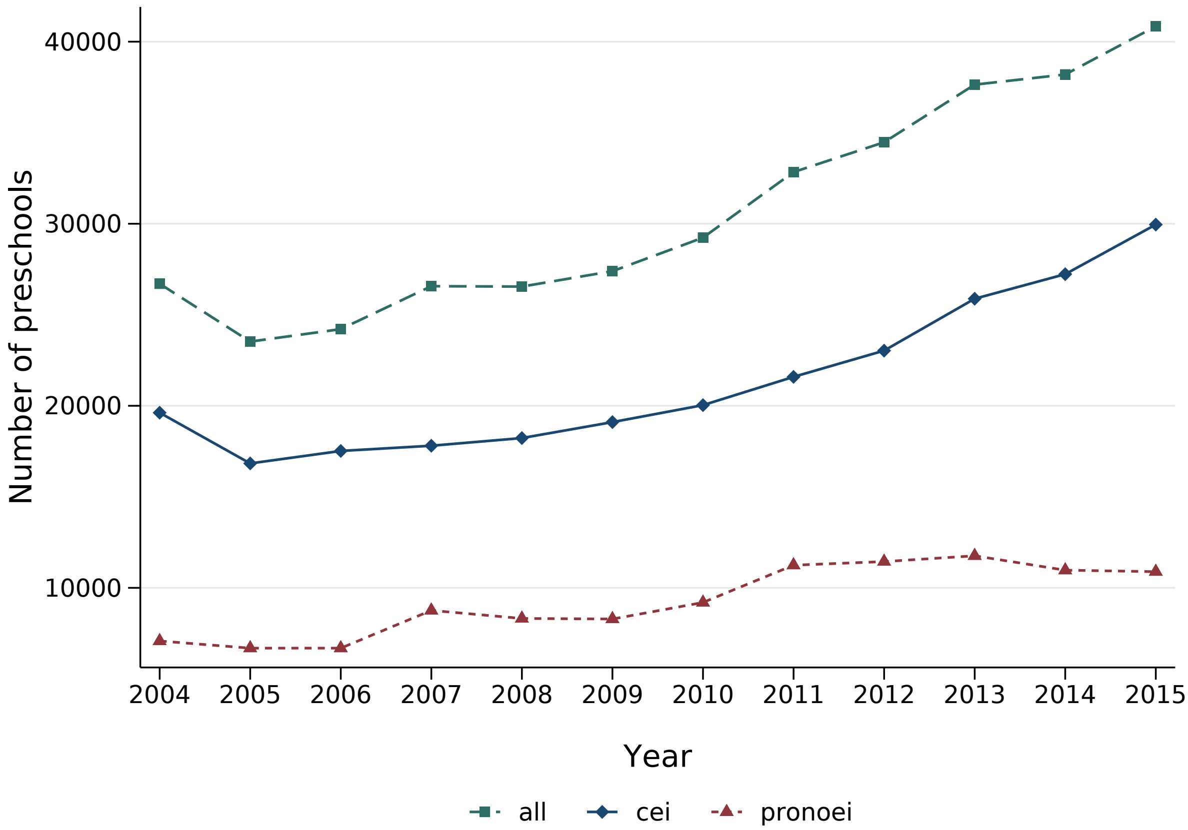
While the National Ministry designed and financed the program, the decision of where to create a school was determined by the local school board, UGEL (the local education administrative unit, Unidad de Gestión Escolar Local, in Spanish). UGELs first selected towns with potential unmet demand (defined as the number of preschool-age children without access to preschool), and carried out "demand studies" where teams would visit prioritized towns and conduct a census of the student population by going door to door and counting all preschool or soon-to-be preschool aged children in the town. Each UGEL then prioritized towns according to unmet demand using the results from these studies, and the top towns were selected to receive a school.

Once the towns were selected, the *type* of school that got built was also determined by the number of preschool-age children in a given town, according to the following rule:

- 15 or more students: CEI - Regular preschool with designated classroom and at least one fully trained government teacher.
- 8-14 students: PRONOEI- Community preschool where local mother (promotora) runs the school under occasional supervision of a teacher coordinator.
- 7 or fewer: PZD- Household visits where a trained teacher visits parents once in a while to teach them how to play with the child.

Figure 1 shows the expansion of the preschool program since 2004 by type of school. It shows that the expansion involved the creation of both numerous regular preschools (CEI) as well as community schools (PRONOEI). Figure 2 shows how the expansion was associated with sharp increases in enrollment rates, as reported in the national household survey (ENAHO). Panel A shows that the urban-rural gap had closed by 2015, converging at a national preschool enrollment rate of 84%. Panel B disaggregates this enrollment data by age group, showing that enrollment rates for 4 and 5-year-olds rose above 90%, while, despite some improvements, there remain large gaps in the enrollment of 3-year-olds.

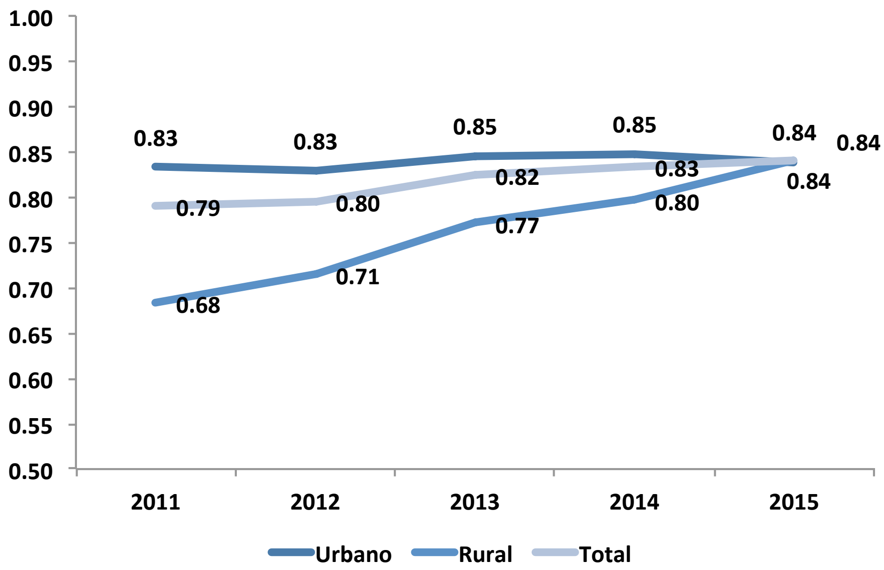
Figure 1: Expansion of preschools by type (2004-2015)



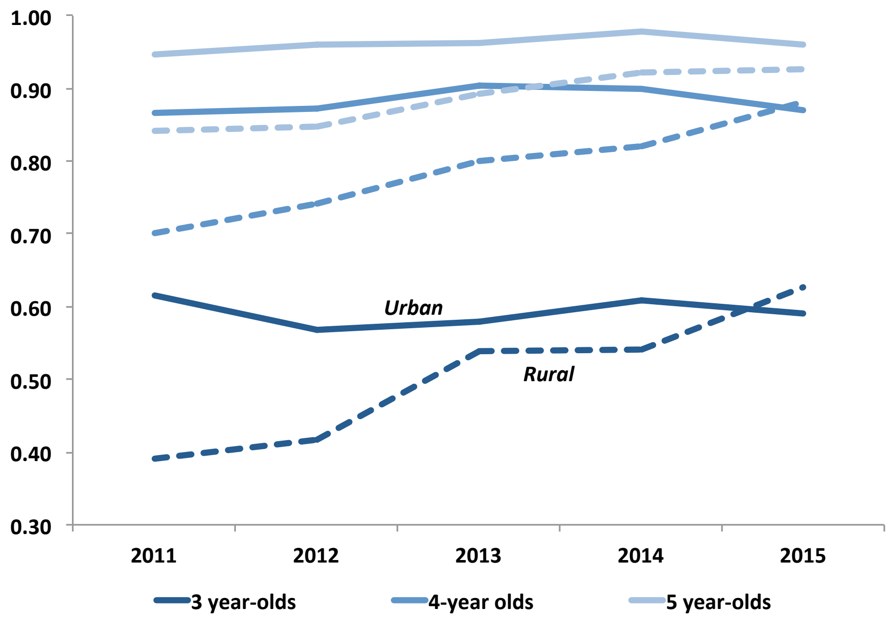
Source: Padron de Escuelas Escolarizado y No Escolarizado, MINEDU.

Preschool in Peru is available to children ages 3-5, and public preschools are supposed to be free, with parents paying only for school uniforms. In practice, there is anecdotal evidence that some public teachers charge parents small fees to enroll their students. In most public preschools, meals are provided by a national school-feeding program called Qali Warma, which has extensive coverage although it varies by region. Regular preschools extend for 5 hours a day, while the community preschools normally receive students for a minimum of 4 hours a day. Community preschools are led by a local community member called a *promotora*, who is not a trained teacher but receives training, supervision and support from a trained teacher who serves as a coordinator (docente coordinadora). However, the degree of supervision and support varies strongly with coordinators often having to travel long distances to provide support and oversight to their *promotoras*.

Figure 2: Urban-Rural Preschool Enrollment Gap



(a) Reduction in Urban-Rural Enrollment Gap



(b) Rural-Urban Gap by Age Group

Source: National Household Survey (ENAHO).

Table 1: Descriptive statistics by preschool type

	Private CEI	Public CEI	PRONOEI
No. Students	45.4	43.9	9.97
No. Teachers	3.3	2.2	0.98
Rural (dummy)	0.01	0.60	0.52
Rurality (Scale)	1.86	5.95	5.30
Income Quintile	3.91	1.97	2.32
N (Active)	12,164	23,003	18,626

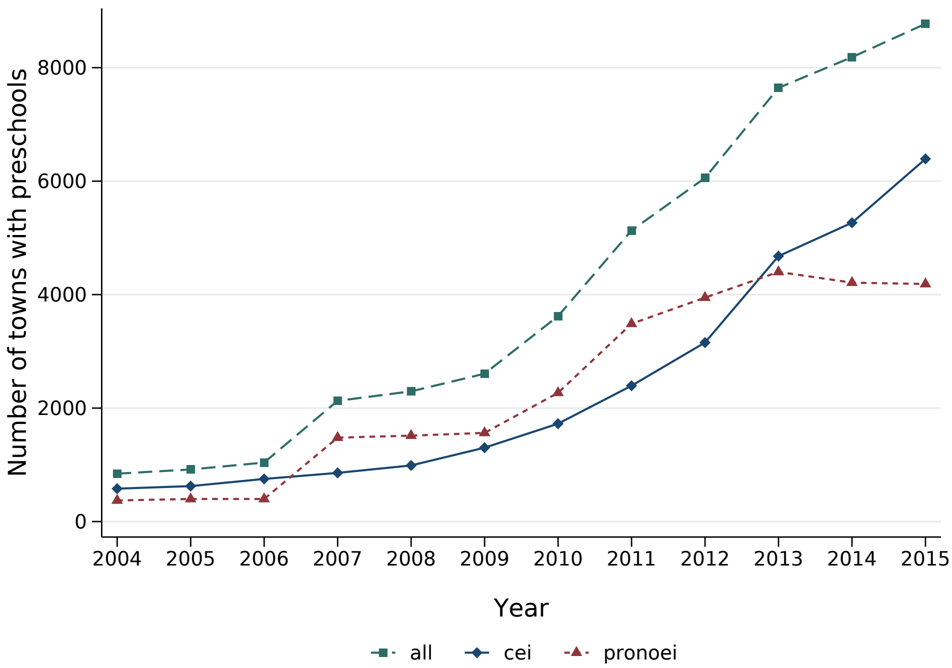
Source: SIAGIE 2015

Table 1 shows some descriptive statistics for the two types of schools, differentiating between private and public regular preschools. As expected, CEI's are on average larger than PRONOEIs, which are intended for small towns with fewer than 15 preschool-aged students. Nevertheless, PRONOEIs are often found in urban areas and only 52% of PRONOEIs are rural, which is due to the presence of large numbers of PRONOEIs in Lima, where there should be none. Private preschools are overwhelmingly urban, while most public preschools (60%) are located in rural areas. Private preschool students are also much wealthier than their public counterparts, as expected.

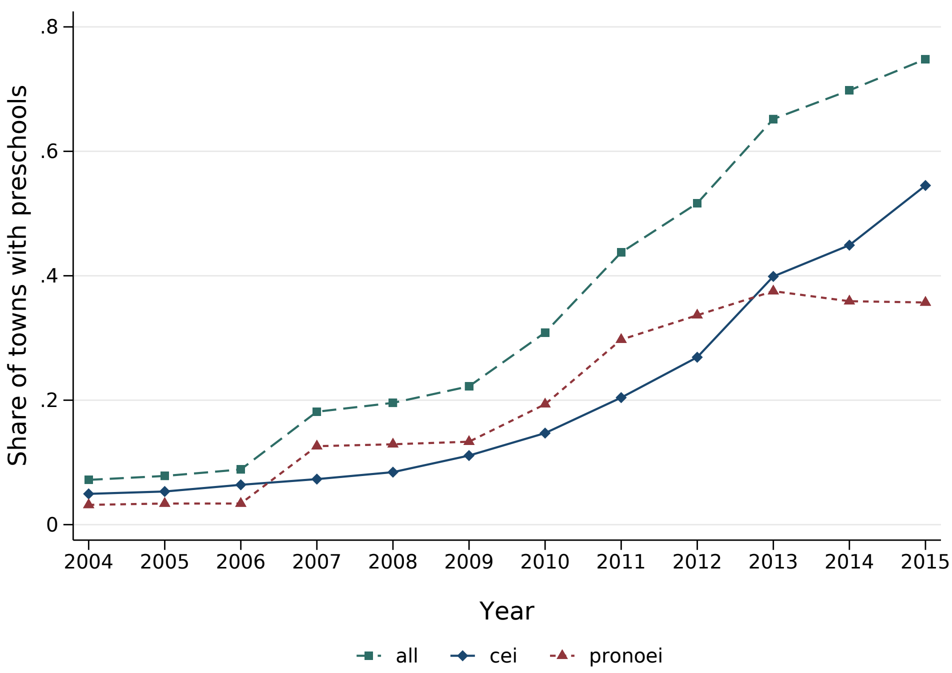
The rollout of preschools affected towns at different points in time between 2004 and 2015. Figure 3 describes the changes across time in the number and in the share of towns that had a least one preschool, by type. At the beginning of our sample period, less than 10% of towns in Peru had access to any type of preschool, but by 2015 almost 80% of towns did. The share of towns with access to a regular preschool increased by approximately 50 percentage points, while the increase in access to community preschools was a bit lower, around 35 percentage points. Our identification strategy exploits the variation in exposure to preschool generated by the steep rise in the creation of preschools all over Peru.

Figure 3: Availability of preschools at town level (2004-2015)

(a) Number of towns treated



(b) Share of towns treated



Sample of towns exposed to preschool any time after 2003.

3 Data

This paper uses administrative data from the Ministry of Education in order to estimate the effect of preschool on learning outcomes, and to estimate the relative effectiveness of the different types of preschools.

Short term learning outcomes are measured using student-level scores on 2nd grade standardized exams (the Evaluación Censal Estudiantil, or ECE) taken by all students in Peru since 2007. The exam is carried out in all schools with more than 5 students enrolled in 2nd grade, and is carried out in both Spanish as well as native languages for students whose first language is not Spanish. The exam tests students on quantitative and reading comprehension skills, and assigns students to one of 3 categories: "Satisfactory", "Below Satisfactory" and "Beginning" the lowest score. A standardized Rasch score is also assigned to each student, which has a mean of 500 and standard deviation of 100. Our sample of students is made up of students who took this exam between the years 2007 and 2018.

In addition, we are able to measure longer term learning outcomes by following students throughout their school progression and observing them at several additional points in time. Besides the second grade of primary school, certain cohorts of students also took standardized exams in 4th grade of primary school, and then again in 2nd grade of secondary school. In addition, we can track students in the administrative databases of the school system and observe grade progression, repetition or drop-out up until 2015.

We identify siblings using a rolling population census (called SISFOH) carried out berween 2012 and 2013 that covers most of rural (and a significant amount of urban) Peruvian households. The census gathered data on socioeconomic characteristics, educational attainment, household composition, and labor market participation of households. School aged individuals are then matched with the test scores for the years 2007-2018. We are able to match approximately 2.5 million students who appear in the census and have taken the test between 2007 and 2018, of which more than half (1.3 million) are in families where we were able to match at least two siblings to the test data.

Data on school characteristics including geocoded location, information on teachers, inputs and students come from administrative datasets of the Ministry of Education. The date of creation of each school comes from the administrative Padron of schools, which has information on the creation and closures of all schools. Before 2011, PRONOEIs were not included in the regular padron records and were instead found in a different part of the school census. We are able to recreate their dates of creation by going through these dataset and assigning the original date of creation to each PRONOEI (this means some of the PRONOEI creation dates may be more subject to measurement error).

Figure 2 describes our sample of students by whether they lived in towns that had access to a preschool during our study period. Column 1 presents descriptive statistics for students who lived in a town that never got a preschool during our sample period (until 2015); column 2 refers to students who lived in towns that at the beginning of our study period did not have a preschool and received at least one between 2004 and 2015; column 3 refers to students who lived in towns that had access to preschool prior to 2004. Students living in towns that did not get exposure (column 1) or got exposure during the study period (column 2) are similar along a large set of observable characteristics. However, students who lived in towns with preschools available prior to 2004 (column 3) are generally wealthier than the other two categories: they belong to smaller and richer families and their parents are on average more educated. This makes sense since these are likely to be residents of larger cities, where some form of preschool has been available for a long time.

Table 2: Students characteristics by level of exposure to preschool

	(1) Never	(2) Changers	(3) Always	(1)-(2)	(1)-(3)	(2)-(3)
Female	0.4958	0.4919	0.4924	-0.0038	-0.0034	0.0004
	(0.5000)	(0.4999)	(0.4999)	(0.0015)	(0.0013)	(0.0009)
	[164,599]	[308,569]	[4373682]			
Family size	5.5791	5.6418	5.1570	0.0628	-0.4220	-0.4848
	(1.8906)	(1.8728)	(1.8179)	(0.0079)	(0.0065)	(0.0045)
	[81,655]	[180,749]	[1668733]			
Mother no education	0.2258	0.2607	0.0727	0.0349	-0.1530	-0.1879
	(0.4181)	(0.4390)	(0.2597)	(0.0018)	(0.0010)	(0.0007)
	[80,757]	[177,287]	[2210621]			
Mother primary education	0.5593	0.6008	0.3235	0.0415	-0.2358	-0.2773
	(0.4965)	(0.4897)	(0.4678)	(0.0021)	(0.0017)	(0.0012)
	[80,757]	[177,287]	[2210621]			
Mother secondary educ.	0.2150	0.1385	0.6038	-0.0764	0.3888	0.4652
	(0.4108)	(0.3455)	(0.4891)	(0.0016)	(0.0017)	(0.0012)
	[80,757]	[177,287]	[2210621]			
Father no education	0.0779	0.0873	0.0229	0.0094	-0.0550	-0.0645
	(0.2680)	(0.2823)	(0.1495)	(0.0012)	(0.0006)	(0.0004)
	[74,153]	[163,626]	[1928887]			
Father primary education	0.6018	0.6718	0.2745	0.0701	-0.3272	-0.3973
	(0.4895)	(0.4695)	(0.4463)	(0.0021)	(0.0017)	(0.0012)
	[74,153]	[163,626]	[1928887]			
Father secondary educ.	0.3204	0.2408	0.7026	-0.0795	0.3822	0.4617
	(0.4666)	(0.4276)	(0.4571)	(0.0019)	(0.0017)	(0.0012)
	[74,153]	[163,626]	[1928887]			
Mother housework	0.8282	0.8828	0.6631	0.0546	-0.1650	-0.2196
	(0.3772)	(0.3217)	(0.4726)	(0.0014)	(0.0017)	(0.0011)
	[80,748]	[177,356]	[2209777]			
Poverty category	1.8991	1.8606	2.1203	-0.0385	0.2213	0.2597
	(0.9262)	(0.9154)	(0.9067)	(0.0039)	(0.0033)	(0.0023)
	[79,256]	[175,883]	[2148880]			
School walking dist.	0.7639	0.8719	0.7255	0.1080	-0.0384	-0.1464
	(0.4247)	(0.3342)	(0.4463)	(0.0051)	(0.0055)	(0.0034)
	[6,786]	[18,047]	[308,376]			

Note: This figure reports student-level characteristics by whether the town they live in was never exposed to preschool during our study period (Column 1), was exposed for the first time between 2004 and 2015 (Column 2) or had been exposed already prior 2004 (Column 3). The poverty category is a categorical variable that takes values 1 if the household is categorized as extremely poor, 2 if poor and 3 if non-poor. Standard deviations are in parenthesis and observations are in brackets.

4 Empirical Strategy

In order to estimate the impact of preschool access on student learning, we take advantage of the gradual rollout of preschool across rural Peru in two different ways. First, we exploit variation across towns and years through an event study design; second, we take advantage of the variation in the exposure of siblings to the rollout of preschools through a family fixed effects model.

4.1 Event Study Estimator

Our first empirical strategy relies on cross-town variation in exposure to preschool. As shown in 3, the preschool expansion program led to a steep rise in the number of towns that had access to preschool over a ten year period. In 2006, only 10% of towns had preschools, while by 2015 more than three quarters of towns had access to at least some type of preschool. As described previously, the decentralized nature of the roll-out means that there was no systematic spatial clustering in the creation of new schools, and similar towns across different school districts are likely to have received preschools at different points in time.

We exploit this variation in exposure through an event study, which compares the evolution of standardized test scores of pupils located in towns where a preschool gets built to towns where the preschool had not yet been built at that time.

In order to identify the dynamic causal effect of the rollout of schooling outcomes, we estimate the following regression model:

$$Score_{ity} = \alpha_i + \delta_y + \sum_{j=-6}^{7} \beta_j Preschool_t \times T_{i,y=y^*+j} + \epsilon_{ity} \tag{1}$$

where $Preschool$ is a dummy that takes value 1 if a preschool was available in the town, t of student i who took the standardized test score in year y; γ_y are year fixed effects; $\mathbf{X}$ is a vector of controls that includes linear trends for region and for the number of students at town level at baseline.

Students may start preschool at age of 3 and attend until the age of 5, and students take the standardized test scores in the second grade of primary school, around age 7. As a result, we code student i as treated, if her town t received a preschool at least 3 years before she took the standardized test score. That is we code her as treated if a school was built in her town when she was four years old, meaning that it was available to her when she was 5 years old thus allowing her to have at least one year of treatment. In other words, we code cohorts of students as treated three years after the school is built in their town, starting in year y^*.

The coefficient β_j captures the treatment effect of the preschool on student i's outcomes $j+1$ years after the potential exposure to preschool when $j \geq 0$. For $j < 0$, we estimate the placebo treatment effects j years before the initial rollout in a given town. The outcome variable is the score of the student i on the standardized 2nd-grade exam ECE. Standard errors are clustered by town, which is the level of treatment.

Our event study estimator relies on the traditional difference-in-difference parallel trends assumption that in the absence of treatment (a preschool being built), standardized test scores would have evolved similarly in the control and treatment groups (in this case, towns that received and did not receive preschools). We use the periods prior to the preschool being built to provide evidence consistent with this assumption.

We restrict our sample to towns that at the beginning of our time period (2004) had not yet received a preschool. In other words, we exclude towns which had preschools during our entire sample since these towns, particularly larger cities, are likely to be very different than those that received preschools during our sample period. Table 2 compares towns that always had a preschool during our sample against those who either acquired a preschool during that time or who by 2016 had not yet received one. It shows that the earlier group that always had preschools is systematically different from the other two. On the other hand, towns that received a preschool during our evaluation window and those that had not yet done so are much more similar to each other.[3] This is important both to inform the external validity limitations of our study (we cannot speak to the effect of preschools in large cities), as well as to suggest that given their similarity, it may reasonable to extrapolate our findings (with some caveats) to the towns not yet reached.

[3]The table 2 shows the difference in means between these three groups. While most differences are statistically significant due to the large sample size, the magnitude of the differences between those that never had and first-time receivers is a fraction of what they are between them and those who always had a preschool.

The recent difference-in-difference literature has raised a number of concerns regarding the validity of the TWFE estimator when treatment is staggered and effects may be heterogeneous across units or across time. As a robustness check, we complement the standard two-way fixed effect estimator with the estimator proposed by (de Chaisemartin and D'Haultfœuille, 2020), which corrects for the "forbidden comparisons" of using already treated groups as control for later treated units, which can introduce bias into the TWFE estimator (see section 5.2).

4.2 Family Fixed Effects Estimator

As a second identification strategy, we exploit variation in the exposure of siblings to the rollout of preschools through a family fixed effects estimator. The estimator is also a difference-in-difference estimator, but one which uses older siblings as a control group to treated younger siblings similar to the one used by Deming (2009)[4] or at a more aggregated level by Duflo (2001). This estimator exploits two sources of variation: the difference between younger and older siblings (cohort variation), and the difference between families that experienced a change in the availability of preschool and families that did not (family or town variation). Within family differences control for any observable and unobservable omitted variables that are common to both siblings (parent's education, pedagogical practices, etc.), while across family differences control for changes across time, as well as systematic differences between younger and older siblings. Figure A.1 in the Appendix shows a graphical representation of this empirical strategy.

The identification assumption in this second estimator is once again the difference-in-difference common trends assumption. In this case, the assumption is that in the absence of treatment (had they not received access to a preschool) younger siblings in treated families would have followed similar trend lines (relative to their older siblings) as those families that did not observe changes in preschool availability over the time that the siblings were preschool-age. While this parallel trend assumption cannot be formally tested, it can be indirectly tested by looking at families with three or more siblings, and using middle untreated siblings as a placebo test.

The main specification for the within family estimator is the following:

[4] However, while Deming (2009) used a within-family estimator, he did not have exogenous variation of the availability of preschool, and had to demonstrate that this selection into take-up of preschool within siblings did not bias his estimates.

$$Score_{ifct} = \alpha_f + \delta_c + \beta Preschool_{ifct} + \mathbf{X'}\gamma + \varepsilon_{ifct} \tag{2}$$

where $Preschool$ is a dummy that takes value 1 if a preschool was available in the town, t, of student i from family f and cohort c when she was 5 years old (this is coded in the same way explained for the town event study); α_f is a family fixed effect, δ_c is a cohort fixed effect; $\mathbf{X}$ is a vector of controls that vary within family (like gender of the child). The outcome variable is the score of student i on the standardized 2nd grade exam ECE. Standard errors are clustered by family.

5 Results

5.1 Main Results

We find evidence that the exposure to preschool has a positive and modest impact on learning outcomes in both reading comprehension and in mathematics.

5.1.1 Event Study Results

We run the event study estimation disaggregating by the type of preschool that the town received (regular vs. community preschools). We find a positive, albeit moderate, effect of acquiring a regular preschool on student learning, but find no statistically significant effect of receiving a community preschool on learning outcomes through the event study.

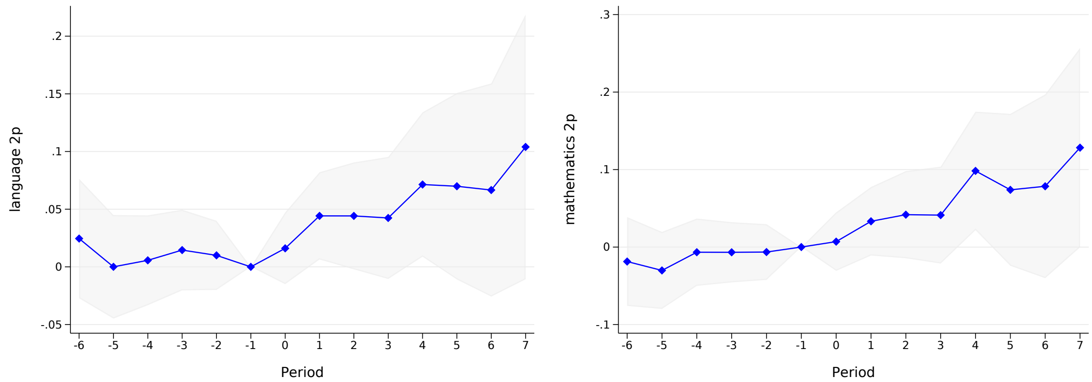
Figures 4 and 5 show the results of the event study design of equation 1 for towns that received a regular preschool and a community preschool respectively.[5] The figures show that in the years prior to receiving the preschool, there were no significant differences between treated and untreated towns, which is consistent with the parallel trend assumption. Then, for those towns that receive a regular preschool, once it is built the coefficients increase slowly over time and become statistically significant between two to four years after students start being exposed. While not all coefficients are significant, being exposed to a regular preschool appears to increase learning outcomes by between 0.05 and 0.1 standard deviations for both reading and mathematics test scores. However, these results do not hold for community preschools. There appears to be no effect of receiving a community preschool on learning outcomes according to the event-study.

Table 3 shows the TWFE estimator, aggregating all the post-treatment coefficients, and it shows an increase of 0.05 standard deviations on student learning for students in towns that received a regular preschool, which is consistent with the magnitudes shown in the event study. Once again, the results hold only for those towns that received a regular preschool, and we find no effect of the creation of a community preschool on learning outcomes using this identification strategy.

[5] Figure A.2 in the appendix shows the effect of having any of the two schools, without differentiating between the two types.

Figure 4: Effect of CEI preschool on learning outcomes

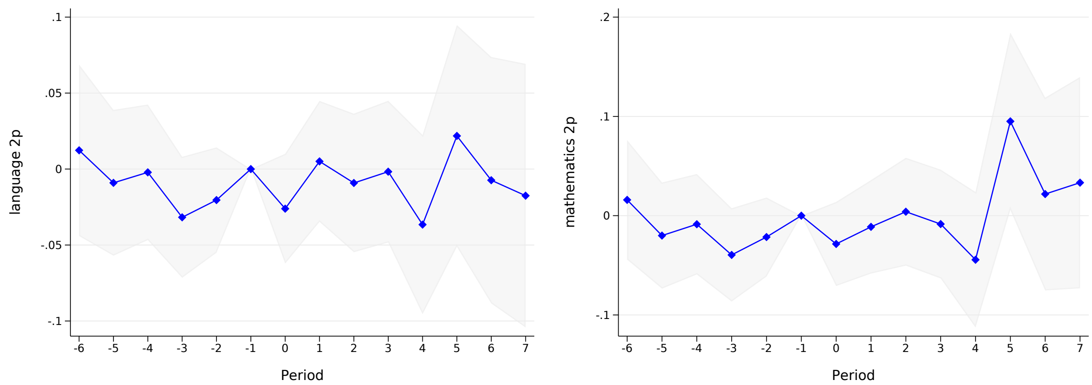
(a) Reading (b) Mathematics



Sample of towns exposed to preschool starting from 2004.

Figure 5: Effect of PRONOEI preschool on learning outcomes

(a) Reading (b) Mathematics



Sample of towns exposed to preschool any time after 2003.

Table 3: Effect of Preschool Availability - Town Two Way Fixed Effects Estimator

	Regular		Community		Any	
	Reading (1)	Math (2)	Reading (3)	Math (4)	Reading (5)	Math (6)
Regular Preschool	0.0470***	0.0439**				
	[0.0150]	[0.0172]				
Community Preschool			-0.0185	-0.0231		
			[0.0149]	[0.0173]		
Any Preschool					0.0201	0.0131
					[0.0127]	[0.0146]
Observations	416454	417108	416454	417108	416454	417108
R-squared	0.368	0.295	0.368	0.295	0.368	0.295
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Note: This table shows the result of the two way fixed effect difference in difference estimator on learning outcomes in second grade. Columns 1 and 2 show the effect of receiving a regular preschool, columns 3-4 show the effect of receiving a community preschool, while columns 5-6 group both types of preschools together. Standard errors in brackets. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

5.1.2 Family Fixed Effects Results

Table 4 shows the results of the family fixed effects specification shown in equation 2. Columns 3 and 4 show that having access to a regular preschool improves test scores by 0.07 standard deviations in math and 0.12 s.d. in reading comprehension. On the other hand, for students who received a community preschool in their towns, there is a smaller but still statistically significant impact of 0.04 s.d. and 0.06 s.d. on math and reading test scores respectively.[6]

It is important to highlight that the point estimates for both identification strategies measure the effect of the *availability* of preschool regardless of whether the students actually attended preschool or not, and therefore represent an intent to treat estimate. This *ITT* effect would need to be scaled up by the proportion of students who actually attended preschool conditional on having access to it, but unfortunately this data is not available.

While we do not know the proportion of students that actually went to preschool con-

[6]Columns 1 and 2 group together both types of preschool and show that on average having access to any preschool improves 2nd grade standardized test scores by 0.03 and 0.08 s.d. for math and reading scores respectively.

Table 4: Effect of Preschool Availability - Family fixed effects estimator

	(1) math	(2) reading	(3) math	(4) reading	(5) math	(6) reading
Any preK avail	0.0340***	0.0772***				
	[0.0126]	[0.0114]				
CEI avail			0.0663***	0.118***		
			[0.0137]	[0.0123]		
PRONOEI avail					0.0423***	0.0620***
					[0.0153]	[0.0137]
Observations	228132	227824	228132	227824	228132	227824
R-squared	0.708	0.752	0.708	0.752	0.708	0.751

Standard errors in brackets

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: This table shows the effect of having access to a preschool (radius of 1km around town) on early learning outcomes as measured by second grade standardized test scores. Columns 1-2 include all types of preschools, while Columns 3-4 and 5-6 break it down by type of preschool, regular preschools (CEI) and community preschools (PRONOEI) respectively. All regressions include family fixed-effects, cohort fixed effects and gender. Robust standard errors clustered by family. Scores are standardized so coefficients can be interpreted as standard deviations.

ditional on having access, there are two ways to try to estimate this proportion. First, national statistics of preschool enrollment during this period suggest that roughly 58% of rural students attended preschool. Using the family fixed effect estimates, this would scale up our treatment effects for the regular preschool to 0.2 standard deviations for reading and 0.1 s.d. for math. However, this national school enrollment data could reflect the fact that some towns had schools and others did not, while telling us little about which students of those who had access to preschool actually chose not to go. In order to approximate this more closely, we take advantage of the fact that preschool enrollment data was collected starting in 2013 in the administrative database of SIAGIE. We find that 75% of students that had a preschool available in the town where they were enrolled in elementary school attended preschool. While this is an estimate for 2013-2014 and attendance rates have almost certainly increased over the years, this yields a conservative average treatment effect of 0.16 standard deviations for reading comprehension, and 0.09 for math. The ATT is therefore likely to fall somewhere between these upper and lower bounds giving estimates of 0.16 to 0.2 for reading and 0.09 and 0.1 for math.

Finally, while the results are robust to the inclusion of region specific time trends and other robustness checks (see next section), it is important to keep in mind that as a fixed effects estimator, we are estimating the effect of having access to preschool for the subset of the sample that has observed changes in preschool availability in the last decade or so (those in Column 2 of table 2. This means that our ability to extrapolate the findings to other populations is limited. For instance, it is evident that our estimates should not be extrapolated to students who had access to preschools in Lima (highly urban centers that had preschools long before the sample begins) and are excluded from our sample. And while one should also be cautious in extrapolating the estimated effect to the control group (in this case, towns that had not yet received a preschool), the similarity in characteristics between the two groups suggests that these may be a more reasonable extrapolation.

5.2 Robustness Checks

We run several robustness checks on both identification strategies to rule out some potential sources of bias, and test the sensitivity of our results to certain parameters.

Family Fixed Effects Estimator: For the family fixed effects estimator, one potential concern is that once a preschool is created in the town that only the younger siblings are young enough to attend, parents will dedicate more resources to the younger sibling who now has an advantage. This would be particularly problematic if these resources come from substituting away from an older sibling, since the control group would now be negatively treated. To address potential substitution between siblings that could affect the control group, as a robustness check, we restrict my sample of siblings to those who are more than three years apart so that the older sibling would have already taken the exam by the time the younger sibling started attending preschool. While this does not rule out substitution, it prevents any substitution from biasing our estimate since it would have come after the older sibling had already taken the exam.

This robustness for the family fixed effect estimator are included in Table 5. Panel A presents the results of limiting the sample to siblings who are at least three years apart in age and show virtually identical results to the main specification, despite the smaller sample. As an additional robustness check, Panel B shows the main specification with the inclusion of region by cohort fixed effect to allow for the possibility that time trends are region and not nation-specific. The point estimate is identical for most of the specifications, and only lowers for community preschools, where it is no longer significant for math test scores.

Table 5: Robustness Checks- Family Fixed Effects Estimator

	(1) Math	(2) Reading	(3) Math	(4) Reading	(5) Math	(6) Reading
Panel A. Siblings with 3 years age difference						
Any prek available	0.0297**	0.0718***				
	[0.0124]	[0.0112]				
CEI available			0.0562***	0.109***		
			[0.0133]	[0.0119]		
PRONOEI available					0.0378**	0.0508***
					[0.0148]	[0.0132]
Observations	97573	97454	97573	97454	97573	97454
R-squared	0.528	0.586	0.528	0.587	0.528	0.586
Panel B. Region by year fixed effects						
Any prek available	0.0365***	0.0729***				
	[0.0129]	[0.0117]				
CEI available			0.0781***	0.119***		
			[0.0141]	[0.0127]		
PRONOEI availble					0.0235	0.0439***
					[0.0156]	[0.0141]
Observations	228132	227824	228132	227824	228132	227824
R-squared	0.712	0.755	0.712	0.756	0.711	0.755

Standard errors in brackets

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: This table shows two robustness checks: Panel A restricts the sample to siblings more than three years apart in age, while Panel B includes region by year fixed effects to account for the possibility that time trends vary by region. Standard errors are clustered by family. Scores are standardized and can be interpreted as standard deviations.

Event study estimator: As mentioned earlier, for the across-town event study, one potential concern is that the nature of the staggered roll-out alongside heterogeneous treatment effects could induce bias if we already-treated units are used as controls (see (de Chaisemartin and D'Haultfœuille, 2020) for a more detailed explanation). In order to test the robustness of our event study estimator we implement the estimator proposed by (de Chaisemartin and D'Haultfœuille, 2020). Figure ?? in the appendix show the results of the estimator. The Chaisemartin and D'Haultfoeuille estimator shows that the results are robust for reading comprehension test scores, but not for mathematics. We run additional robustness checks, including testing the sensitivity of our results to variations in the catchment area (the maximum distance that a preschool can be from a town centroid and still be considered as having access to it) from 0 meters to 1.5 km. We find that our results are robust to variations in this catchment area as well as to the inclusion or exclusion of controls.

6 Does preschool narrow socioeconomic gaps?

One of the main motivations for the provision of public preschool education is that cognitive skills are developed at an early age so that kids from disadvantaged households are already behind their peers by the time they enter primary school. Schady et al. (2015) measure the socioeconomic gaps between the richest and poorest quartiles in early childhood cognitive development in 5 countries, including Peru, and find large socioeconomic gaps across all of them ranging from 0.78 to 1.23 standard deviations. Peru falls somewhere in the middle of the sample with a gap of 0.95 s.d. in urban students and 0.77 for rural students. Public preschool, by providing early learning and cognitive stimulation to all children, is envisioned as a way of closing socioeconomic achievement gaps.

However, it is not obvious that preschool will have a leveling effect on early childhood achievement and cognitive skills. If parental stimulation and behaviors are complements to preschool, preschool may have a larger benefit for those kids with more educated and wealthier parents who can dedicate time and effort to complementary activities. Furthermore, if an important part of the impact of preschool is through freeing mothers to join the labor market (Berlinski and Galiani, 2007; Berlinski et al., 2011; Hallman et al., 2005), those who start off with higher levels of education may be able to get higher returns on their labor, exacerbating initial levels of inequality.

Even if preschool does help close socioeconomic gaps, the question remains to what extent it can make up for the difference in stimulating environments at home, and if there is a minimum level above which children are equally prepared for primary school.

Table 6 presents the results of the family-level difference in difference estimator in Table 4, disaggregating by the socioeconomic status of the students. It interacts the availability of preschool variable in the main specification with a socioeconomic index from the Ministry of Social Protection (MIDIS). The index codes families as poor, extremely poor and non-poor. We compare families coded as poor or extreme poor against those coded as non-poor. The table shows that the effect of preschool availability is *smaller* in mathematics for families characterized as poor, reducing by half the effect for those families coded as non-poor. While the coefficient for reading comprehension is also negative, we find no significant difference in reading across the different socioeconomic groups.

Table 6: Heterogeneous Effects by Poverty- Family Fixed Effects Estimator

	(1) Math	(2) Reading	(3) Math	(4) Reading	(5) Math	(6) Reading
Any prek available	0.0626***	0.0916***				
	[0.0185]	[0.0165]				
Any prek available × poor	-0.0430**	-0.0237				
	[0.0210]	[0.0186]				
CEI available			0.0946***	0.137***		
			[0.0209]	[0.0186]		
CEI available × poor			-0.0439*	-0.0300		
			[0.0246]	[0.0218]		
PRONOEI available					0.0522**	0.0573***
					[0.0238]	[0.0212]
PRONOEI available × poor					-0.0151	0.00409
					[0.0287]	[0.0254]
Observations	220205	219906	220205	219906	220205	219906
R-squared	0.706	0.751	0.706	0.751	0.706	0.751

Standard errors in brackets

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: This table shows the effect of having access to a preschool (radius of 1km around town) interacted with a poverty measure where poor takes value 1 if the family is characterized by the Ministry of Social Protection as being "extremely poor" and 0 if "poor". Columns 1-2 include all types of preschools, while Columns 3-4 and 5-6 break it down by type of preschool, regular preschools (CEI) and community preschools (PRONOEI) respectively. All regressions include family fixed-effects, cohort fixed effects and gender. Robust standard errors clustered by family. Scores are standardized so coefficients can be interpreted as standard deviations.

Table 7 shows another way of calculating the socioeconomic test score gap by preschool attendance for a more recent sample of students for whom we have preschool attendance data. In this case, we compare the performance of students who attend the same primary school, by whether they attended preschool or not and their socioeconomic characteristics. The gap is the average difference in test scores between a poor and non-poor student. The variable “Poor” measures the socioeconomic test-score gap for students who did not attend preschool, while the interaction term measures the change in this gap for students who did attend preschool. Columns 1 and 4 show this effect aggregated for students who attended any preschool, while columns 2, 3, 5 and 6 disaggregate the sample by students who attended only regular preschool (CEI) and for students who attended community preschools (PRONOEI).

As expected, on average non-poor students perform much better than poor students on the second grade standardized test scores. Even comparing within schools (with a school fixed effect), this socioeconomic gap is 0.8 SD for mathematics, and 0.12 SD for reading comprehension. Consistent with the findings of our two main specifications, students who attended preschool perform on average 0.2 SD better than students who did not, a difference that holds even for those who only attended a PRONOEI (keeping in mind that this is not a causal estimate). What is perhaps surprising, however, is that the interaction term, which denotes how much the socioeconomic gap widens or shrinks for students who went to preschool, is negative, suggesting that the socioeconomic gap actually widens among students who went to preschool. In other words, the average gap between poor and non-poor students is 0.05 s.d larger for students who attended preschool than those who did not (controlling for gender and primary school). This holds for both types of preschools but the gap widens more for students who attended community preschools (0.09 s.d.) rather than regular preschools (0.04 s.d.).

While it is important to note that this effect is not the causal impact of preschool on socioeconomic gaps (since the preschool is not randomly assigned to different socioeconomic groups, nor is orthogonal to them), it is suggestive that for students in the same primary school (and therefore likely to have similar conditions), preschool attendance is not enough to close socioeconomic gaps in early learning outcomes. In fact, preschool appears to widen them so that careful thought should be given to the quality of the preschools, as well as how to boost complementary inputs like nutrition or health outcomes for the most vulnerable students. This is consistent with efforts by certain countries to try to provide holistic early childhood interventions that include health and nutrition components, and not just basic educational ones. Importantly, this widening of the socioe-

Table 7: Socioeconomic Gaps by Preschool Attendance

	Mathematics			Language		
	All (1)	Regular (2)	Community (3)	All (4)	Regular (5)	Community (6)
Poor	-0.081*** (0.012)	-0.088*** (0.012)	-0.071*** (0.015)	-0.119*** (0.012)	-0.126*** (0.012)	-0.104*** (0.015)
Any	0.226*** (0.011)			0.230*** (0.011)		
Any × Poor	-0.053*** (0.013)			-0.047*** (0.013)		
Regular		0.232*** (0.012)			0.238*** (0.011)	
Reg. × Poor		-0.040** (0.013)			-0.034** (0.013)	
Community			0.214*** (0.020)			0.197*** (0.019)
Com. × Poor			-0.093*** (0.022)			-0.078*** (0.021)
Obs.	169043	151200	52645	169093	151248	52673

Note: This table shows the socioeconomic test score gap by preschool attendance. Columns 1 and 4 use any kind of preschool, while columns 2 and 5 restrict it to regular preschools, and Columns 3 and 6 to community preschools. All regressions include school fixed effects and robust standard errors clustered by school. Data come from SIAGIE so the sample is restricted to 2013-2016, when preschool attendance started being recorded.

conomic gap could also reflect differences in preschool quality among poor and non-poor students, which are in turn driving these differences. While here we are controlling for the type of school, it could be that there are serious quality differences even within each type of school. This needs to be explored further.

7 Exploring Mechanisms

Developing countries are starting to roll out large preschool programs in order to improve learning outcomes by investing in early childhood, where investments are likely to have the largest returns. However, many developing countries have large rural populations that live in sparsely populated areas where it may not be cost-effective to build brick and mortar preschools. Like Peru, countries are starting to experiment with alternative methods of delivering preschool education including training local mothers to deliver care in their homes. Can the benefits of early childhood education be achieved through these more cost-effective means?

We explore the variation in the types of preschools built in each town in Peru in order to shed light on the relative contribution of the different early childhood education inputs to learning outcomes. The regular preschools provide a trained teacher and formal school infrastructure, while the community preschools provide day care and interactions with peers, but with much more limited pedagogical training. In Table 3, we find that once you disaggregate the overall effect, the type of preschool matters greatly: while regular preschools impact student learning by around 0.05 SD, there is no effect of community preschools on learning. Using the family fixed estimators we find effects for both types of preschools, but the effect of going to a regular preschool is significantly larger than for a community preschool. This is consistent with the findings of other studies in Peru that have looked at differences between CEI and PRONOEIs and find that students that attend CEIs tend to have better outcomes than students who attended PRONOEIs (Cueto and Diaz, 1999; Diaz, 2006; Cueto et al., 2016). A recent study by the Ministry that tested various abilities of 5 year old students found that those who attended CEIs performed better in numerical, and language skills, and in certain social skills than those in community preschools (Ministerio de Educación del Perú, 2013). While this suggests that the type of preschool matters, this differential impact is not causal. While it could be that these findings reflect a lower quality of community preschools, it is also possible that the underlying conditions of these communities (which are smaller, poorer and more dispersed) make preschool ineffective. Perhaps these students are missing necessary complementary inputs (i.e. parental education or adequate nutrition) such that no type of preschool would work well for them.

Attending one or the other type of preschool is strongly correlated with socioeconomic characteristics of the family. The Ministry study surveyed the parents of students in both CEI and PRONOEIs and found substantial differences in the educational attainment and labor profiles of parents: for example, only 44% of mothers of PRONOEI students had completed *any* post-primary school, compared to 71% for mothers whose children were studying in a regular preschool (Ministerio de Educación del Perú, 2013) [7]. Similarly, parents of PRONOEI students were more than 10 percentage points more likely to be unskilled workers or to not work outside the home, and be poorer as measured by the possession of household goods (Ministerio de Educación del Perú, 2013). As a result, some or perhaps much of the difference in performance between students with access to one over the other type of preschool could relate to their underlying socioeconomic conditions, rather than the school. In fact, a recent randomized study of a similar context in Colombia found that moving students from community preschools to regular preschools had negative results in cognitive development, a positive impact on nutrition, and no impact on socioemotional development (Bernal et al., 2019). This suggests that it is not entirely obvious that regular preschools should outperform community ones.

This second half of the paper exploits the decision rule used to determine which school to build in rural Peru to provide well-identified evidence of the differential benefits of providing a regular preschool relative to a community school. This will shed light on components of a high quality preschool education that can yield the type of dividends that developing country governments are hoping to find.

7.1 Empirical Strategy: RD

The application of a decision rule based on the town population prior to building a school allows us to estimate the differential causal impact of attending a regular preschool relative to a community preschool through a regression discontinuity design.

[7] Breaking down the figures further, 32% of mothers of PRONEI did not complete primary school, while for those of CEI this was only 15%. These figures are similar for father's education

As previously discussed, the population of preschool-age or soon to be preschool-age children in each town at the time of the construction of the preschool determined which type of preschool was built. Teams of surveyors were sent to towns where local school boards suspected that there was unmet preschool demand (preschool age children without access to preschools). These surveyors completed what was called a "demand study" which involved going door to door and counting children who were of preschool age at the time.[8] The data gathered in these demand studies was then used to determine if the town received a CEI (regular preschool) or a PRONOEI (community preschool): towns with 15 or more students received a CEI, while those between 8 and 14 received a PRONOEI.

While the data from the original demand studies was never centralized and is not available in a systematic way, we are able to recreate an approximation of the data using the 2012/2013 rolling census SISFOH. The SISFOH was a one-time rolling census undertaken of almost all rural and most urban households in Peru by the Ministry of Social Protection. Surveyors went door to door in a similar fashion as that used in the demand studies and listed all members of the household with dates of birth. While this is a cross-sectional database, and represents a snapshot of rural Peru in 2012/2013, we are able exploit the availability of dates of birth to estimate the number of children that would have been present in the town (assuming little mobility) at each year up until 2012.

An alternative strategy to recreate the estimate of the number of preschool age students that would have been living in these towns at the time the demand surveys were completed is through the official school enrollment records. While preschool enrollment remains low, enrollment in primary school is universal. Using administrative data on school enrollment, which includes birthdays, we are able to back out the number of preschool age children that would have been living in the town at the time the preschool was created. This strategy will work only for towns with primary school (widely available across most rural areas of Peru so it is reasonable to assume most towns had one), and assuming that students come from the same or nearby town that would have also been considered in the construction of the preschool.

The regression discontinuity design then allows us to use the discontinuity in the treatment generated by the decision rule in order to estimate the causal effect of having a regular preschool over a community preschool. The identification strategy assumes that towns

[8] It is unclear whether the demand studies projected demand going forward or focused on current students.

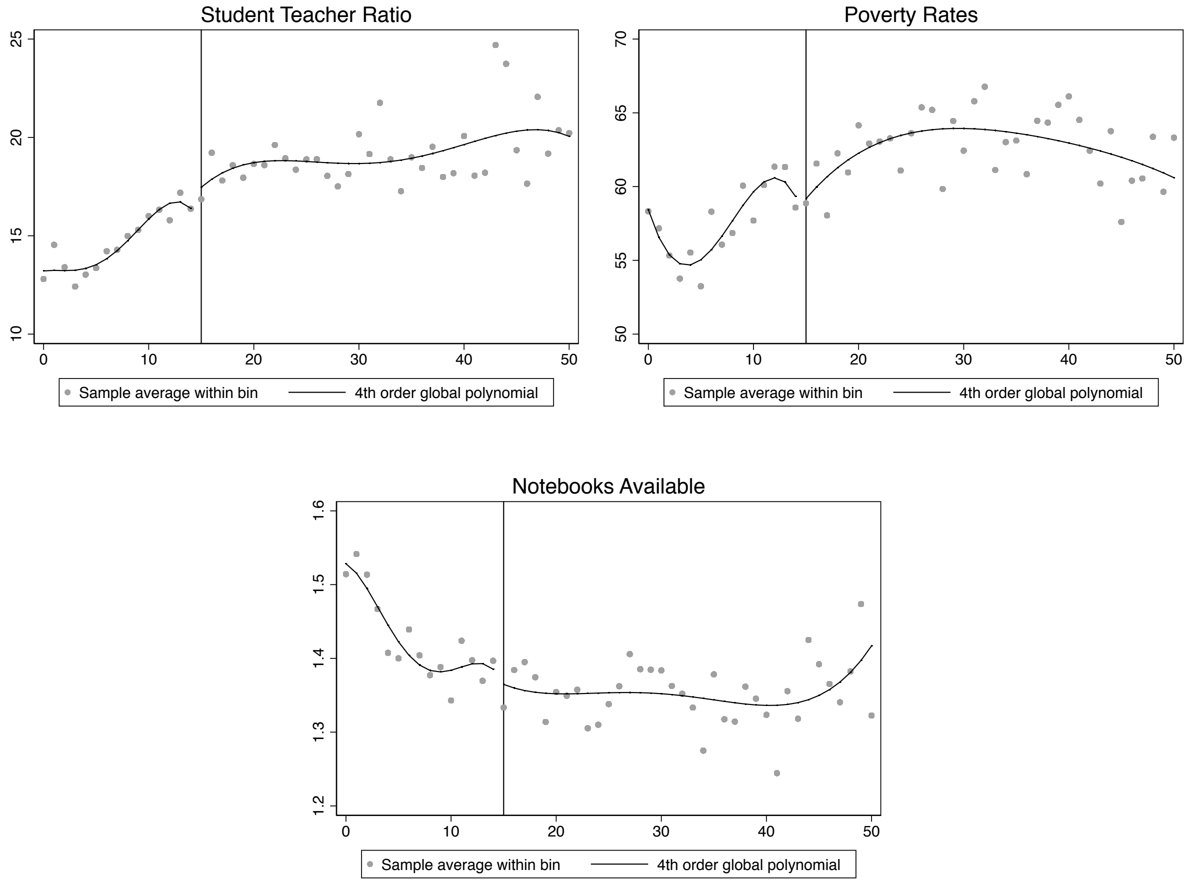


Figure 6: Balance of Covariates

Note: These figures show balance of available covariates across the discontinuity for characteristics of the primary school in the town, as well as socioeconomic characteristics of the town residents. Data come from Censo Escolar.

that are close to the threshold of 15 preschool-age students are similar in all characteristics that are relevant for performance on the standardized exam. Figure 6 shows balance on available characteristics for towns close to the cutoff. It shows that schools around the boundary are balanced on student-teacher ratio, the percentage of students who are in poverty, and the number of workbooks available per student. Figure 7 also shows that the density of the running variable around the cutoff is relatively smooth, which suggests that there is no sorting around the cut-off (while this is possible in practice when the surveyors counted children, the way we are calculating the town population would remove any potential bias, generating an instrument that is not manipulated).

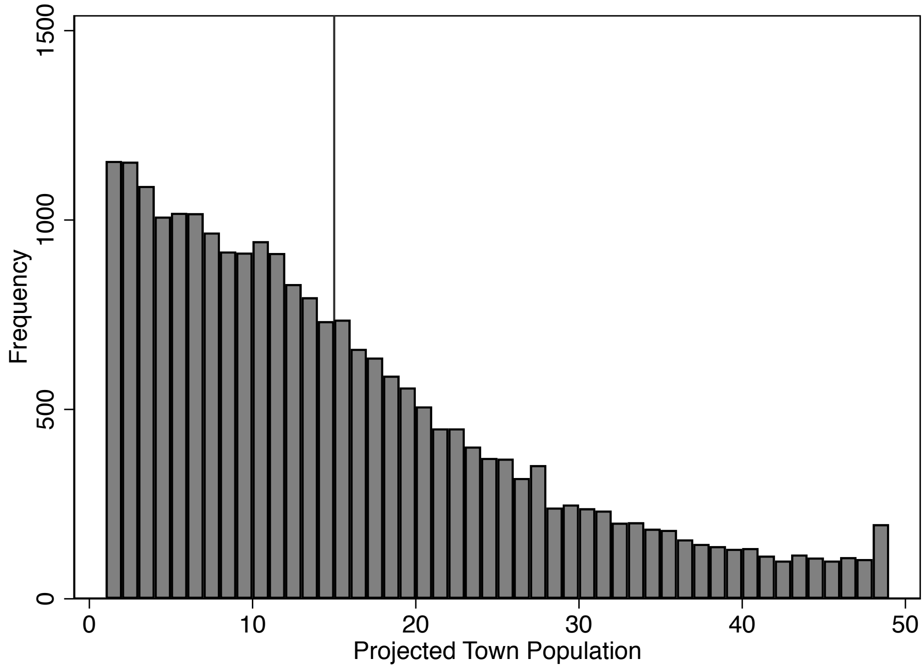


Figure 7: Histogram of Running Variable

Note: This is a histogram of the preschool-aged population of towns two years prior to the preschool being built. The data is constructed from SIAGIE administrative records.

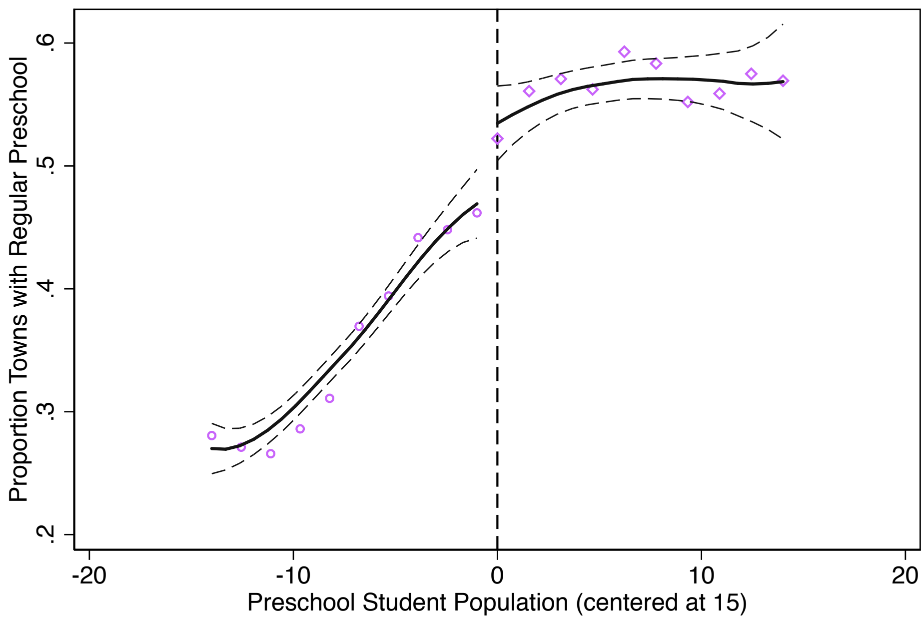


Figure 8: First Stage Graph of Assignment to Preschool by Predicted Town Population

Note: This figure shows the probability of being assigned a regular preschool as a function of the population of the town two years prior to the school construction, showing the discontinuity in assignment around the 15 student threshold.

7.2 RD Results

Figure 8 shows the first stage of the RD using the reconstructed population for each town 2 years before the preschool was built. While the variable is noisy, we are able to see a jump in the treatment assignment of around 8 percentage points, which is statistically significant as shown in the Table 8 with F-statistics above the 10 benchmark for most of the bandwidths presented.

We find some evidence that being assigned to a regular preschool rather than a community preschool increases test scores for students near the cutoff. Table 9 shows the results of the reduced form effect on test scores, using a variety of local linear, nonparametric and polynomial specifications. It is important to note that given the fact that the first stage gets very weak when the bandwidth around the threshold gets very small, we face a tradeoff between losing power close to the threshold, and potential bias as we increase the bandwidth and get farther away from the threshold. For transparency, we have included

Table 8: First Stage Results

	CEI			
	(1)	(2)	(3)	(4)
Running	0.016	0.014***	0.015***	0.015***
	(0.011)	(0.003)	(0.002)	(0.001)
Above	0.072**	0.077***	0.077***	0.086***
	(0.036)	(0.023)	(0.019)	(0.018)
Interact	-0.005	-0.005	-0.009***	-0.009***
	(0.014)	(0.004)	(0.002)	(0.002)
F-Stat	3.96	11.7	19.1	21.9
BW	5	10	15	20
Obs.	2987	5985	8578	9764
R^2	0.112	0.130	0.163	0.170

Note: This table shows the first stage results of the regression discontinuity, showing the probability that a regular preschool gets built in a given town as a function of the population in the town when the decision was made to build it. Column 1 uses a bandwidth of 5, Column 2 of 10, Column 3 of 15 and Column 4 of 20. The dependent variable is whether the town has a CEI built, and the running variable is the preschool age population of the town three years before the preschool is inaugurated.

a variety of bandwidths, which show this tradeoff. Very close to the cutoff, the estimates become very noisy and are not statistically different from zero. Once we use a slightly larger bandwidth, we find positive and significant effects on learning outcomes of around 0.1 SD for both reading comprehension and mathematics. Running a 2SLS in Table 10, we find that a town that received a regular preschool improves learning outcomes by about 0.9 standard deviations relative to students in towns that received a community preschool close to the cut-off, but it is important to note that for many specifications we lack power to be able to provide a precise estimate.

While we therefore find some evidence suggesting that regular preschools are indeed better for student learning than community preschools, it would be helpful to dig up the demand studies to have a stronger first stage in order to be able to estimate these point estimates closer to the threshold.

8 Conclusion

This paper provides quasi-experimental evidence of the impact of a massive expansion of preschool education in Peru on student learning outcomes. We use two different difference-in-difference estimators that take advantage of the progressive rollout of preschool to towns across rural Peru that generates across-town and within-family variation in exposure to preschool. We find that having access to a regular preschool increases learning outcomes as measured by standardized test scores in 2nd grade, while having a community preschool has smaller effects on learning, which are insignificant in our event study specification.

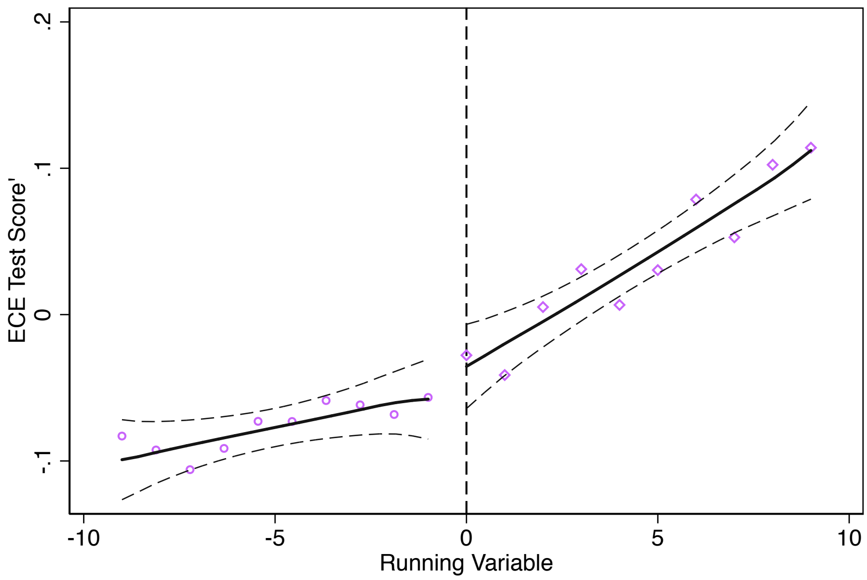
However, in order to causally identify this differential effect , we implement a regression discontinuity approach, which takes advantage of the decision rule that assigned types of preschools based on the preschool-aged population of each town. We find some suggestive evidence of a positive impact of being assigned a normal preschool relative to a community preschool, which suggests that the quality of the preschool inputs (like a trained teacher and appropriate infrastructure) matter for learning outcomes.

Finally, we present evidence that suggests that while preschool is important for learning outcomes, it may actually widen socioeconomic learning gaps, so that other complementary and more focused interventions are needed to help close the gaps. This, of course, given the above findings could also reflect the fact that poorer students attend lower qual-

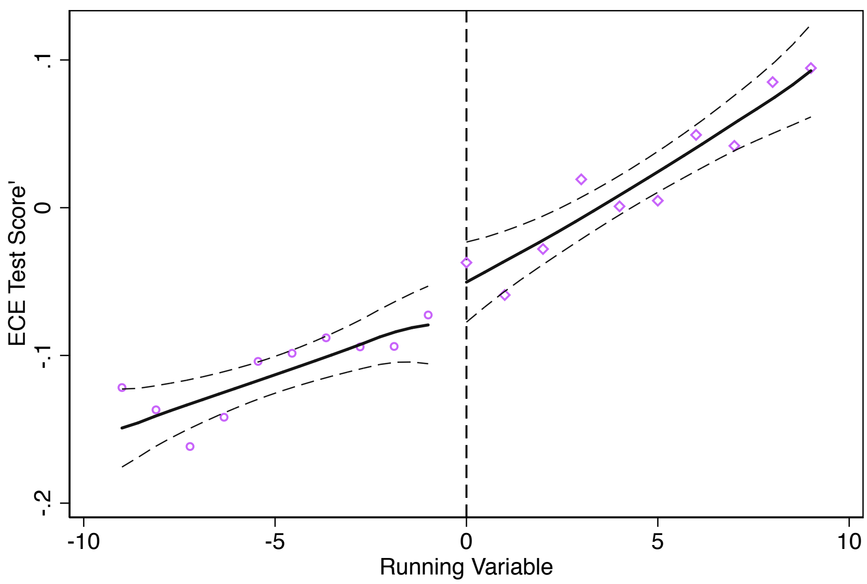
Table 9: Reduced Form Results for Regression Discontinuity

	Language				Mathematics			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Local Linear (Rectangular Kernel)				
Above	-0.004	0.071	0.099*	0.119**	-0.046	0.056	0.100*	0.123**
	(0.089)	(0.060)	(0.051)	(0.048)	(0.094)	(0.064)	(0.053)	(0.051)
R^2	0.053	0.052	0.050	0.047	0.021	0.019	0.019	0.018
				Local Linear (Epanechnikov Kernel)				
Above	0.023	0.049	0.087*	0.104**	-0.024	0.023	0.085	0.109**
	(0.092)	(0.062)	(0.053)	(0.049)	(0.096)	(0.066)	(0.056)	(0.052)
R^2	0.051	0.050	0.050	0.049	0.019	0.019	0.019	0.018
				Quadratic Polynomial				
Above	0.236	-0.009	0.039	0.057	0.189	-0.075	0.009	0.027
	(0.189)	(0.095)	(0.075)	(0.070)	(0.195)	(0.100)	(0.079)	(0.074)
R^2	0.054	0.052	0.050	0.048	0.022	0.020	0.019	0.018
				Cubic Polynomial				
Above	-0.611	0.107	0.016	0.002	-0.341	0.027	-0.088	-0.074
	(0.516)	(0.154)	(0.107)	(0.100)	(0.557)	(0.159)	(0.113)	(0.105)
R^2	0.059	0.053	0.050	0.048	0.024	0.020	0.020	0.019
Observations	2321	4214	5597	6357	2321	4215	5598	6358
Towns	2139	3740	4826	5392	2139	3741	4827	5393
Bandwidth	5	10	15	20	5	10	15	20

Note: This table shows the reduced form results of the regression discontinuity of being assigned a regular preschool compared to a community preschool on learning outcomes. Each panel shows a different specification (local linear or polynomial) and each column shows results by different bandwidths. Robust standard errors clustered by town. Scores are standardized so they can be interpreted as standard deviations.



(a) Mathematics



(b) Reading Comprehension

Figure 9: Reduced form of Preschool Assignment on Test Scores

Note: This figure shows the average test score of the students in a town as a function of the running variable (population of the town two years prior to the school construction). Test scores are standardized and standard errors are clustered by town.

Table 10: Two Stage Least Square Regression Discontinuity

	Language				Mathematics			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Local Linear (Rectangular Kernel)				
CEI	-0.053	0.835	0.949*	1.008**	-0.560	0.657	0.962*	1.040**
	(1.089)	(0.772)	(0.542)	(0.450)	(1.163)	(0.786)	(0.564)	(0.470)
				Local Linear (Epanechnikov Kernel)				
CEI	0.238	0.593	0.898	0.967*	-0.257	0.281	0.878	1.009*
	(0.983)	(0.788)	(0.605)	(0.505)	(1.009)	(0.799)	(0.628)	(0.528)
				Quadratic Polynomial				
CEI	1.288	-0.106	0.500	0.743	1.031	-0.894	0.109	0.338
	(1.229)	(1.125)	(0.996)	(1.007)	(1.193)	(1.255)	(1.013)	(1.004)
				Cubic Polynomial				
CEI	16.158	0.980	0.194	0.042	9.011	0.247	-1.078	-0.881
	(104.228)	(1.584)	(1.338)	(1.210)	(59.821)	(1.476)	(1.489)	(1.322)
Obs.	2321	4213	5595	6356	2321	4214	5596	6357
BW	5	10	15	20	5	10	15	20

Note: This table shows the two stage least squares results of the regression discontinuity of being assigned a regular preschool compared to a community preschool on learning outcomes. Each panel shows a different specification (local linear or polynomial) and each column shows results by different bandwidths. The first stage results are shown in Table 8. Robust standard errors clustered by town. Scores are standardized so they can be interpreted as standard deviations.

ity preschools, or that preschool attendance is complementary to other family inputs like nutrition and early childhood stimulation. In this case, efforts should be made to complement those inputs for poorer students in order to close early socioeconomic achievement gaps.

Ultimately, this paper provides evidence that preschool expansion, even when implemented at scale by the government in vulnerable communities, has modest impacts on learning outcomes but that quality of the preschool matters. This should make us cautious about extrapolating from intensive and comprehensive programs in developing countries, but also should encourage governments to improve not only the access to preschool, but also its quality, particularly for the most vulnerable students.

References

Behrman, J. R., Cheng, Y., and Todd, P. E. (2004). Evaluating preschool programs when length of exposure to the program varies: A nonparametric approach. *The Review of Economics and Statistics*, 86(1):108–132.

Berlinski, S. and Galiani, S. (2007). The effect of a large expansion of pre-primary school facilities on preschool attendance and maternal employment. *Labour Economics*, 14(3):665–680.

Berlinski, S., Galiani, S., and Gertler, P. (2009). The effect of pre-primary education on primary school performance. *Journal of Public Economics*, 93(1-2):219–234.

Berlinski, S., Galiani, S., and Manacordad, M. (2008). Giving children a better start: Preschool attendance and school-age profiles. *Journal of Public Economics*, 92:1416–1440.

Berlinski, S., Galiani, S., and McEwan, P. J. (2011). Preschool and Maternal Labor Market Outcomes: Evidence from a Regression Discontinuity Design. *Economic Development and Cultural Change*, 59(2):313–344.

Bernal, R., Attanasio, O., Peña, X., and Vera-Hernández, M. (2019). The effects of the transition from home-based childcare to childcare centers on children's health and development in Colombia. *Early Childhood Research Quarterly*, (47):418–431.

Bernal, R. and Fernandez, C. (2013). Subsidized childcare and child development in Colombia: Effects of Hogares Comunitarios de Bienestar as a function of timing and length of exposure. *Social Science and Medicine*, 97:241–249.

Britto, P. R., Lye, S. J., Proulx, K., Yousafzai, A. K., Matthews, S. G., Vaivada, T., Perez-Escamilla, R., Rao, N., Ip, P., Fernald, L. C. H., MacMillan, H., Hanson, M., Wachs, T. D., Yao, H., Yoshikawa, H., Cerezo, A., Leckman, J. F., and Bhutta, Z. A. (2017). Nurturing care: promoting early childhood development. *The Lancet*, 389(10064):91 – 102.

Cueto, S. and Diaz, J. J. (1999). Impacto de la educacion inicial en el rendimiento en primer grado de primaria en escuelas publicas urbanas de lima. *Revista de Psciolgia de la Universidad Catolica del Peru*, XVII(1):74–91.

Cueto, S., León, J., Miranda, A., Dearden, K., Crookston, B. T., and Behrman, J. R. (2016). Does pre-school improve cognitive abilities among children with early-life stunting? a longitudinal study for peru. *International Journal of Educational Research*, 75:102 – 114.

Cunha, F., Heckman, J. J., Lochner, L. J., and Masterov, D. V. (2006). Interpreting the evidence on life cycle skill formation. In Welch, E. A. H. . F., editor, *Handbook of the Economics of Education*, page 697–812. North-Holland, Amsterdam.

de Chaisemartin, C. and D'Haultfœuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, 110(9):2964–96.

Deming, D. (2009). Early Childhood Intervention and Life-Cycle Skill Development: Evidence from Head Start. *American Economic Journal: Applied Economics*, 1(3):111–134.

Diaz, J. J. (2006). Pre-school education and schooling outcomes in peru.

Duflo, E. (2001). Schooling and Labor Market Consequences of Schooling Construction in Indonesia: Evidence from an Unusual Policy Experiment. *American Economic Review*, 91(4):795–81.

Engle, P., Fernald, L. C., Alderman, H., Behrman, J., O'Gara, C., Yousafzai, A., de Mello, M. C., Hidrobo, M., Ulkuer, N., Ertem, I., and Ilthus, S. (2011). Strategies for reducing inequalities and improving developmental outcomes for young children in low-income and middle-income countries. *The Lancet*, 378(9799):1339–1353.

Hallman, K., Quisumbing, A. R., Ruel, M., and de la Brière, B. (2005). Mothers' work and child care: Findings from the urban slums of guatemala city. *Economic Development and Cultural Change*, 53(4):855–885.

Heckman, J. J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science*, 312:1900–1902.

Heckman, J. J., Moon, S. H., Pinto, R., Savelyez, P., and Yavitz, A. (2010). The Rate of Return to the HighScope Perry Preschool Program. *Journal of Public Economics*, 94(1-2):114–128.

McCoy, D. C., Yoshikawa, H., Ziol-Guest, K. M., Duncan, G. J., Schindler, H. S., Magnuson, K., Yang, R., Koepp, A., and Shonkoff, J. P. (2017). Impacts of Early Childhood

Education on Medium- and Long-Term Educational Outcomes. *Educational Researcher*, 46(8):474–487.

Ministerio de Educación del Perú (2013). Estudio de educación inicial: Un acercamiento a las niñas y los niños de cinco años de edad. *Serie Investigaciones MINEDU*.

Nores, M. and Barnett, W. S. (2010). Benefits of early childhood interventions across the world: (under) investing in the very young. *Economics of Education Review*, 29(2):271 – 282. Special Issue in Honor of Henry M. Levin.

Richter, L. M., Daelmans, B., Lombardi, J., Heymann, J., Boo, F. L., Behrman, J. R., Lu, C., Lucas, J. E., Perez-Escamilla, R., Dua, T., Bhutta, Z. A., Stenberg, K., Gertler, P., and Darmstadt, G. L. (2017). Investing in the foundation of sustainable development: pathways to scale up for early childhood development. *The Lancet*, 389(10064):103 – 118.

Schady, N., Behrman, J., Araujo, M. C., Azuero, R., Bernal, R., Bravo, D., Lopez-Boo, F., Macours, K., Marshall, D., Paxson, C., and Vakis, R. (2015). Wealth Gradients in Early Childhood Cognitive Development in Five Latin American Countries. *Journal of Human Resources*, 50(2):446–463.

Schweinhart, L. J., Montie, J., Xiang, Z., Barnett, W., Belfield, C., and Nores, M. (2005). *Lifetime effects: The High/Scope Perry Preschool Study through age 40*. High/Scope Press, Ypsilanti MI.

Yoshikawa, H., Weiland, C., Brooks-Gunn, J., Burchinal, M. R., Espinosa, L. M., Gormley, W. T., Ludwig, J., Magnuson, K. A., Phillips, D., and Zaslow, M. J. (2013). Investing in Our Future: The Evidence Base on Preschool Education. *Foundation for Child Development Report*.

A Supplementary Material

Table A.1: Summary Statistics of Students by allprek

	Never	Changers	Always
Language score ECE 2P	478.27	486.24	556.21
	(97.97)	(86.69)	(90.54)
Mathematics score ECE 2P	489.73	498.71	549.31
	(115.91)	(110.76)	(113.69)
Language score ECE 4P	438.19	416.46	490.33
	(100.38)	(88.01)	(93.90)
Mathematics score ECE 4P	438.30	421.87	482.08
	(102.02)	(92.53)	(93.90)
Language score ECE 2S	523.64	518.77	577.89
	(64.55)	(57.30)	(69.42)
Mathematics score ECE 2S	518.82	515.42	570.63
	(76.04)	(69.70)	(85.39)
Social Science score ECE 2S	455.64	450.91	509.44
	(91.06)	(84.75)	(96.75)
Science and Tech score ECE 2S	459.03	453.72	511.97
	(99.55)	(93.69)	(97.38)

Note: appendix.

Figure A.1: Graphical Representation of Empirical Strategy

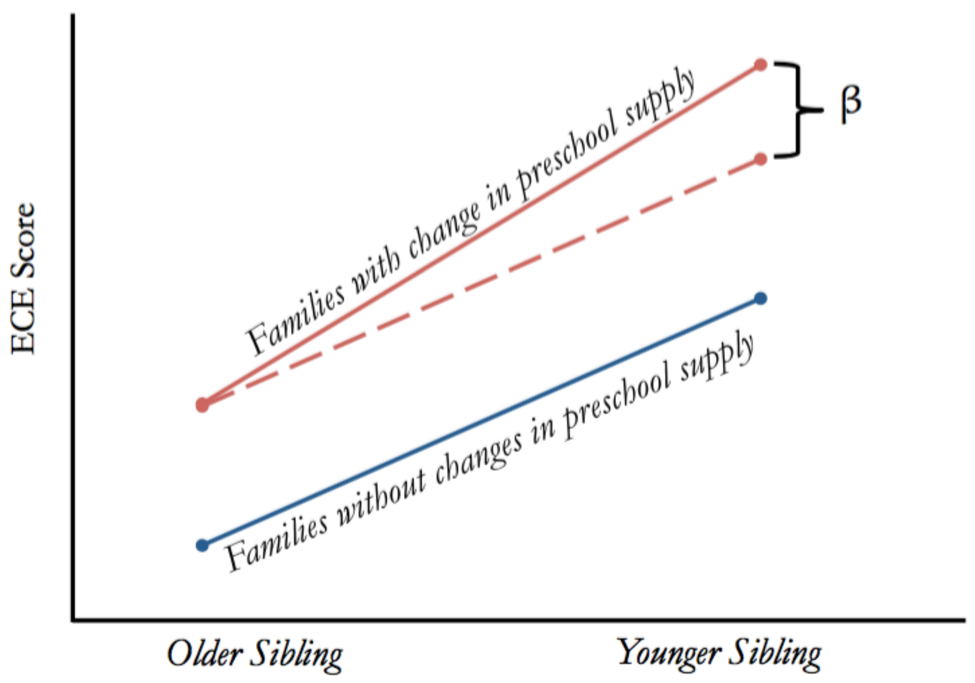
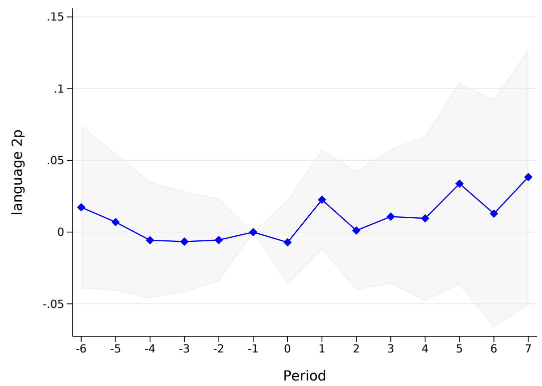
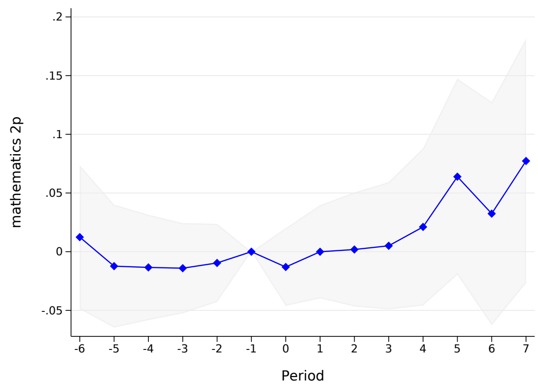


Figure A.2: Effect of any preschool on learning outcomes

(a) Reading

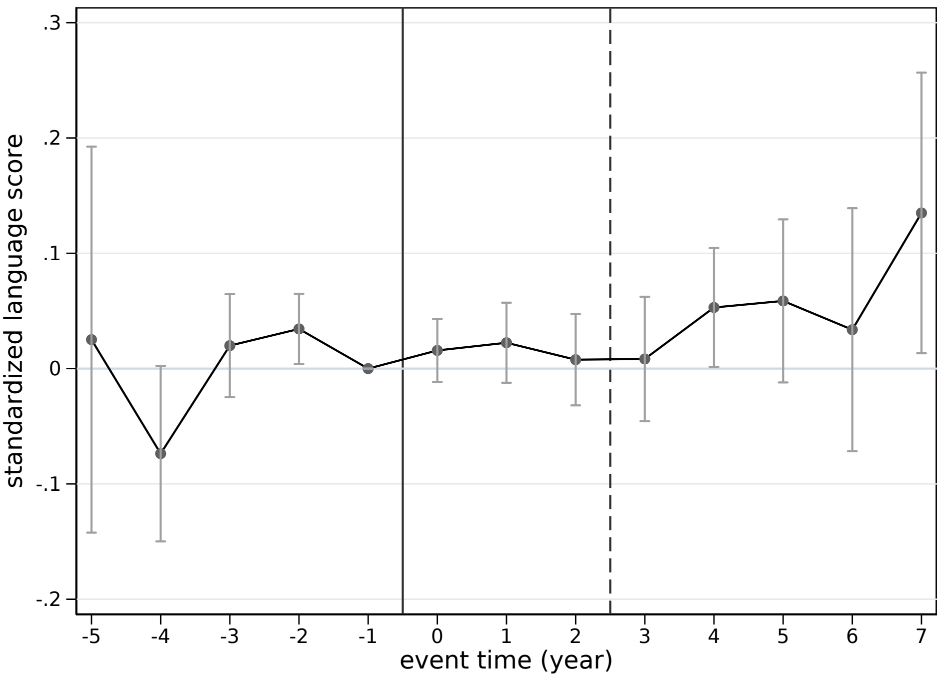


(b) Mathematics

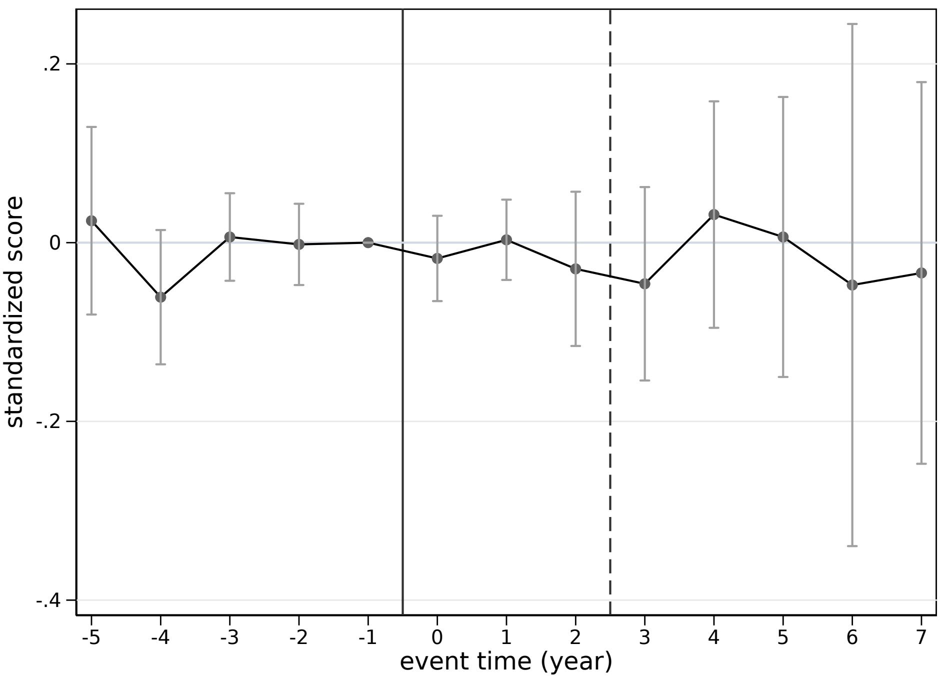


Sample of towns exposed to preschool any time after 2003.

Figure A.3: TWFE de Chaisemartin and D'Haultfoeuille estimator



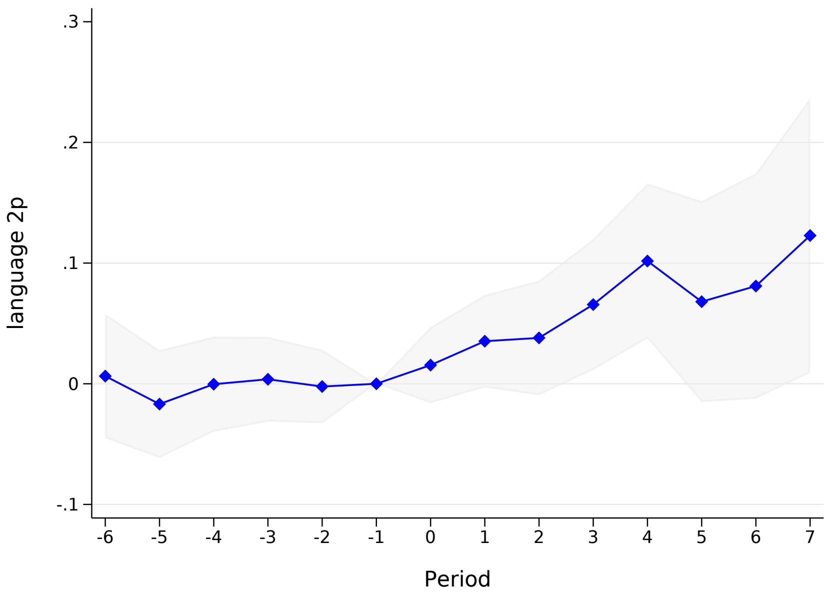
(a) Reading



(b) Mathematics

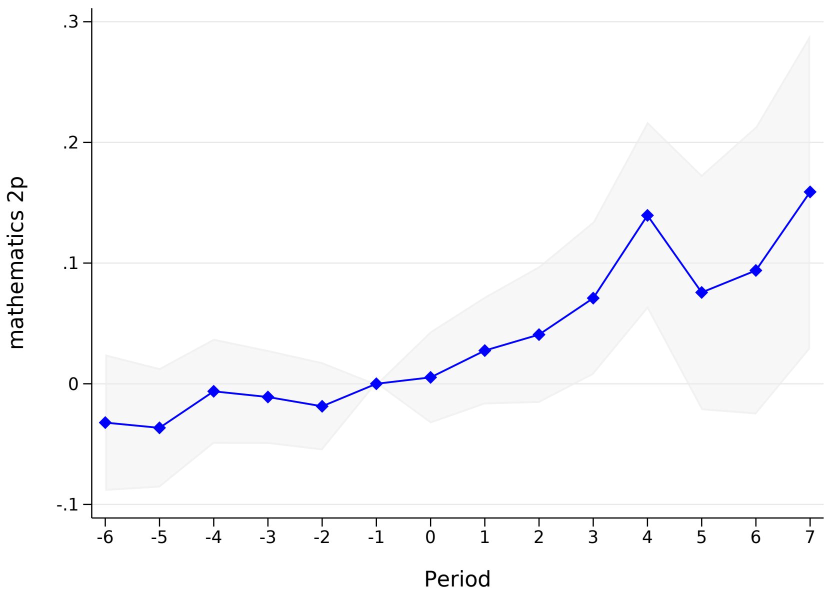
Note: This figures shows the dynamic TWFE estimator from de Chaisemartin and D'Haultfoeuille (2020). Outcomes are test scores at second grade of primary school in reading and mathematics. Standard errors are clustered at town level.

Figure A.4: Event Study: Reading CEI no controls



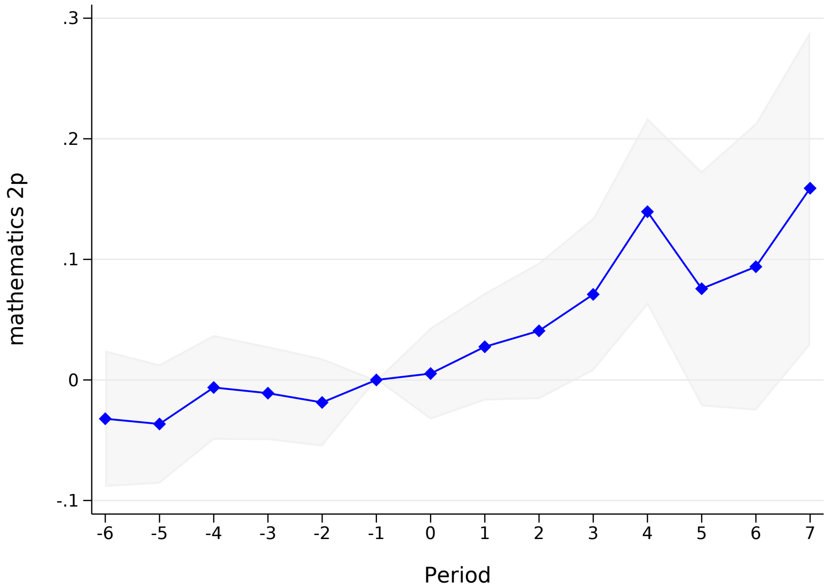
Note:

Figure A.5



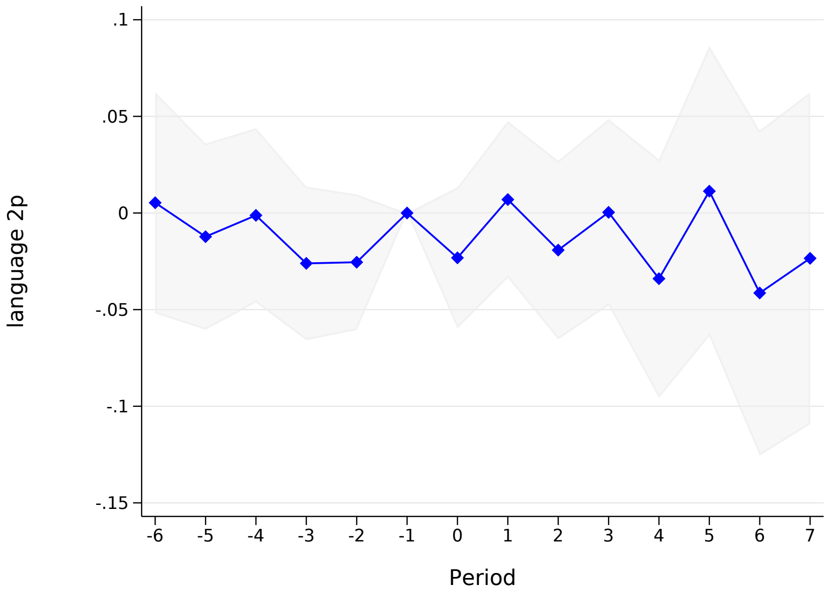
Note:

Figure A.6: Event Study: Math CEI no controls



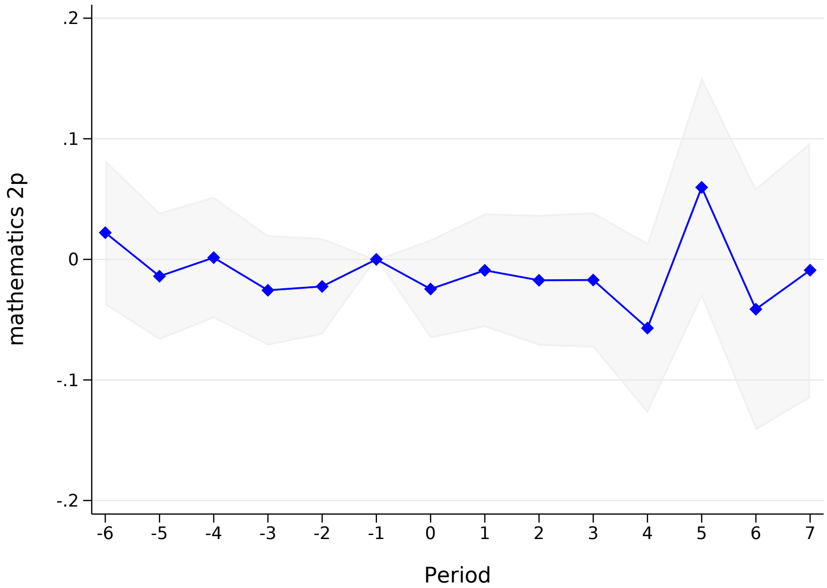
Note:

Figure A.7: Event Study: Reading PRONOEI no controls



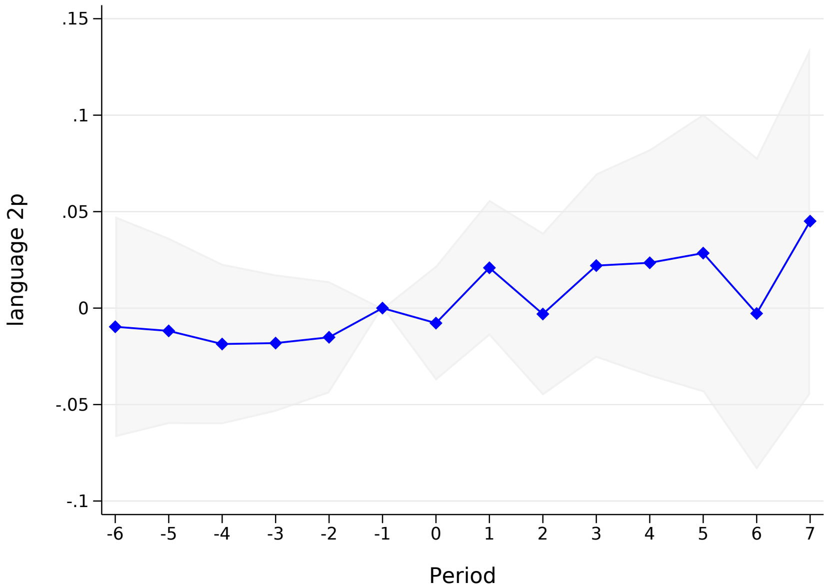
Note:

Figure A.8: Event Study: Math PRONOEI no controls



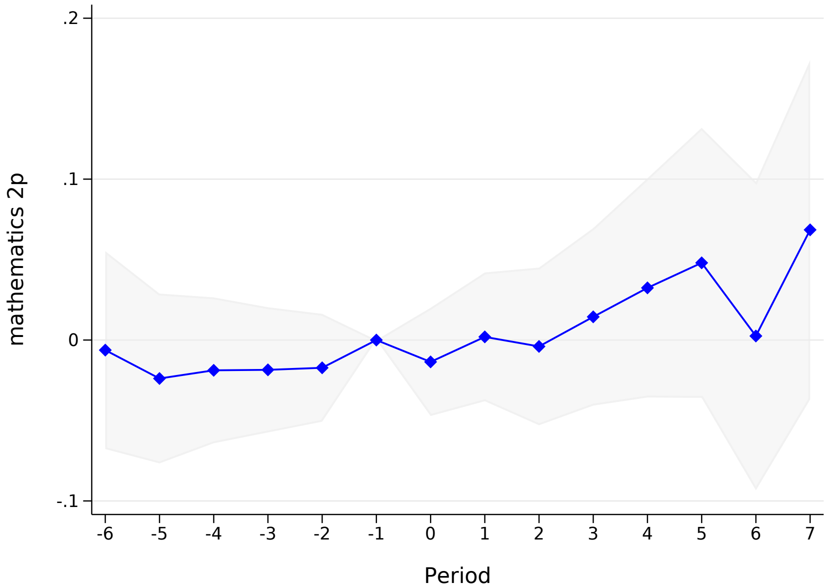
Note:

Figure A.9: Event Study: Reading all preschools no controls



Note:

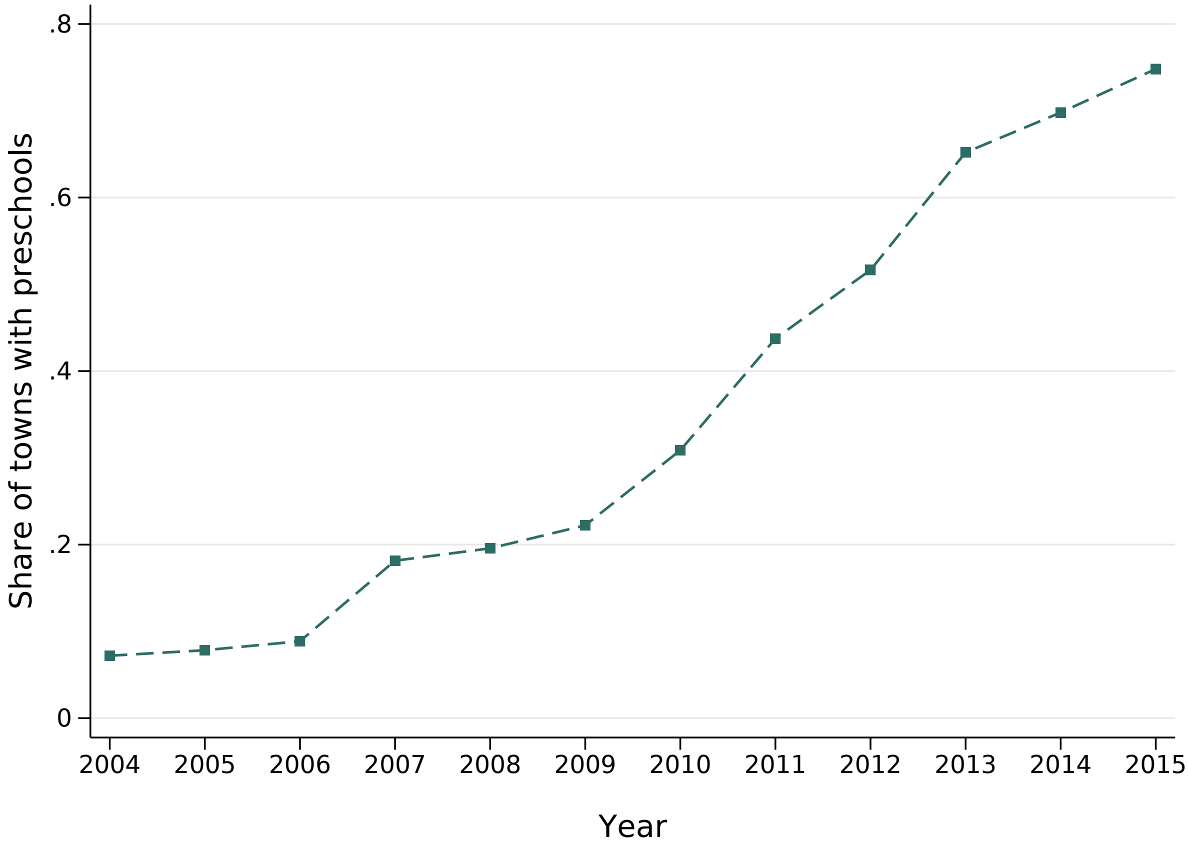
Figure A.10: Event Study: Math all preschools no controls



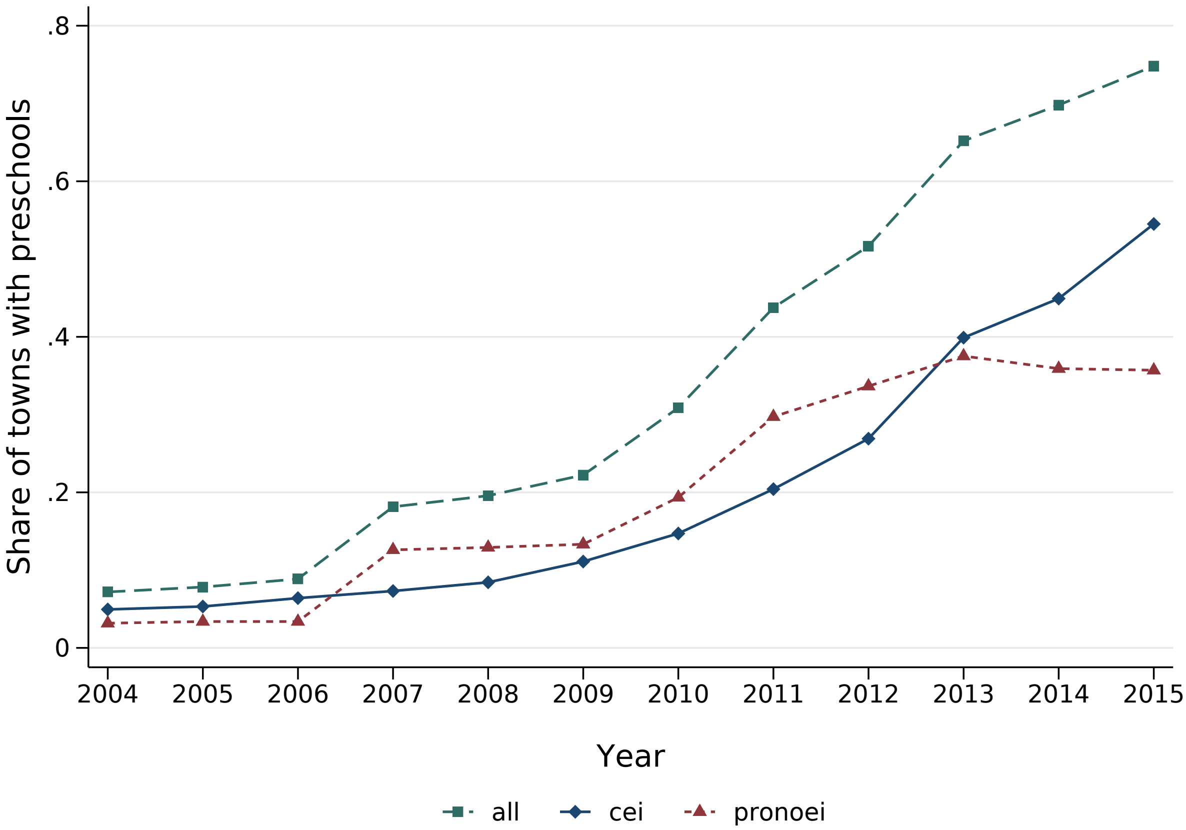
Note:

B Material not for Publication

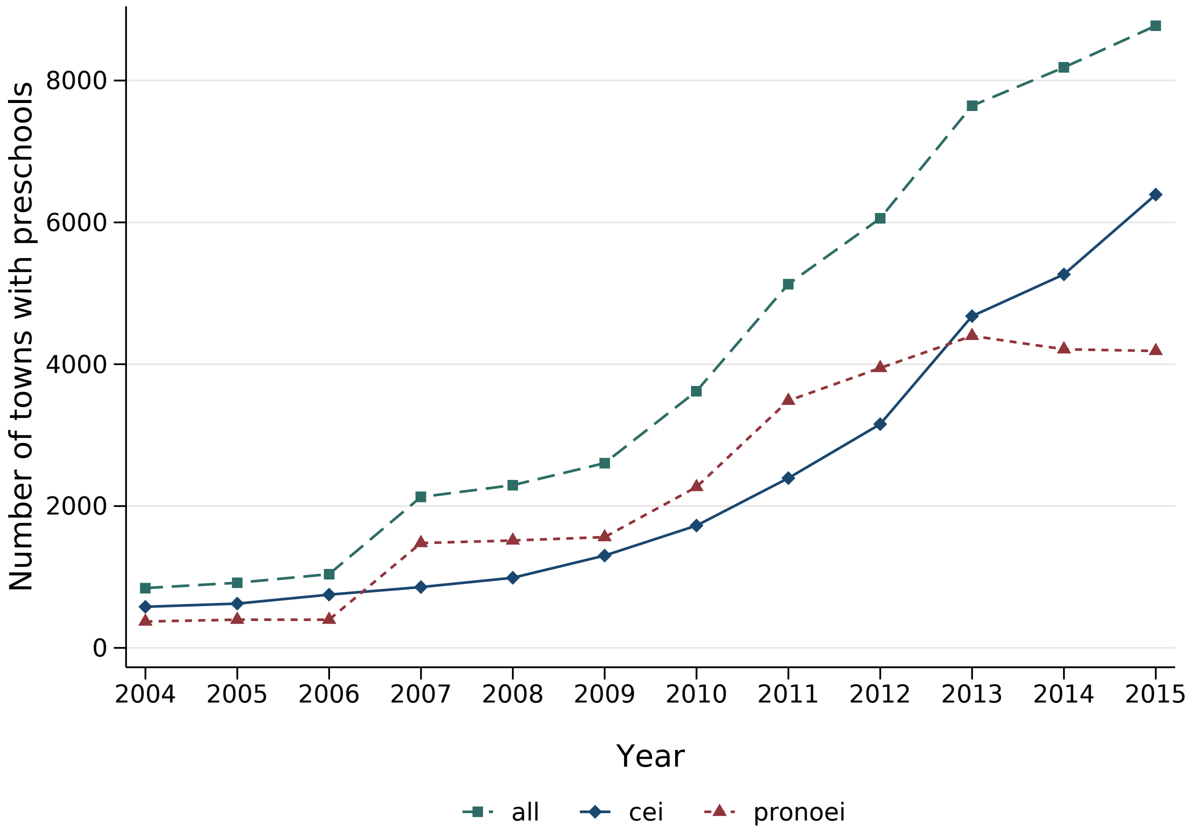
B.1 Summary Statistics



Note:


Share of towns with preschools
.8
.6
.4
.2
0
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
Year
all
cei
pronoei

Note:


Number of towns with preschools
8000
6000
4000
2000
0
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
Year
all
cei
pronoei

Note:

Table B.2: Summary Statistics of Students by cei

	Never	Changers	Always
student is female	0.49	0.49	0.49
	(0.50)	(0.50)	(0.50)
student age	9.55	9.00	8.72
	(2.93)	(2.92)	(2.84)
family size	5.58	5.61	5.13
	(1.87)	(1.87)	(1.81)
number of kids	3.54	3.59	2.98
	(1.74)	(1.75)	(1.59)
no. of siblings	1.93	2.12	1.65
	(0.95)	(1.03)	(0.79)
monther no educ	0.22	0.25	0.07
	(0.42)	(0.44)	(0.25)
monther primary educ	0.59	0.59	0.31
	(0.49)	(0.49)	(0.46)
monther secondary educ	0.19	0.15	0.63
	(0.39)	(0.36)	(0.48)
father no educ	0.08	0.08	0.02
	(0.27)	(0.28)	(0.14)
father primary educ	0.62	0.65	0.26
	(0.49)	(0.48)	(0.44)
father secondary educ	0.31	0.27	0.72
	(0.46)	(0.44)	(0.45)
mother mainly doing housework	0.84	0.87	0.65
	(0.36)	(0.33)	(0.48)
father mainly doing housework	0.02	0.01	0.01
	(0.14)	(0.11)	(0.09)
Poverty SISFOH category	1.91	1.88	2.13
	(0.92)	(0.91)	(0.91)
school at walk distance	0.82	0.87	0.72
	(0.39)	(0.34)	(0.45)

Note: .

Table B.3: Summary Statistics of Students by pronoei

	Never	Changers	Always
student is female	0.49	0.49	0.49
	(0.50)	(0.50)	(0.50)
student age	9.02	8.79	8.70
	(2.91)	(2.86)	(2.84)
family size	5.34	5.29	5.11
	(1.82)	(1.82)	(1.84)
number of kids	3.27	3.21	2.94
	(1.67)	(1.65)	(1.60)
no. of siblings	1.87	1.81	1.60
	(0.93)	(0.89)	(0.76)
monther no educ	0.16	0.13	0.05
	(0.37)	(0.33)	(0.22)
monther primary educ	0.50	0.44	0.26
	(0.50)	(0.50)	(0.44)
monther secondary educ	0.34	0.43	0.69
	(0.47)	(0.50)	(0.46)
father no educ	0.05	0.04	0.01
	(0.23)	(0.21)	(0.12)
father primary educ	0.49	0.42	0.21
	(0.50)	(0.49)	(0.40)
father secondary educ	0.46	0.53	0.78
	(0.50)	(0.50)	(0.41)
mother mainly doing housework	0.79	0.77	0.61
	(0.40)	(0.42)	(0.49)
father mainly doing housework	0.01	0.01	0.01
	(0.12)	(0.10)	(0.09)
Poverty SISFOH category	1.95	1.95	2.19
	(0.92)	(0.91)	(0.89)
school at walk distance	0.81	0.76	0.70
	(0.39)	(0.43)	(0.46)

Note: .

Table B.4: Summary Statistics of Students by allprek

	Never	Changers	Always
student is female	0.50	0.49	0.49
	(0.50)	(0.50)	(0.50)
student age	9.69	9.11	8.74
	(2.93)	(2.94)	(2.85)
family size	5.58	5.64	5.16
	(1.89)	(1.87)	(1.82)
number of kids	3.52	3.62	3.00
	(1.76)	(1.76)	(1.60)
no. of siblings	1.88	2.11	1.66
	(0.93)	(1.03)	(0.81)
monther no educ	0.23	0.26	0.07
	(0.42)	(0.44)	(0.26)
monther primary educ	0.56	0.60	0.32
	(0.50)	(0.49)	(0.47)
monther secondary educ	0.21	0.14	0.60
	(0.41)	(0.35)	(0.49)
father no educ	0.08	0.09	0.02
	(0.27)	(0.28)	(0.15)
father primary educ	0.60	0.67	0.27
	(0.49)	(0.47)	(0.45)
father secondary educ	0.32	0.24	0.70
	(0.47)	(0.43)	(0.46)
mother mainly doing housework	0.83	0.88	0.66
	(0.38)	(0.32)	(0.47)
father mainly doing housework	0.02	0.01	0.01
	(0.13)	(0.12)	(0.10)
Poverty SISFOH category	1.90	1.86	2.12
	(0.93)	(0.92)	(0.91)
school at walk distance	0.76	0.87	0.73
	(0.42)	(0.33)	(0.45)

Note: .

Table B.5: Summary Statistics of Students by allprek

	Never	Changers	Always
Language score ECE 2P	478.27	486.24	556.21
	(97.97)	(86.69)	(90.54)
Mathematics score ECE 2P	489.73	498.71	549.31
	(115.91)	(110.76)	(113.69)
Language score ECE 4P	438.19	416.46	490.33
	(100.38)	(88.01)	(93.90)
Mathematics score ECE 4P	438.30	421.87	482.08
	(102.02)	(92.53)	(93.90)
Language score ECE 2S	523.64	518.77	577.89
	(64.55)	(57.30)	(69.42)
Mathematics score ECE 2S	518.82	515.42	570.63
	(76.04)	(69.70)	(85.39)
Social Science score ECE 2S	455.64	450.91	509.44
	(91.06)	(84.75)	(96.75)
Science and Tech score ECE 2S	459.03	453.72	511.97
	(99.55)	(93.69)	(97.38)

Note: appendix.

Table B.6: Summary Statistics of Students by cei

	Never	Changers	Always
Language score ECE 2P	481.79	490.69	559.03
	(92.54)	(86.07)	(89.72)
Mathematics score ECE 2P	493.84	502.32	551.33
	(113.67)	(110.19)	(113.42)
Language score ECE 4P	429.52	421.32	492.41
	(94.52)	(88.00)	(93.42)
Mathematics score ECE 4P	431.40	426.83	483.73
	(96.87)	(92.92)	(93.52)
Language score ECE 2S	523.03	520.26	580.15
	(60.79)	(58.05)	(68.94)
Mathematics score ECE 2S	518.85	517.19	572.72
	(72.12)	(70.96)	(85.34)
Social Science score ECE 2S	456.12	452.58	511.63
	(87.55)	(85.73)	(96.59)
Science and Tech score ECE 2S	458.21	455.42	514.20
	(95.50)	(93.97)	(96.96)

Note: .

Table B.7: Summary Statistics of Students by pronoei

	Never	Changers	Always
Language score ECE 2P	512.55	532.58	566.72
	(94.91)	(90.75)	(87.92)
Mathematics score ECE 2P	518.32	534.04	556.16
	(116.17)	(113.81)	(112.41)
Language score ECE 4P	453.99	465.76	499.91
	(96.00)	(92.23)	(92.51)
Mathematics score ECE 4P	452.77	463.00	489.56
	(97.70)	(93.88)	(92.34)
Language score ECE 2S	545.22	556.74	586.49
	(68.29)	(66.82)	(68.26)
Mathematics score ECE 2S	540.10	551.45	578.52
	(81.80)	(81.02)	(85.32)
Social Science score ECE 2S	479.08	489.48	517.08
	(94.10)	(93.32)	(96.89)
Science and Tech score ECE 2S	481.22	490.76	520.26
	(98.92)	(96.14)	(96.32)

Note: .

Table B.8: Summary Statistics of Students by pronoei

	Never	Changers	Always
none w/ any prek	0.14	0.00	0.00
	(0.34)	(0.00)	(0.00)
some w/ any prek	0.12	0.28	0.00
	(0.32)	(0.45)	(0.00)
all w/ any prek	0.74	0.72	1.00
	(0.44)	(0.45)	(0.00)
no. of any prek in pre-policy	4.80	14.05	101.23
	(11.53)	(25.17)	(121.63)
no. of CEI in pre-policy	4.31	13.76	74.75
	(10.67)	(25.13)	(104.62)
no. of PRONOEI in pre-policy	0.49	0.29	26.47
	(2.27)	(0.93)	(33.16)
no. of prek in post-policy	5.29	20.00	75.65
	(10.45)	(27.02)	(66.12)
no. of CEI in post-policy	4.51	14.83	46.17
	(8.89)	(20.16)	(32.42)
no. of PRONOEI in post-policy	0.78	5.17	29.48
	(2.59)	(8.24)	(37.53)
school at walk distance	0.81	0.76	0.70
	(0.39)	(0.43)	(0.46)

Note: .

Table B.9: Summary Statistics of Students by cei

	Never	Changers	Always
none w/ any prek	0.49	0.00	0.00
	(0.50)	(0.00)	(0.00)
some w/ any prek	0.27	0.68	0.00
	(0.44)	(0.47)	(0.00)
all w/ any prek	0.25	0.32	1.00
	(0.43)	(0.47)	(0.00)
no. of any prek in pre-policy	1.31	0.90	76.97
	(4.56)	(3.28)	(111.74)
no. of CEI in pre-policy	0.73	0.35	57.58
	(3.26)	(2.20)	(94.27)
no. of PRONOEI in pre-policy	0.58	0.55	19.39
	(1.70)	(1.47)	(30.62)
no. of prek in post-policy	2.58	2.17	59.11
	(6.07)	(4.57)	(63.69)
no. of CEI in post-policy	1.70	1.79	36.77
	(4.90)	(3.71)	(32.88)
no. of PRONOEI in post-policy	0.88	0.38	22.34
	(1.66)	(1.44)	(34.29)
school at walk distance	0.82	0.87	0.72
	(0.39)	(0.34)	(0.45)

Note: .

Table B.10: Summary Statistics of Students by allprek

	Never	Changers	Always
none w/ any prek	1.00	0.00	0.00
	(0.00)	(0.00)	(0.00)
some w/ any prek	0.00	1.00	0.00
	(0.00)	(0.00)	(0.00)
all w/ any prek	0.00	0.00	1.00
	(0.00)	(0.00)	(0.00)
no. of any prek in pre-policy	1.60	0.41	73.75
	(5.88)	(2.79)	(110.38)
no. of CEI in pre-policy	1.19	0.27	55.13
	(4.28)	(2.00)	(92.95)
no. of PRONOEI in pre-policy	0.41	0.13	18.62
	(1.90)	(0.96)	(30.18)
no. of prek in post-policy	2.93	1.98	56.68
	(7.78)	(3.99)	(63.36)
no. of CEI in post-policy	2.23	1.51	35.26
	(6.46)	(3.21)	(32.95)
no. of PRONOEI in post-policy	0.70	0.47	21.42
	(1.72)	(1.23)	(33.83)
school at walk distance	0.76	0.87	0.73
	(0.42)	(0.33)	(0.45)

Note: .

Table B.11: Summary Statistics of Students by exposure to preK

	No PreK	Any PreK
student is female	0.47	0.50
	(0.499)	(0.500)
student age	9.65	7.97
	(2.545)	(2.596)
family size	5.70	5.10
	(1.916)	(1.795)
number of kids	3.62	2.96
	(1.774)	(1.587)
no. of siblings	2.00	1.67
	(0.978)	(0.808)
monther no educ	0.21	0.07
	(0.410)	(0.258)
monther primary educ	0.54	0.33
	(0.499)	(0.469)
monther secondary educ	0.25	0.60
	(0.433)	(0.490)
father no educ	0.07	0.02
	(0.256)	(0.149)
father primary educ	0.56	0.28
	(0.496)	(0.450)
father secondary educ	0.37	0.70
	(0.483)	(0.460)
mother mainly doing housework	0.80	0.68
	(0.399)	(0.467)
father mainly doing housework	0.01	0.01
	(0.110)	(0.097)
Poverty SISFOH category	1.78	2.13
	(0.894)	(0.905)
school at walk distance	0.75	0.73
	(0.434)	(0.442)
Language score ECE 2P	506.46	569.67

	(85.269)	(87.983)
Mathematics score ECE 2P	508.43	567.59
	(108.924)	(119.386)
Language score ECE 4P	429.20	490.10
	(88.838)	(94.721)
Mathematics score ECE 4P	429.74	482.48
	(93.081)	(94.507)
Language score ECE 2S	532.56	578.77
	(62.922)	(70.331)
Mathematics score ECE 2S	526.38	568.06
	(69.637)	(80.652)
Social Science score ECE 2S	464.23	510.76
	(88.722)	(96.647)
Science and Tech score ECE 2S	467.98	511.64
	(93.021)	(96.924)
none w/ any prek	0.12	0.02
	(0.325)	(0.134)
some w/ any prek	0.18	0.05
	(0.381)	(0.210)
all w/ any prek	0.70	0.94
	(0.457)	(0.246)
no. of any prek in pre-policy	28.37	67.59
	(70.945)	(105.159)
no. of CEI in pre-policy	20.52	50.45
	(56.775)	(88.030)
no. of PRONOEI in pre-policy	7.85	17.14
	(20.982)	(29.089)
no. of prek in post-policy	25.80	52.94
	(48.812)	(61.479)
no. of CEI in post-policy	16.09	33.25
	(26.758)	(32.428)
no. of PRONOEI in post-policy	9.71	19.69
	(24.179)	(32.470)

Table B.12: Descriptive Statistics: students by type of preK

	All	CEI	PRONOEI
student is female	0.49	0.49	0.49
	(0.50)	(0.50)	(0.50)
student age	8.80	8.74	8.72
	(2.86)	(2.85)	(2.84)
family size	5.22	5.18	5.16
	(1.83)	(1.83)	(1.84)
number of kids	3.06	3.03	2.99
	(1.63)	(1.62)	(1.61)
no. of siblings	1.70	1.68	1.64
	(0.84)	(0.82)	(0.79)
monther no educ	0.09	0.08	0.06
	(0.29)	(0.27)	(0.24)
monther primary educ	0.35	0.33	0.30
	(0.48)	(0.47)	(0.46)
monther secondary educ	0.56	0.59	0.64
	(0.50)	(0.49)	(0.48)
father no educ	0.03	0.03	0.02
	(0.17)	(0.16)	(0.14)
father primary educ	0.32	0.29	0.25
	(0.46)	(0.45)	(0.43)
father secondary educ	0.65	0.68	0.73
	(0.48)	(0.47)	(0.44)
mother mainly doing housework	0.68	0.67	0.64
	(0.46)	(0.47)	(0.48)
father mainly doing housework	0.01	0.01	0.01
	(0.10)	(0.10)	(0.09)
Poverty SISFOH category	2.09	2.11	2.15
	(0.91)	(0.91)	(0.90)
school at walk distance	0.73	0.73	0.71
	(0.44)	(0.44)	(0.45)

Note: .

Table B.13: Descriptive Statistics: all vs family fixed-effect samples

	All	Family FE
student is female	0.49	0.49
	(0.50)	(0.50)
student age	9.29	9.28
	(2.95)	(2.94)
family size	5.67	5.67
	(1.88)	(1.88)
number of kids	3.64	3.64
	(1.77)	(1.77)
no. of siblings	2.06	2.06
	(1.01)	(1.01)
monther no educ	0.26	0.26
	(0.44)	(0.44)
monther primary educ	0.61	0.61
	(0.49)	(0.49)
monther secondary educ	0.13	0.13
	(0.34)	(0.34)
father no educ	0.09	0.09
	(0.28)	(0.28)
father primary educ	0.68	0.68
	(0.47)	(0.47)
father secondary educ	0.23	0.23
	(0.42)	(0.42)
mother mainly doing housework	0.88	0.88
	(0.32)	(0.32)
father mainly doing housework	0.02	0.02
	(0.13)	(0.13)
Poverty SISFOH category	1.86	1.86
	(0.92)	(0.92)
school at walk distance	0.86	0.87
	(0.35)	(0.34)

Note: .

Table B.14: Balance Table - all preK

	(1) Never	(2) Changers	(3) Always	(1)-(2)	(1)-(3)	(2)-(3)
Female	0.4958	0.4919	0.4924	-0.0038	-0.0034	0.0004
	(0.5000)	(0.4999)	(0.4999)	(0.0015)	(0.0013)	(0.0009)
	[164,599]	[308,569]	[4373682]			
Family size	5.5791	5.6418	5.1570	0.0628	-0.4220	-0.4848
	(1.8906)	(1.8728)	(1.8179)	(0.0079)	(0.0065)	(0.0045)
	[81,655]	[180,749]	[1668733]			
Mother no education	0.2258	0.2607	0.0727	0.0349	-0.1530	-0.1879
	(0.4181)	(0.4390)	(0.2597)	(0.0018)	(0.0010)	(0.0007)
	[80,757]	[177,287]	[2210621]			
Mother primary education	0.5593	0.6008	0.3235	0.0415	-0.2358	-0.2773
	(0.4965)	(0.4897)	(0.4678)	(0.0021)	(0.0017)	(0.0012)
	[80,757]	[177,287]	[2210621]			
Mother secondary educ.	0.2150	0.1385	0.6038	-0.0764	0.3888	0.4652
	(0.4108)	(0.3455)	(0.4891)	(0.0016)	(0.0017)	(0.0012)
	[80,757]	[177,287]	[2210621]			
Father no education	0.0779	0.0873	0.0229	0.0094	-0.0550	-0.0645
	(0.2680)	(0.2823)	(0.1495)	(0.0012)	(0.0006)	(0.0004)
	[74,153]	[163,626]	[1928887]			
Father primary education	0.6018	0.6718	0.2745	0.0701	-0.3272	-0.3973
	(0.4895)	(0.4695)	(0.4463)	(0.0021)	(0.0017)	(0.0012)
	[74,153]	[163,626]	[1928887]			
Father secondary educ.	0.3204	0.2408	0.7026	-0.0795	0.3822	0.4617
	(0.4666)	(0.4276)	(0.4571)	(0.0019)	(0.0017)	(0.0012)
	[74,153]	[163,626]	[1928887]			
Mother housework	0.8282	0.8828	0.6631	0.0546	-0.1650	-0.2196
	(0.3772)	(0.3217)	(0.4726)	(0.0014)	(0.0017)	(0.0011)
	[80,748]	[177,356]	[2209777]			
Poverty category	1.8991	1.8606	2.1203	-0.0385	0.2213	0.2597
	(0.9262)	(0.9154)	(0.9067)	(0.0039)	(0.0033)	(0.0023)
	[79,256]	[175,883]	[2148880]			
School walking dist.	0.7639	0.8719	0.7255	0.1080	-0.0384	-0.1464
	(0.4247)	(0.3342)	(0.4463)	(0.0051)	(0.0055)	(0.0034)
	[6,786]	[18,047]	[308,376]			

Note: .

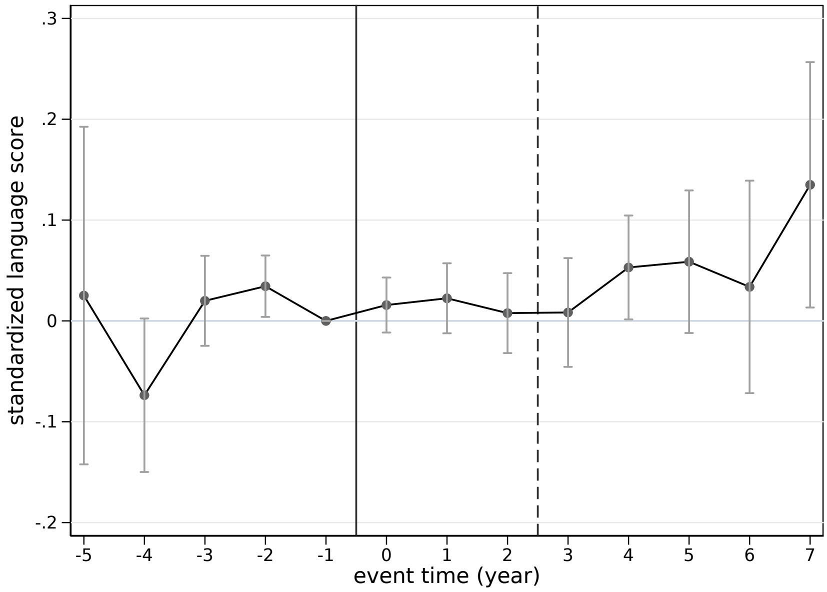
Table B.15: Balance Table - cei

	(1) Never	(2) Changers	(3) Always	(1)-(2)	(1)-(3)	(2)-(3)
student is female	0.4958	0.4919	0.4924	0.0010	0.0010	0.0000
	(0.5000)	(0.4999)	(0.4999)	(0.0012)	(0.0009)	(0.0009)
	[164,599]	[308,569]	[4373682]			
student age	9.6861	9.1080	8.7425	0.5466	0.8302	0.2836
	(2.9331)	(2.9358)	(2.8467)	(0.0095)	(0.0069)	(0.0067)
	[82,770]	[181,239]	[2272423]			
family size	5.5791	5.6418	5.1570	-0.0363	0.4468	0.4831
	(1.8906)	(1.8728)	(1.8179)	(0.0061)	(0.0045)	(0.0044)
	[81,655]	[180,749]	[1668733]			
number of kids	3.5249	3.6220	3.0020	-0.0542	0.5623	0.6165
	(1.7589)	(1.7583)	(1.6046)	(0.0057)	(0.0039)	(0.0038)
	[82,842]	[181,396]	[2273753]			
no. of siblings	1.8770	2.1147	1.6628	-0.1901	0.2831	0.4731
	(0.9282)	(1.0271)	(0.8060)	(0.0032)	(0.0020)	(0.0019)
	[82,841]	[181,396]	[2273752]			
monther no educ	0.2258	0.2607	0.0727	-0.0300	0.1598	0.1898
	(0.4181)	(0.4390)	(0.2597)	(0.0014)	(0.0007)	(0.0006)
	[80,757]	[177,287]	[2210621]			
monther primary educ	0.5593	0.6008	0.3235	-0.0053	0.2761	0.2814
	(0.4965)	(0.4897)	(0.4678)	(0.0016)	(0.0011)	(0.0011)
	[80,757]	[177,287]	[2210621]			
monther secondary educ	0.2150	0.1385	0.6038	0.0353	-0.4359	-0.4712
	(0.4108)	(0.3455)	(0.4891)	(0.0012)	(0.0012)	(0.0011)
	[80,757]	[177,287]	[2210621]			
father no educ	0.0779	0.0873	0.0229	-0.0070	0.0558	0.0629
	(0.2680)	(0.2823)	(0.1495)	(0.0009)	(0.0004)	(0.0004)
	[74,153]	[163,626]	[1928887]			
father primary educ	0.6018	0.6718	0.2745	-0.0297	0.3607	0.3903
	(0.4895)	(0.4695)	(0.4463)	(0.0017)	(0.0011)	(0.0011)
	[74,153]	[163,626]	[1928887]			
father secondary educ	0.3204	0.2408	0.7026	0.0367	-0.4165	-0.4532
	(0.4666)	(0.4276)	(0.4571)	(0.0016)	(0.0012)	(0.0011)
	[74,153]	[163,626]	[1928887]			
mother housework	0.8282	0.8828	0.6631	-0.0286	0.1895	0.2181
	(0.3772)	(0.3217)	(0.4726)	(0.0011)	(0.0012)	(0.0011)
	[80,748]	[177,356]	[2209777]			
father housework	0.0178	0.0140	0.0092	0.0065	0.0100	0.0035
	(0.1322)	(0.1175)	(0.0954)	(0.0004)	(0.0003)	(0.0002)
	[74,143]	[163,546]	[1928221]			
Poverty SISFOH category	1.8991	1.8606	2.1203	0.0336	-0.2203	-0.2539
	(0.9262)	(0.9154)	(0.9067)	(0.0030)	(0.0022)	(0.0022)
	[79,256]	[175,883]	[2148880]			
school at walk distance	0.7639	0.8719	0.7255	-0.0482	0.0976	0.1458
	(0.4247)	(0.3342)	(0.4463)	(0.0039)	(0.0038)	(0.0032)
	[6,786]	[18,047]	[308,376]			

Note: .

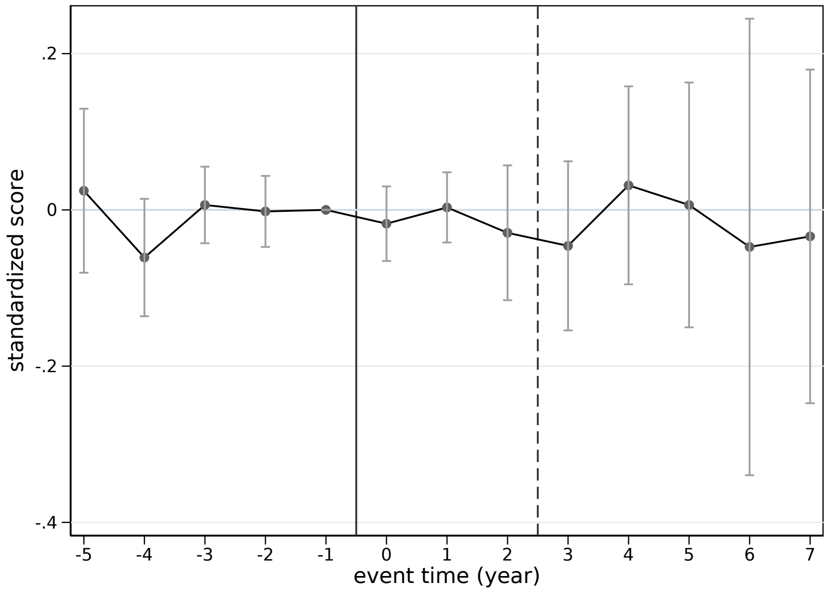
B.2 Event Study Graphs

Figure B.11



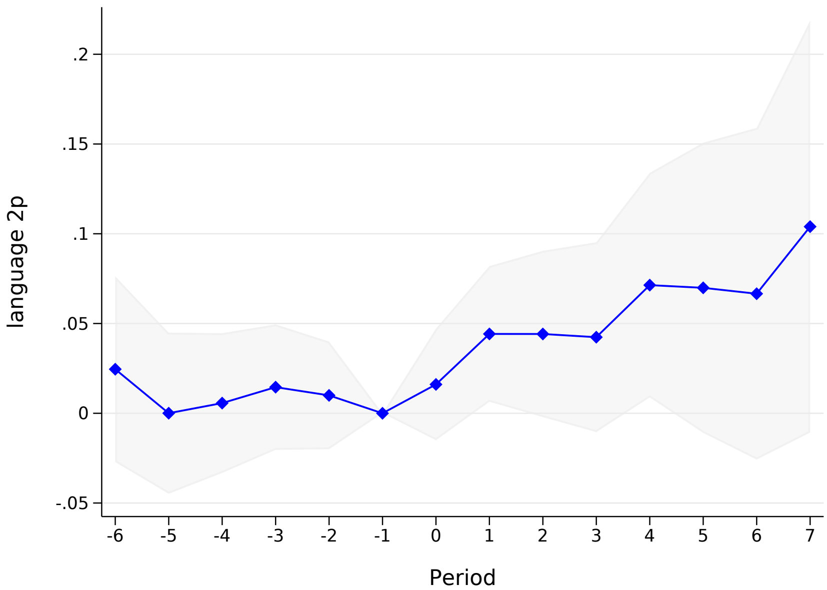
Note:

Figure B.12



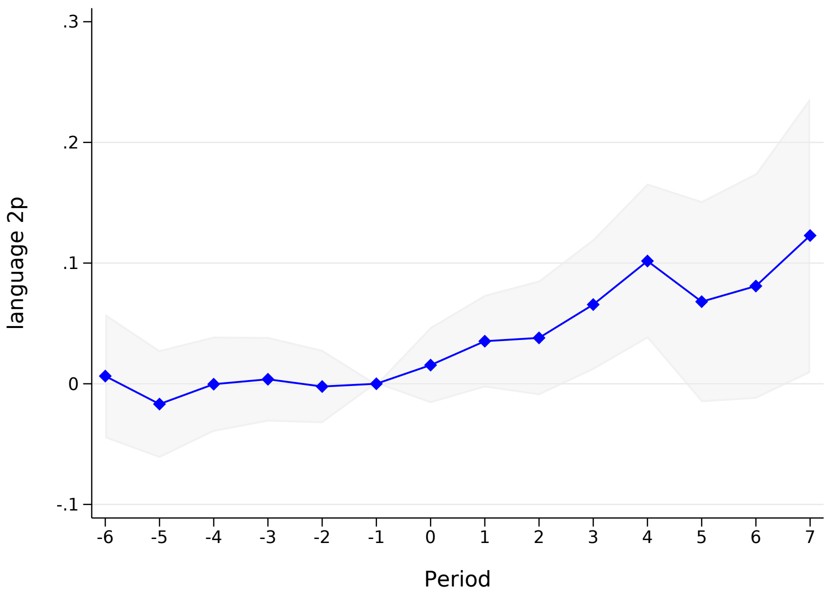
Note:

Figure B.13



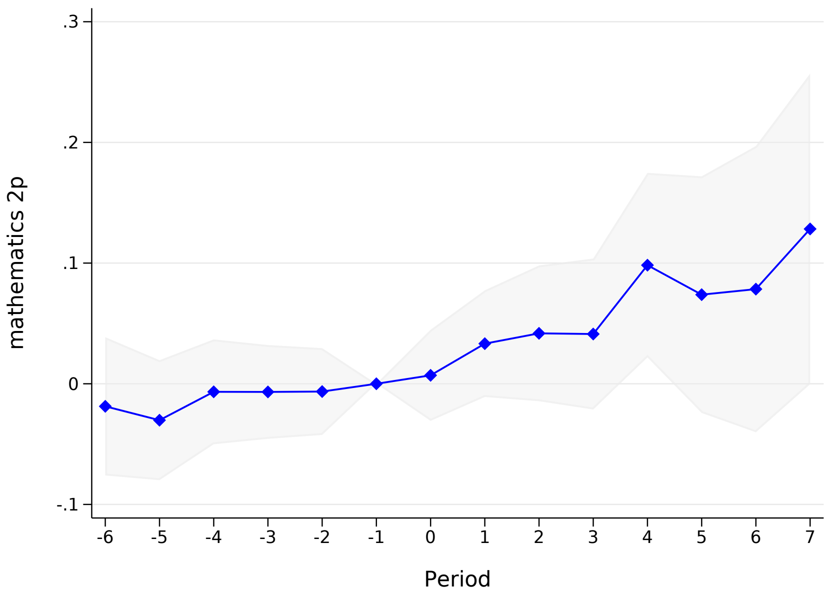
Note:

Figure B.14



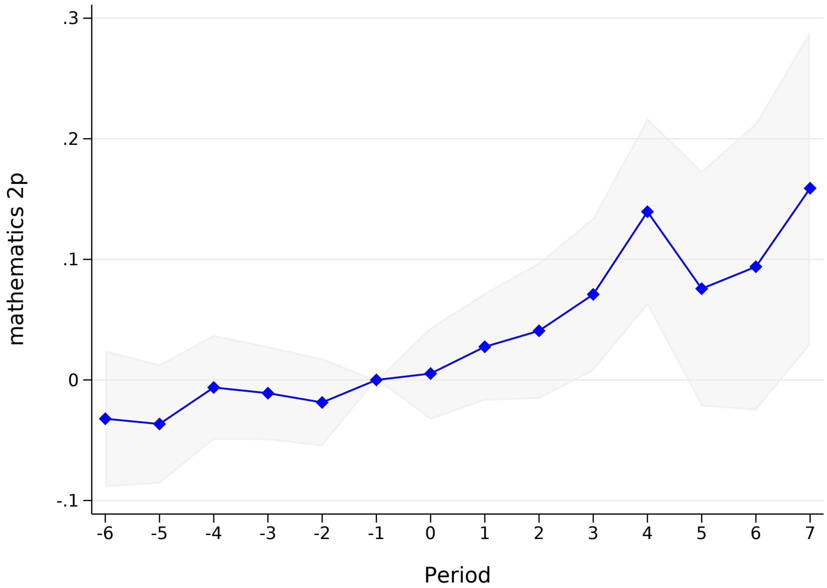
Note:

Figure B.15



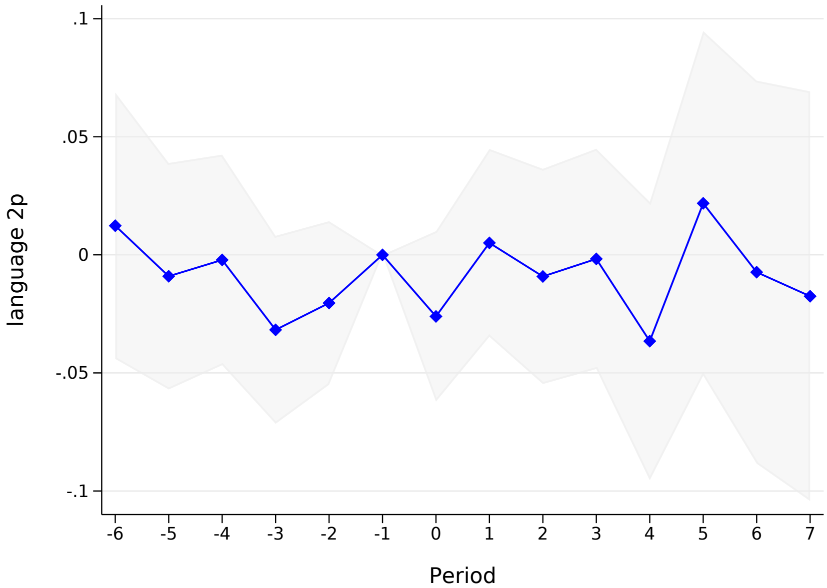
Note:

Figure B.16



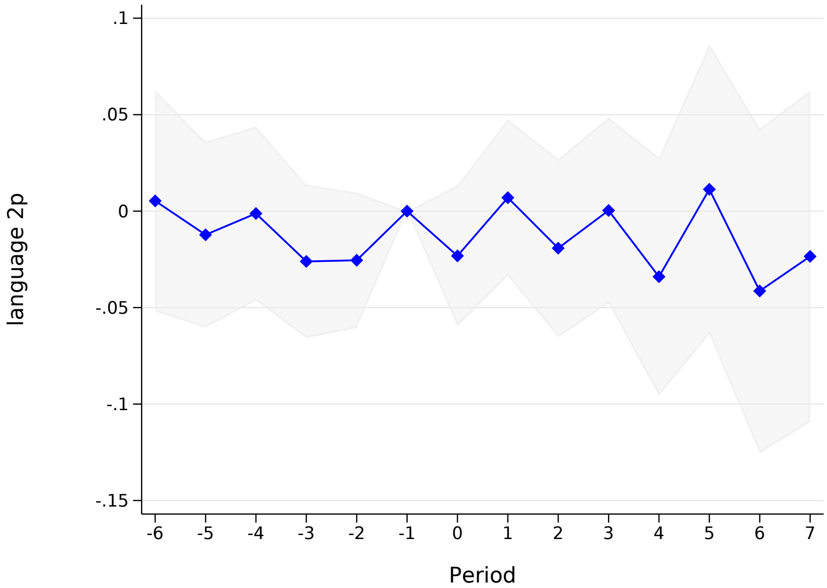
Note:

Figure B.17



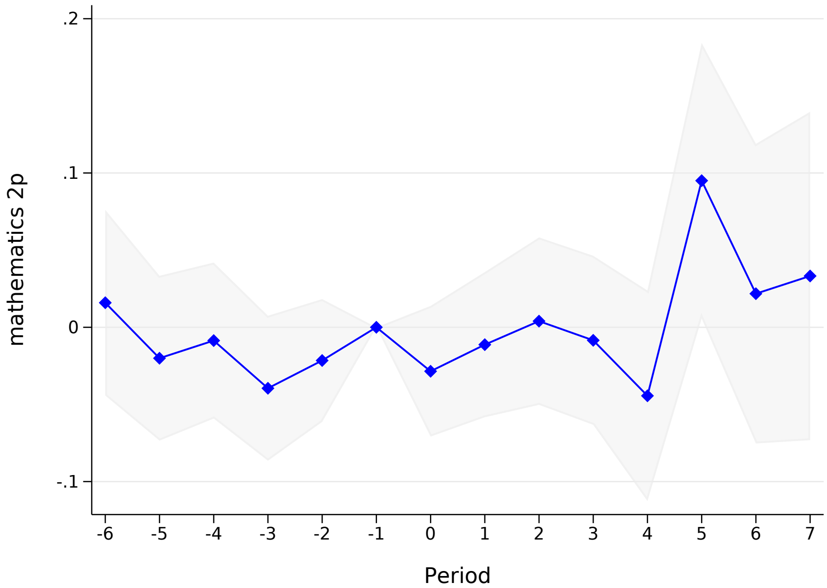
Note:

Figure B.18

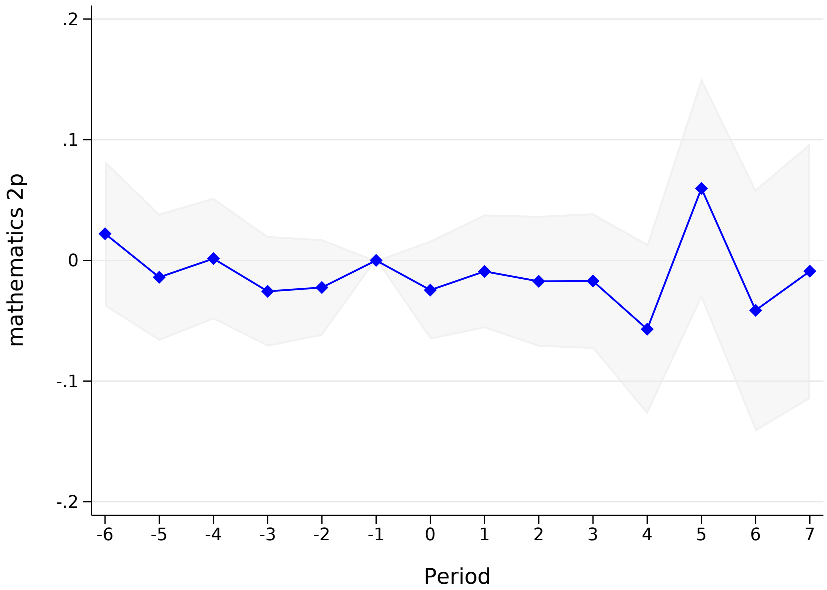


Note:

Figure B.19

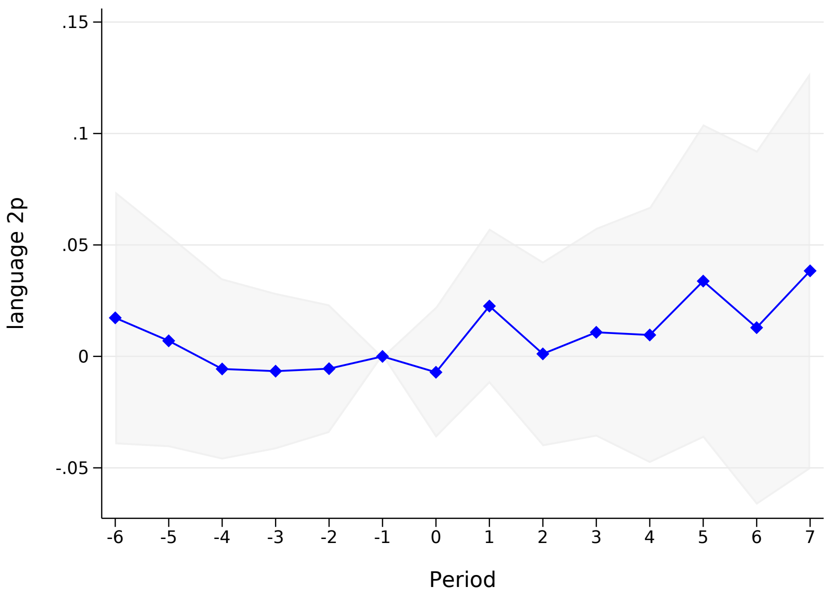


Note:



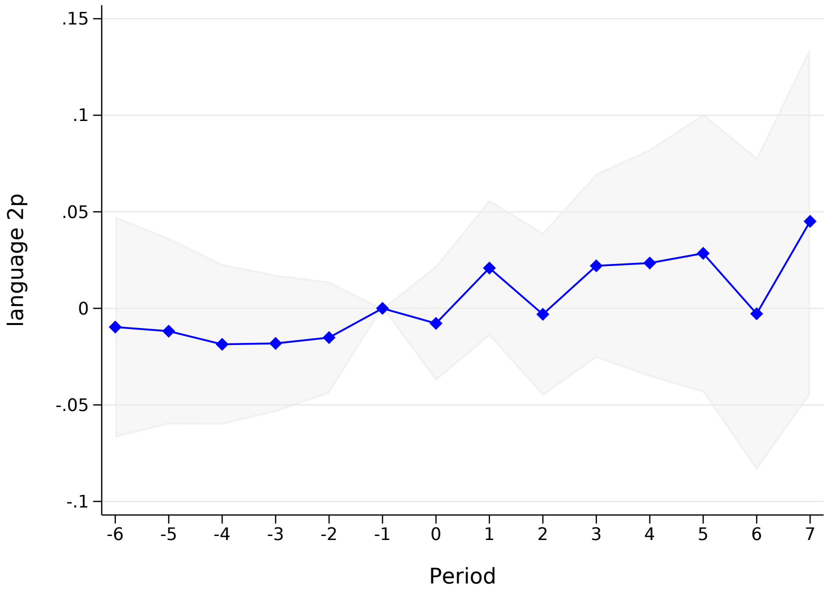
Note:

Figure B.20



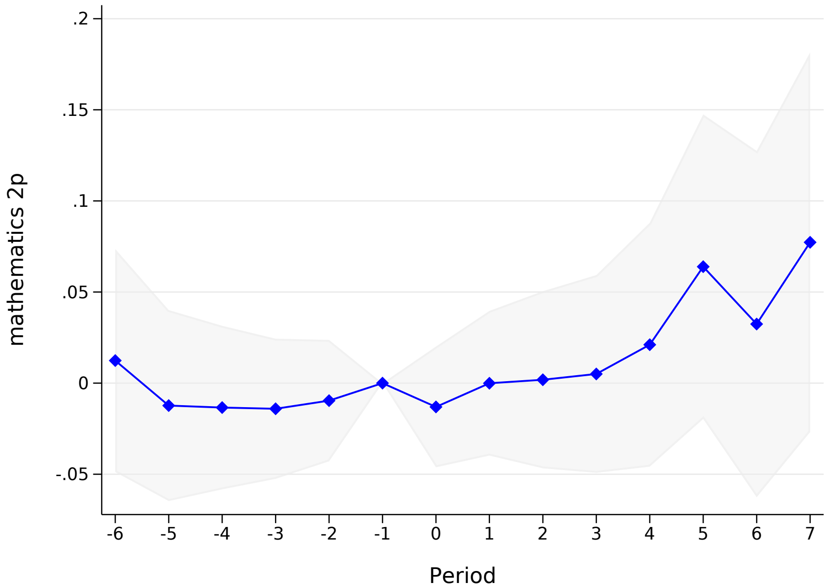
Note:

Figure B.21



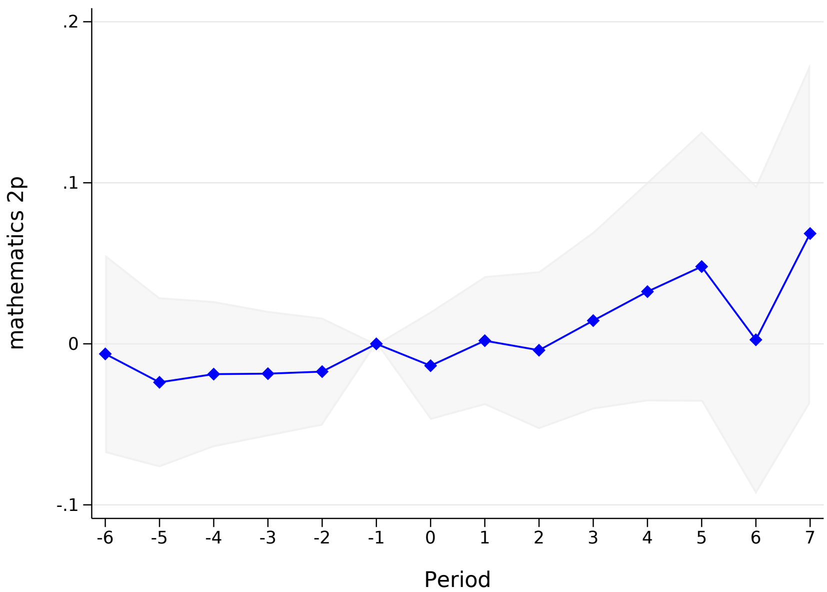
Note:

Figure B.22



Note:

Figure B.23



Note:

B.3 Two-Way Fixed Effects estimation

Table B.16: Naive regressions

	Diff in Means		School Fixed Effects		Socioec.	
	Math	Reading	Math	Reading	Math	Reading
	(1)	(2)	(3)	(4)	(5)	(6)
Preschool	60.15***	54.20***	26.53***	19.62***	22.82***	16.50***
	(0.449)	(0.305)	(0.407)	(0.280)	(1.282)	(0.736)
CEI	62.27***	57.12***	28.34***	21.24***	24.78***	18.07***
	(0.458)	(0.307)	(0.426)	(0.294)	(1.327)	(0.777)
PRONOEI	38.95***	24.84***	18.51***	11.31***	17.50***	11.05***
	(0.762)	(0.528)	(0.726)	(0.499)	(1.785)	(1.022)
N	490,934	491,068	490,934	491,068	337,897	337,897

Note: These are naive regressions that represent the difference in test scores between students with and without preschool. Columns 1 and 2 show a simple difference in means between students with and without preschool, by type of preschool. Columns 3 and 4 add school fixed effects to control for town or school specific differences. Columns 5 and 6 in addition include controls for socioeconomic characteristics. Robust standard errors in parentheses; Each coefficient represents a separate regression.

Table B.17: OLS TWFE - ECE 2P, 4P, 2S, prek by type with controls, catch area 0m

	(1) lang 2p	(2) math 2p	(3) lang 2p	(4) math 2p	(5) lang 2p	(6) math 2p
cei Prek	0.0499***	0.0448***				
	[0.0150]	[0.0171]				
pronoei Prek			-0.0149	-0.0180		
			[0.0148]	[0.0172]		
any Prek					0.0232*	0.0152
					[0.0127]	[0.0146]
Observations	416454	417108	416454	417108	416454	417108
R-squared	0.368	0.295	0.368	0.295	0.368	0.295
Controls	yes	yes	yes	yes	yes	yes

Standard errors in brackets
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

	Regular		Community		Any	
	Reading	Math	Reading	Math	Reading	Math
	(1)	(2)	(3)	(4)	(5)	(6)
Regular Preschool	0.0470***	0.0439**				
	[0.0150]	[0.0172]				
Community Preschool			-0.0185	-0.0231		
			[0.0149]	[0.0173]		
Any Preschool					0.0201	0.0131
					[0.0127]	[0.0146]
Observations	416454	417108	416454	417108	416454	417108
R-squared	0.368	0.295	0.368	0.295	0.368	0.295
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Table B.18: OLS TWFE - ECE 2P, all prek, catch area 0m

	(1) lang 2p	(2) lang 2p	(3) math 2p	(4) math 2p
any Prek	0.0267**	0.0232*	0.0147	0.0152
	[0.0127]	[0.0127]	[0.0145]	[0.0146]
Observations	416454	416454	417108	417108
R-squared	0.362	0.368	0.288	0.295
Controls	no	yes	no	yes

Standard errors in brackets

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table B.19: OLS TWFE - ECE 2P, all prek, catch area 1000m

	(1) lang 2p	(2) lang 2p	(3) math 2p	(4) math 2p
any Prek	0.0234*	0.0201	0.0126	0.0131
	[0.0127]	[0.0127]	[0.0146]	[0.0146]
Observations	416454	416454	417108	417108
R-squared	0.362	0.368	0.288	0.295
Controls	no	yes	no	yes

Standard errors in brackets

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table B.20: OLS TWFE - ECE 2P, cei with controls

	(1) lang 2p	(2) lang 2p	(3) math 2p	(4) math 2p
cei Prek	0.0534***	0.0499***	0.0412**	0.0448***
	[0.0148]	[0.0150]	[0.0170]	[0.0171]
Observations	416454	416454	417108	417108
R-squared	0.362	0.368	0.288	0.295
Controls	no	yes	no	yes

Standard errors in brackets

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table B.21: OLS TWFE - ECE 2P, cei with controls

	(1) lang 2p	(2) lang 2p	(3) math 2p	(4) math 2p
cei Prek	0.0504***	0.0470***	0.0400**	0.0439**
	[0.0149]	[0.0150]	[0.0171]	[0.0172]
Observations	416454	416454	417108	417108
R-squared	0.362	0.368	0.288	0.295
Controls	no	yes	no	yes

Standard errors in brackets

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table B.22: OLS TWFE - ECE 2P, pronoei prek, catch area 0m

	(1) lang 2p	(2) lang 2p	(3) math 2p	(4) math 2p
pronoei Prek	-0.0121	-0.0149	-0.00665	-0.0180
	[0.0146]	[0.0148]	[0.0172]	[0.0172]
Observations	416454	416454	417108	417108
R-squared	0.362	0.368	0.288	0.295
Controls	no	yes	no	yes

Standard errors in brackets

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table B.23: OLS TWFE - ECE 2P, pronoei prek, catch area 1000m

	(1) lang 2p	(2) lang 2p	(3) math 2p	(4) math 2p
pronoei Prek	-0.0162	-0.0185	-0.0121	-0.0231
	[0.0147]	[0.0149]	[0.0173]	[0.0173]
Observations	416454	416454	417108	417108
R-squared	0.362	0.368	0.288	0.295
Controls	no	yes	no	yes

Standard errors in brackets

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table B.24: OLS TWFE - ECE 4P, all prek, catch area 0m

	(1) lang 4p	(2) lang 4p	(3) math 4p	(4) math 4p
any Prek	-0.0724***	-0.0675***	-0.0595**	-0.0570**
	[0.0220]	[0.0228]	[0.0257]	[0.0268]
Observations	55713	55692	55699	55678
R-squared	0.404	0.407	0.428	0.431
Controls	no	yes	no	yes

Standard errors in brackets

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table B.25: OLS TWFE - ECE 4P, all prek, catch area 1000m

	(1) lang 4p	(2) lang 4p	(3) math 4p	(4) math 4p
any Prek	-0.0750***	-0.0701***	-0.0609**	-0.0584**
	[0.0220]	[0.0228]	[0.0258]	[0.0268]
Observations	55713	55692	55699	55678
R-squared	0.404	0.407	0.428	0.431
Controls	no	yes	no	yes

Standard errors in brackets

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table B.26: OLS TWFE - ECE 4P, cei prek, catch area 0m

	(1) lang 4p	(2) lang 4p	(3) math 4p	(4) math 4p
cei Prek	-0.0533**	-0.0428*	-0.0291	-0.0227
	[0.0224]	[0.0234]	[0.0261]	[0.0271]
Observations	55713	55692	55699	55678
R-squared	0.404	0.407	0.428	0.431
Controls	no	yes	no	yes

Standard errors in brackets

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table B.27: OLS TWFE - ECE 4P, cei prek, catch area 1000m

	(1) lang 4p	(2) lang 4p	(3) math 4p	(4) math 4p
cei Prek	-0.0535**	-0.0430*	-0.0291	-0.0225
	[0.0225]	[0.0235]	[0.0262]	[0.0271]
Observations	55713	55692	55699	55678
R-squared	0.404	0.407	0.428	0.431
Controls	no	yes	no	yes

Standard errors in brackets

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table B.28: OLS TWFE - ECE 4P, pronoei prek, catch area 0m

	(1) lang 4p	(2) lang 4p	(3) math 4p	(4) math 4p
pronoei Prek	-0.0313	-0.0417	-0.0473	-0.0465
	[0.0316]	[0.0275]	[0.0317]	[0.0312]
Observations	55713	55692	55699	55678
R-squared	0.404	0.407	0.428	0.431
Controls	no	yes	no	yes

Standard errors in brackets

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table B.29: OLS TWFE - ECE 4P, pronoei prek, catch area 1000m

	(1) lang 4p	(2) lang 4p	(3) math 4p	(4) math 4p
pronoei Prek	-0.0369	-0.0480*	-0.0540*	-0.0544*
	[0.0326]	[0.0283]	[0.0325]	[0.0319]
Observations	55713	55692	55699	55678
R-squared	0.404	0.407	0.428	0.431
Controls	no	yes	no	yes

Standard errors in brackets

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table B.30: OLS TWFE - ECE 2S, all prek, catch area 0m

	(1)	(2)	(3)	(4)
	lang 2s	lang 2s	math 2s	math 2s
any Prek	0.00303	-0.00473	0.00551	0.00234
	[0.0116]	[0.0116]	[0.0132]	[0.0135]
Observations	104960	104943	104942	104925
R-squared	0.317	0.319	0.294	0.297
Controls	no	yes	no	yes

Standard errors in brackets

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table B.31: OLS TWFE - ECE 2S, all prek, catch area 1000m

	(1)	(2)	(3)	(4)
	lang 2s	lang 2s	math 2s	math 2s
any Prek	0.00141	-0.00615	0.00452	0.00126
	[0.0117]	[0.0116]	[0.0132]	[0.0135]
Observations	104960	104943	104942	104925
R-squared	0.317	0.319	0.294	0.297
Controls	no	yes	no	yes

Standard errors in brackets

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table B.32: OLS TWFE - ECE 2S, cei prek, catch area 0m

	(1)	(2)	(3)	(4)
	lang 2s	lang 2s	math 2s	math 2s
cei Prek	0.0224	0.00607	0.0275	0.0145
	[0.0152]	[0.0159]	[0.0173]	[0.0188]
Observations	104960	104943	104942	104925
R-squared	0.317	0.319	0.294	0.297
Controls	no	yes	no	yes

Standard errors in brackets

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table B.33: OLS TWFE - ECE 2S, cei prek, catch area 1000m

	(1) lang 2s	(2) lang 2s	(3) math 2s	(4) math 2s
cei Prek	0.0205	0.00420	0.0262	0.0128
	[0.0154]	[0.0160]	[0.0174]	[0.0189]
Observations	104960	104943	104942	104925
R-squared	0.317	0.319	0.294	0.297
Controls	no	yes	no	yes

Standard errors in brackets

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table B.34: OLS TWFE - ECE 2S, pronoei prek, catch area 0m

	(1) lang 2s	(2) lang 2s	(3) math 2s	(4) math 2s
pronoei Prek	-0.0211	-0.0150	-0.0171	-0.00760
	[0.0145]	[0.0146]	[0.0166]	[0.0170]
Observations	104960	104943	104942	104925
R-squared	0.317	0.319	0.294	0.297
Controls	no	yes	no	yes

Standard errors in brackets

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table B.35: OLS TWFE - ECE 2S, pronoei prek, catch area 1000m

	(1) lang 2s	(2) lang 2s	(3) math 2s	(4) math 2s
pronoei Prek	-0.0183	-0.0114	-0.0135	-0.00376
	[0.0143]	[0.0144]	[0.0165]	[0.0169]
Observations	104960	104943	104942	104925
R-squared	0.317	0.319	0.294	0.297
Controls	no	yes	no	yes

Standard errors in brackets

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$