HCEO WORKING PAPER SERIES

Early Skill Effects on Types of Parental Investments and Long-Run Outcomes

Sebastian Gallegos
Pablo Celhay

Working Paper

2020-014

03/2020

The University of Chicago
1126 E. 59th Street Box 107
Chicago IL 60637

Early Skill Effects on Types of Parental Investments and Long-Run Outcomes[1]

Sebastian Gallegos[a] and Pablo A. Celhay[b]

[a]Inter-American Development Bank
[b]Pontificia Universidad Católica de Chile

March 13, 2020

Abstract

This paper examines the effects of skill advantages at age six on different types of parental investments, and long-run outcomes up to age 27. We exploit exogenous variation in skills due to school entry rules, combining 20 years of Chilean administrative records with a regression discontinuity design. Our results show higher in-school performance and college entrance scores, and sizable effects on college attendance and enrollment at selective institutions, particularly for low-income children. We find that parental time investments are neutral to early skills gaps, while monetary investments are reinforcing and likely to be mediating the long-run effects.

Keywords: Early Life Shocks; Long-run Outcomes; Skills; Parental Investments; College Attendance, Test Scores, Low-income, Developing Country

JEL codes: I21, I26, I28, J24, J31.

[1]Gallegos (corresponding author): Inter-American Development Bank, 1300 New York Avenue, NW, Washington, DC 20577, and member of the HCEO Network, University of Chicago. Email: sebagallegos@uchicago.edu. Celhay: School of Government, Pontificia Universidad Católica de Chile; Vicuna Mackenna 4860, Macul, Region Metropolitana, Chile. Email: pacelhay@uc.cl. We thank the Ministry of Education (MINEDUC) of the Government of Chile and the DEMRE at Universidad de Chile for access to the sources of data. We thank Philip Oreopoulos, Christopher Neilson, Arnaud Maurel, Ofer Malamud, Susan Mayer, Pietro Biroli, María Rosales, Oscar Mitnik, Diego Vera and seminar participants at the University of Chicago, Princeton University, the Inter-American Development Bank, and the APPAM and SREE annual meetings for helpful comments. This paper is dedicated to the memory of our professor Robert J. Lalonde. Celhay acknowledges financial support from CONICYT (FONDECYT/Iniciación/11180416). Gallegos also thanks the Center for Studies of Conflict and Social Cohesion (CONICYT/FONDAP/15130009). All errors are our own; the views expressed in this paper are those of the authors and should not be attributed to the Inter-American Development Bank or to any other institution.

1 Introduction

In developing and emerging countries, human capital is the most important asset that children acquire to escape poverty in the long run. Yet, several studies show that children's development in these places is particularly vulnerable to early life shocks (see, e.g., Bharadwaj et al., 2013), with lasting effects on adult outcomes (Almond et al., 2018). Investments by parents can mediate these long term consequences and there is growing interest in understanding how parents' behavior respond to shocks on their children's human capital.[1] Parental investments have been linked to beliefs about their returns (e.g., Carneiro et al., 2019; Boneva and Rauh, 2018; Cunha et al., 2013), information frictions (Dizon-Ross, 2019), time and budget constraints (e.g., Bono et al., 2016, Dahl and Lochner, 2012), or preferences (e.g., Beuermann and Jackson, 2018; Bharadwaj et al., 2018). Despite the growing literature in this area, several questions are in need of more evidence. These questions are related to whether parents' choices of investments vary by type (e.g., cash or time inputs) and how responses exacerbate (if at all) long term effects on their children. These questions become more relevant in emerging economies, where children have less resources to cope with early life disparities.

This paper examines the effects of skill advantages at age six on multiple types of parental investments and educational outcomes up to age 27. We first document effects on in-school measures of performance. Second, we show how parental beliefs on children's future human capital, financial investments, and time investments respond to these gaps in school performance. Third, we estimate effects on long run outcomes, following children over twenty years. We place our findings within a human capital accumulation framework that connects parental responses to the effects of early skill gaps on children's future human capital.

Our research design mimics a local experiment where children are exogenously allocated to start school at different ages due to birth date cutoff rules. Figure 1 shows that these age differences translate into large disparities in multiple skills, measured just before school entry.[2] We exploit this variation in skills at entry combining twenty years of Chilean administrative micro-data with a regression discontinuity design. We supplement the administrative records with survey data containing information on parental investments and beliefs reported by parents and students. Leveraging large sample sizes, we are also able to estimate precisely how our results vary by socioeconomic background.

We find that children who start school with higher ability perform better on several in-school outcomes like GPA and test scores (0.20 standard deviations), measured from the 1st through the 4th grade. In addition, we show that by the end of 4th grade parents are more likely to believe those children will complete post-secondary degrees, such as graduate school, college and technical careers

[1] Parental investments in children have been a topic of study for a long time (e.g., Becker and Tomes, 1976 and Behrman et al., 1982), but interest has risen sharply in more recent years. See, e.g., Almond and Mazumder (2013), Doepke and Zilibotti (2017), and Francesconi and Heckman (2016).

[2] Among other factors, previous research has shown that older children have been exposed to more parenting time and are more mature than their younger peers and so can perform higher in cognitive test scores and can better develop different skills (e.g., Black et al., 2011; Deming, 2009; Dhuey et al., 2019; Lubotsky and Kaestner, 2016).

(effect sizes of 13%, 4% and 2%). We then document that parents do not invest time differentially, but reinforce initial skill gaps by investing an additional 0.11σ of financial resources in children with higher ability at entry.

Following the same students over time, we find that early skill gaps lead to higher take-up (6 percent) and scores (0.08 standard deviations) in the national college entrance exam, and a higher probability of college enrollment, both overall (15 percent) and at selective programs (20 percent). In addition, all of our effects on financial investments and college enrollment are more pronounced for low-income children. The magnitude of the estimates on college enrollment are within those found by the early childhood interventions literature (Elango et al., 2015), suggesting that policy shocks on early skills in developing countries can be as important as programs that are especially designed to bolster children's abilities.

Our paper contributes to the literature of parental beliefs and investments on children. Our findings on the neutrality of parental time investments are consistent with the results by Bharadwaj et al. (2013) for Norway and Chile, in the context of an early health intervention. We add to these results by showing that at the same time monetary investments respond differently to early skills gaps. These results are consistent with parents of high-performing children perceiving that the returns to monetary investments are higher than time investments. We also add to recent research that connects parental investments to parents' beliefs (e.g., Biroli et al., 2018; Boneva and Rauh, 2018; Dizon-Ross, 2019; Attanasio et al., 2019a). Our findings on beliefs suggest that parents interpret the results on in-school performance as signals of their child ability, and adjust their investments according to their perceived return.

Our paper also contributes to the literature that studies long term outcomes of early life disparities. We track students for twenty years with repeated measures across time. Few studies are able to observe outcomes in the middle years of life (the 'missing middle' in Almond et al., 2018), which are important to fully understand effects. For instance, Heckman et al. (2006), Elango et al. (2015) and Beuermann and Jackson (2018) highlight that the effects of early life disparities might fade out in the medium term but emerge in the long run. We measure relevant outcomes at age 6 (GPA), 10 (test scores), 14 (primary school completion), 18 to 20 (high school completion and college entrance exams) and up to 27 (college completion).

Overall, our results suggest that parents respond to signals of children's ability in different ways depending on the type of investment. Those responses may reinforce early gaps with consequences on long-term educational outcomes. We argue that signals would disappear if adjusted by age, but parents and children themselves observe and react based on the unadjusted in-school differences. The natural policy implication of these findings is to provide parents both raw and age-adjusted measures of performance. Schools could also train teachers to communicate results to parents so that they take age into account when assessing children ability, particularly during the early school years when age-performance gaps are larger.

The remaining sections of the paper are organized as follows. In Section 2 we introduce a simple model of human capital accumulation and outline our empirical strategy. In Section 3 we describe the data. Section 4 presents the results and connects our main findings to the conceptual framework.

Section 5 concludes.

2 Methods

In this section we outline a simple model of human capital accumulation that serves to understand how different types of parental investment might respond to early shocks. Then we describe the empirical strategy we use to estimate the causal effect of early disparities on children's outcomes and parental investments.

2.1 Conceptual Framework

We present a conceptual framework that describes the mapping of early childhood shocks and parental investments into a child's future human capital, building on multiple studies from the related literature (e.g., Almond et al., 2018, Boneva and Rauh, 2018, Cunha et al., 2010, and Francesconi and Heckman, 2016). In our model parents have beliefs about their child's ability and can make different types of investments (e.g., spending additional time on educational activities or investing additional money on school-related items). Since the choice of the production function might govern the response to early childhood shocks (Almond et al., 2018), we do not presuppose a particular functional form for preferences or technology relating human capital to later outcomes. This allows different investments to vary in magnitude and sign as a response to early shocks. We consider a simple model with two periods, where the first period is childhood and the second period is child i's young adulthood.

Child i's human capital in the second period is determined by the following production technology

$$h_{2i} = h(\theta_{0i}, I_{1i}^m, I_{1i}^t, \zeta_{1i}), \tag{1}$$

where θ_{0i} represents endowed skills, I_{1i}^m are monetary investments made by parents of child i (such as school-related expenditures) in period 1, I_{1i}^t is child i's parent time investment (such as mentoring activities) in period 1, and ζ_{1i} is a shock during childhood (e.g., a skill advantage in the first grade). We assume that $h(\cdot)$ is differentiable, monotone, weakly increasing, and concave in I_{1i}^m, I_{1i}^t.

Parents have an expectation about the level of human capital that their child will achieve in adulthood, $\widetilde{h_{2i}}$, which depends on their beliefs about child i's skill endowment, $\widetilde{\theta_{0i}}$; parents' investments in their child during the childhood period, I_{1i}^m, I_{1i}^t; and the early shock faced by their child. We introduce these beliefs to point out that parents decide to invest considering their child's expected human capital in adulthood, which may differ from the human capital that they finally acquire (h_{2i}). Importantly, it may be the case that the shock ζ_{1i} does not change the skill endowment of child i, θ_{0i}, but does change parents' beliefs about it. Parents' perceived child's future human capital can be written as

$$\widetilde{h_{2i}} = h(\widetilde{\theta_{0i}}, I_{1i}^m, I_{1i}^t, \zeta_{1i}). \tag{2}$$

During the childhood period, parent i allocates leisure time L_{1i} to child time investment, I_{1i}^{t}, and own leisure time, l_{1i}, so that $L_{1i} = I_{1i}^{t} + l_{1i}$. She also chooses how to allocate available money, M_{1i}, into consumption, C_{1i}, and monetary investment in children, I_{1i}^{m}. Therefore she faces time and budget constraints given by

$$L_{i1} = I_{1i}^{t} + l_{i1} \tag{3}$$

$$M_{1i} = Y_{1i} + w(T - L_{i1}) = C_{1i} + p_I I_{1i}^{m}, \tag{4}$$

where Y_{1i} is non-labor income, w denotes wage in the labor market, T is fixed and represents time available during the day, and p_I is the unit price of monetary investment (e.g., books, computer), with the price of consumption normalized to one. Allowing parents to have preferences on their own leisure time, consumption, and expected child's human capital in adulthood, their maximization problem becomes

$$\max_{I_{1i}^{m}, I_{1i}^{t}} U(l_{1i}, C_{1i}, \widetilde{h_{2i}}) \quad s.t. \ (2), \ (3), \ and \ (4);$$

i.e., the parent chooses different types of investment levels to maximize utility subject to the production technology, budget, and time constraints. The optimal investment strategies in period 1 for the parent of child i and type of investment k are given by

$$I_{1i}^{*k} = I^{k}(\widetilde{\theta_{0i}}, \zeta_{1i}, p_I, Y, w) \quad for \quad k = m, t. \tag{5}$$

Given these optimal investment decisions, the effect of an early shock on human capital in the next period can be decomposed as

$$\underbrace{\frac{\delta h_{2i}^{*}}{\delta \zeta_{1i}}}_{A} = \underbrace{\frac{\delta h(\cdot)}{\delta \zeta_{1i}}}_{B} + \underbrace{\frac{\delta h(\cdot)}{\delta I_{1i}^{*m}} \times \frac{\delta I_{1i}^{*m}}{\delta \zeta_{1i}}}_{C} + \underbrace{\frac{\delta h(\cdot)}{\delta I_{1i}^{*t}} \times \frac{\delta I_{1i}^{*t}}{\delta \zeta_{1i}}}_{D}. \tag{6}$$

The total effect, A, equals to a direct effect of an early shock, B, which can be mitigated or reinforced through behavioral effects of different investment decisions, C and D. Given that we assume that human capital is weakly increasing in investments, $\frac{\delta h(\cdot)}{\delta I_{1i}^{*k}} \geq 0$, the sign of C and D is determined by how parental investments respond, $\frac{\delta I_{1i}^{*k}}{\delta \zeta_{1i}}$.

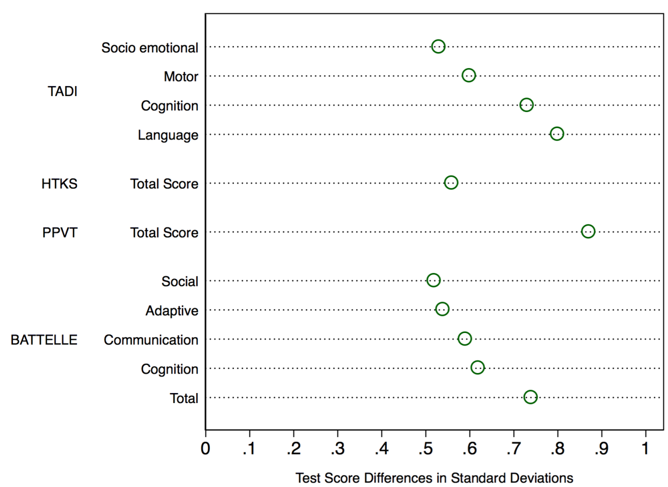
We define a reinforcing investment decision as one that increases investment as a response to a positive shock, $\frac{\delta I_{1i}^{*k}}{\delta \zeta_{1i}} > 0$, while a compensating strategy consists in parents increasing investment as a response to a negative shock, $\frac{\delta I_{1i}^{*k}}{\delta \zeta_{1i}} < 0$.

Parents might respond to shocks differently by type of investment. We hypothesize that the response would differ by the productivity of each investment given the shock and socioeconomic background of the family. For instance, following a negative shock, parents might compensate by investing more time with the child, which is arguably more productive and affordable than buying a computer if the child is lagging behind. These responses imply that $\frac{\delta I_{1i}^{*m}}{\delta \zeta_{1i}} = 0$ and $\frac{\delta I_{1i}^{*t}}{\delta \zeta_{1i}} > 0$. On the other hand, a positive shock $\frac{\delta \widetilde{h}_{2i}^{*}}{\delta \zeta_{1i}} > 0$ may trigger parents' monetary investment, like buying a computer, but not additional mentoring time (because the child is already performing well), so that $\frac{\delta I_{1i}^{*1}}{\delta \zeta_{1i}} > 0$ and $\frac{\delta I_{1i}^{*2}}{\delta \zeta_{1i}} = 0$. Our rich data and research strategy allows us to test these hypotheses in our empirical analysis. We outline our empirical strategy below.

2.2 Empirical Strategy

Our research design resembles a local experiment where children born days apart due to chance start primary school at different ages and thus with very different set of skills at school entry. In Figure 1 we show that there are large differences between older and younger children in a host of tests, measured just before starting school. The age differences translate into skill gaps, measured in standard deviation units (σ) that range from 0.52σ to 0.86σ on a battery of tests commonly used in the early childhood literature (e.g., Rubio-Codina et al., 2016).

Figure 1: Baseline Differences in Skills



Note: Figure 1 plots differences in a host of tests between July- and June-born children, measured just before they start their respective 1^{st} grade. Due to the birth date cutoff rule, July-born children are about a year older than June-born children at school entry. The y-axis shows the measure of different tests and subjects measured. The TADI test is the *Test de Aprendizaje y Desarrollo Infantil*, a test developed by Chilean research centers that specialize in early life development measures. The HTKS test is the Head, Toes, Knees and Shoulders exam; the PPVT test is the Peabody Picture Vocabulary Test, and the BATTELLE test corresponds to the Battelle Developmental Inventory for Young Children. The data comes from the *Encuesta Longitudinal de Primera Infancia* (ELPI), a nationally representative longitudinal survey that follows cohorts of children since birth until early youth.

Our empirical strategy takes advantage of the birth date cutoff rules in Chile, which states that prospective students who are not six years old by June 30 of the academic year should start in the next one. We employ a regression discontinuity (RD) design using exact birth dates for children born in June and July to compare outcomes between children born days apart but with very different skill levels at school entrance.

Our identifying assumptions are standard for RD designs. Essentially, we assume that there are no other changes occurring at the threshold that could confound our analysis. In Appendix A we run a series of robustness tests showing that there are no differences in a host of different covariates at the cutoff and no evidence of manipulation of birth dates around the threshold, and our estimates are stable to using different bandwidths and specifications.

Our main estimating equation is

$$Y_i = \alpha_0 + \alpha_1 Z_i + f(B_i) + \alpha_2 X_i + \mu_i \tag{1}$$

The variable Z_i is equal to one if child i is born in July and is equal to zero if child i is born in June of the same year. $f(B_i)$ is a function of birth date (B_i) interacted with Z_i to allow for different slopes on each side of the cutoff. μ_i represents the error term that we cluster within birth date. We also include a set of predetermined variables as controls in X_i, such as child gender, measures of household socioeconomic status, class size, school rurality, and type of school. All these control variables behave smoothly near the cutoff (see Figure A.2 in Appendix A) and serve mainly to improve precision of our RD estimates. We also add year of birth indicators to control for secular trends common to all children.

Our parameter of interest is α_1, which is the *intention-to-treat* effect of starting school older—with a skill advantage—on the outcome Y_i. We restrict ourselves to these reduced-form effects and do not "scale up" our estimates instrumenting starting age with the threshold because in that case we would need the LATE additional assumptions to hold.[3] In addition, our well identified reduced-form estimates are still conservative as "naive" two-stage least squares (LATE) estimates would mechanically increase the magnitude of our estimated effects. As we describe in the following section, we estimate Equation (1) on many outcomes and therefore simultaneously test multiple hypothesis. To account for the probability of incorrectly rejecting one or more null hypotheses belonging to a family of hypotheses, we follow Anderson (2008) and adjust our standard errors controlling for the family-wise error rate. In the next section, we describe the rich administrative records that we use to implement our empirical strategy.

[3]For instance, we would need to defend that the exclusion restriction holds in this setup (see Jones (2015) for a discussion on this topic), and check the monotonicity assumption (see for example Barua and Lang (2016)).

3 Data

We use administrative data provided by the Ministry of Education (MINEDUC) for the population of students in Chile supplemented with test scores, parental surveys, and student surveys. We link students across their entire school life using an encrypted national identification number and also follow them as they complete high school, take the college entrance exam, enroll in higher education, and graduate from college. We describe our data below.

3.1 Sources

Our primary data source comes from administrative datasets with yearly information on the population of students in primary school (1st to 8th grade) and high school (9th to 12th grade) since year 2002 and up to 2019. Each dataset provides individual data on exact birth date, gender and school characteristics, and in-school outcomes like GPA scores and passing rates.

We supplement these data with standardized test scores from the SIMCE (*Sistema de Medicion de Calidad Escolar*) exam, accompanied by parent and student surveys administered in the fourth grade. We use the surveys to measure parental investments, which we describe in detail in the next subsection. We further combine these data with three additional sources of information to measure long term outcomes. The first comes from the national college entrance exam (*Prueba de Selección Universitaria*, PSU) for years 2004 to 2018. The exam is taken at the end of the high school senior year and is required to get admitted to most universities in the country.[4] The second and third sources are further administrative datasets on higher education enrollment and graduation, respectively. The data is available for years 2007 to 2019. Each year, the MINEDUC collects information from all higher education institutions in the country and produces individual level lists of all students and graduates with information on major, area of study and institution.

3.2 Measures of Parental Investments and Beliefs

We use two surveys to measure parental investment. In one survey, parents provide information on investment in school-related items. For example, if they have a computer or Internet at home, the number of books they own, and the money they spend every month on their child's education. We define binary indicators for these last two variables because they are reported in brackets. We label these variables "higher spending" and "ten or more books", each of which equals one for half of respondents and zero otherwise. Our results are robust to other ways of grouping items of investments, as we show in Appendix B.

In a separate survey, students are asked about the time spent with their parents on educational activities. In particular, children report in a 1-4 Likert scale whether their parents help them study or with their homework, help them understand difficult subjects, whether parents know their grades, and whether parents demand improving grades. Available answers for each item are on the scale of

[4] We signed an agreement with the agency in charge of developing and administering the exam (DEMRE), which provided us the data with the same encrypted identification number contained in the MINEDUC data.

"Never", "Sometimes", "Most of the time", and "Always". We generate variables that equal one if the child answered that her parent does each activity "Always", and zero otherwise. As with financial investments, our analyses are robust to how we group answers.

Figure 2 shows descriptive non-parametric plots of the raw data relating financial and time investment variables to test scores (in standard deviation units). Figure 2a shows that parental financial investments are positively correlated with test scores. Figure 2b also shows a positive correlation between test scores and measures of whether parents always know and congratulate children for their grades. The figure shows a flatter, slightly u-shaped correlation between test scores and measures of whether parents demand good grades, help with homework, and help them study.

We also generated two summary variables for each type of parental investment. The first is a simple average of all investments within type, which is plotted as 'Average Index' in Figure 2a and Figure 2b, respectively. The second is a 'Factor Index', computed using principal components, which reduces the dimensionality of the investment measures to one composite score. We provide details of the composite score computation in Appendix B.

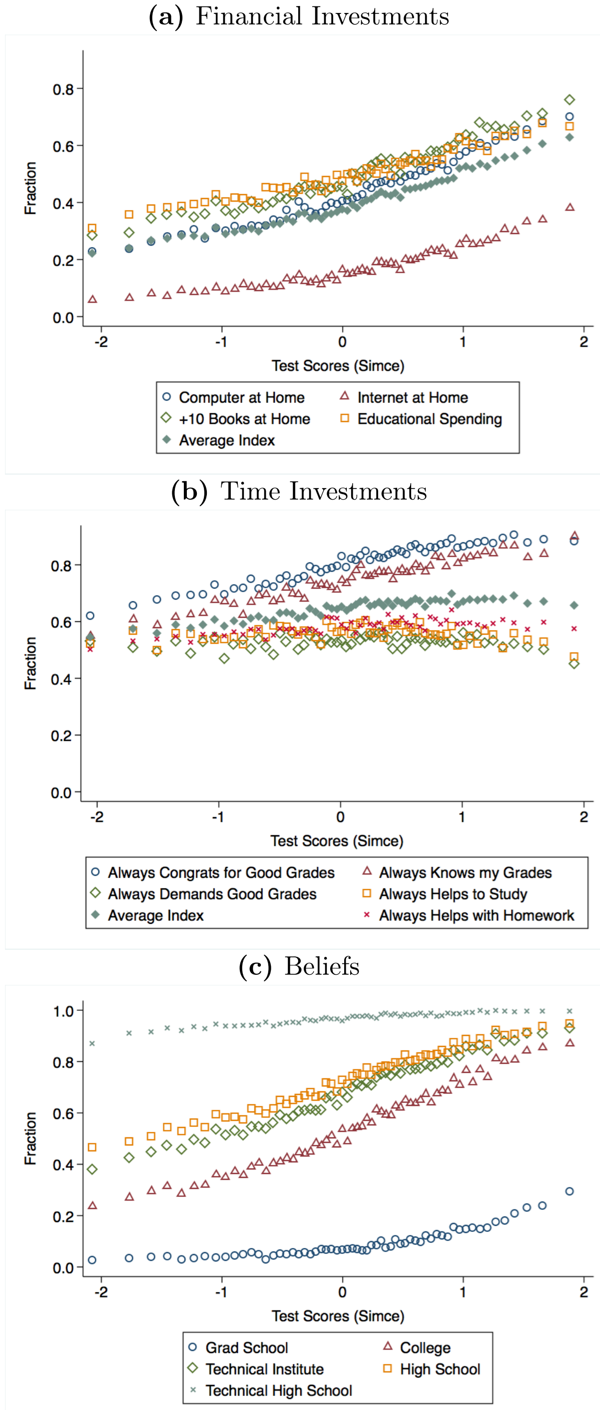
Parents also report their beliefs on their child's educational attainment in the future. We plot their answers vis a vis test scores in Figure 2c. Each answer takes value one when the parent reports that the child will attain at least the respective level of education. The data shows a positive correlation between higher expected educational attainment and test scores. The lower correlation comes from parents who expect their child to complete at least technical high school diploma, because a high fraction (more than 90%) thinks children will reach that educational level.

3.3 Long Run Outcomes

Our main long run outcomes are take-up rates and scores of the national college entrance exam, college enrollment and college graduation. We construct take-up and scores of the entrance exam measuring them up to age 20 for first graders. The college entrance exam is taken by the end of high school senior year, when students are approximately 17 to 18 years old. While every year some test-takers are older (the test-taker median age is 19 years old), a very small fraction of all test-takers (less than 5 percent) take the test after turning 20 years old. Therefore measuring take-up at age 20 is a good proxy of ever taking the entrance exam.

We measure college enrollment similarly for first graders. We define college enrollment as the rate of students who enroll as freshmen in any college in the country. As for the entrance exam take-up, a very small fraction of freshmen are older than 20 years old. We also measure enrollment at selective institutions. The selective universities are non-profit institutions, grouped in the *Council of Rectors of the Universities of Chile* (CRUCH), which receive students with highest scores in the country. Finally we measure college graduation as the rate of students who graduated from any college in the country at age 27.

Figure 2: Parental Investments, Beliefs and Test Scores



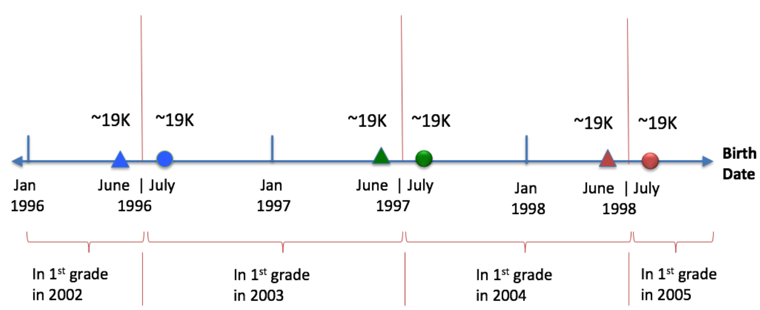
Note: The graphs in Figure 2 plot measures of parental investments and beliefs within equal sized bins of 4th grade test scores (math-language average, in standard deviation units). The y-axis variables are all binary in Figure 2a, Figure 2b and Figure 2c. The variables in Figure 2a are our measures of parental financial investments (having a computer, Internet connection and more than ten books at home, and spending above the median on educational items). The variables in Figure 2b are our measures of parental time investments, and take value one if the child answered that her parent does each activity "Always" and zero otherwise. The variables in Figure 2c are our measures of parental expectations, where parents answer their beliefs about child's expected educational attainment.

3.4 Working Sample

We study two sets of student cohorts. The first cohorts corresponds to first graders in years 2002 to 2005 (born between 1996 to 1998), and the second cohorts consist of eighth graders in the same years (born between 1989 and 1991). The data permits measuring outcomes up to age 20 for first graders because the youngest first graders were born in 1998 and we have data up to year 2019. Analogously, the younger eighth graders were born in 1991, and hence we can follow them until they are 27 years old in 2019.

We build our working samples for first graders and eight graders as follows. After excluding the 7 percent of children enrolled in private schools who do not use the July 1 cutoff to enroll students, our administrative records contain approximately one million children in the first grade (N = 987,264) and eighth grade (N = 1,048,983). The first sample is composed of first graders born in June and July from years 1996 to 1998, and the second contains eighth graders born in June and July from 1989 to 1991. Figure 3 shows our research design for first graders. Those born in July just missed the cutoff date and therefore would start the first grade in the next academic year. For example, those born in June from 1996 would start the first grade in 2002, while those born in July of the same year start in 2003. We exploit the three discontinuities occurring between June and July from years 1996 to 1998 and pool our sample according to month of birth (June of July) in year T, and school starting date, in year T+6 or year T+7. We control for cohort differences in our analyses.

Figure 3: Research Design for 1st Graders



Note: Figure 3 illustrates our research design for first graders. About 19K children were born in either June or July in years 1996 to 1998. According to the age at entry rule those born in June from year T should start the first grade in year T+6, while July-born children should start in year T+7.

Ideally, we would like to have birth records to avoid attrition between birth and first grade enrollment. In addition, if that attrition was differential by month of birth, it would also affect the internal validity of our analysis. We believe that neither is an important problem in the Chilean context because first grade enrollment is mandatory and compliance is very high nationwide. According to official vital statistics (MINSAL, 1996), the number of births was close to 21,000 each month for the years we study. If we exclude 7 percent of the children (those enroll in private schools), the total number of monthly births is very close to our sample of 19,000 per month. In addition, the same source indicates that the number of births was evenly distributed by month of birth, as we also find in our data with first-grade enrollment.

3.5 Summary Statistics

Table 1 presents mean characteristics for students in our working samples. Column (1) presents values for our working sample of first graders, while column (2) does the same for the total population of first graders as a benchmark. Columns (3) and (4) describe eighth graders analogously. Overall, Table 1 shows that our working samples and the student population are fairly similar in a host of individual and school baseline characteristics, suggesting that results using our working samples are not prone to external validity bias (Andrews and Oster, 2019). Table 1 also suggests that births are uniformly distributed by month because half of the students in our working samples were born in June and half in July, and the fraction born each month is 8 percent of the respective benchmark population. This result is consistent with the fact that each month of a given year accounts for approximately 8.3 percent of the births.

Table 1: Summary Statistics

	(1)	(2)	(3)	(4)
	1st Graders		8th Graders	
Variable	Sample	All	Sample	All
Born in June	0.50	0.08	0.50	0.08
Born in July	0.50	0.08	0.50	0.08
Father's Schooling	10.95	10.89	N.A.	N.A.
	(3.61)	(3.63)		
Mother's Schooling	10.88	10.82	N.A.	N.A.
	(3.52)	(3.53)		
Girl	0.49	0.49	0.51	0.50
Class Size	30.27	30.12	32.80	32.35
School Vulnerability (0-100)	29.84	30.00	28.29	29.74
School in Capital Region	0.37	0.37	0.37	0.37
Rural School	0.14	0.14	0.11	0.13
Public School	0.52	0.53	0.57	0.59
Voucher School	0.48	0.47	0.43	0.41
Observations	117,709	987,264	111,664	1,048,983

Notes: Table 1 shows the mean of each variable in rows, with standard deviation in parentheses for non-dichotomic variables. The first two columns describe students in first grade in 2002 to 2005. Column (1) presents values for our working sample of first graders, while column (2) does the same for the full population of 1st graders as a benchmark. Columns (3) and (4) describe 8th graders analogously. The variable 'Born in June (July)' take value 1 if the student was born in June (July). The measures of parental schooling refer to years of completed education and come from SIMCE surveys that have a response rate of 75% for both our working sample and all first graders. Parental schooling is not available for eighth graders since there was no SIMCE survey implemented in the '90s, when these students were in the fourth grade. The school vulnerability index measures percentage of students receiving free or reduced price lunch. The schools type is either public or voucher.

The descriptive statistics situate the sample in a context of a middle-income country. For instance, average parental schooling is close to 11 years, which is less than the 12 years needed to get a high school diploma. Levels of schooling in Chile are higher today,[5] but our data describe students and their parents about 15 years ago, when the country exhibited lower levels of development. The average class size for first graders is about 30 students and 32 students for an average eighth grade class, which again are similar to rates in developing countries. For reference, at about the same time (in the mid-2000s), the class size in primary school was 21 in the US and 27 in Turkey (OECD Stats 2019).

The data also show that students attend schools with a vulnerability index close to 30, on average. This index ranges between 0 and 100 and resembles the percentage of students receiving free or reduced-price lunch, similar to the index often used in the US as a proxy for poverty. In Chile, this index is computed by the government agency responsible for school meal programs (National School Assistance and Scholarship Board, JUNAEB). The index considers poverty of students and risk of dropout as factors.[6] Approximately 37 percent of the students attend schools within the metropolitan region, which includes the national capital, Santiago, and 14 percent of first graders and 11 percent of eighth graders attend schools in rural areas, with the remaining 49 percent of students attending school in urban areas outside the Metropolitan Region. Finally, 52 percent of first graders and 57 percent of eight graders attend public schools, with the remaining fraction attending voucher schools.[7]

4 Results

We start by briefly presenting results on in-school outcomes to focus then on the effects on parental investments and beliefs. Next we describe our estimates on long run outcomes, and finish the results section discussing our findings by socioeconomic status.

4.1 In-school Effects

In Figure 4 and Table 2 we present effects on in-school outcomes for our sample of first graders. July-born children start the first grade 0.48 years older than their June-born counterparts (see Figure 4a) and so are more likely to enjoy a skill advantage over those who start younger, as discussed previously and depicted in Figure 1. Figures 4b, 4c, and 4d show that a skill advantage in the first grade translates into higher GPAs (0.26σ) and higher passing rates (2.4 percentage points (pp)) in the first grade and higher test scores (0.21σ) in the fourth grade.

These results on in-school outcomes are consistent in direction and magnitude with the related literature. In the next sections we complement the findings on educational outcomes exploring how parents react to these signals of children ability. We then supplement the short run effects during school with estimates of long run outcomes.

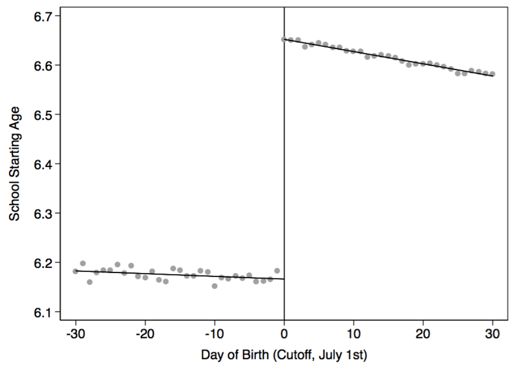
[5] Chile has reached almost universal levels of educational coverage in primary (99 percent) and secondary school (92 percent), well above Latin American countries (Unesco-OECD 2010; IDB 2018).

[6] For details see JUNAEB (2019).

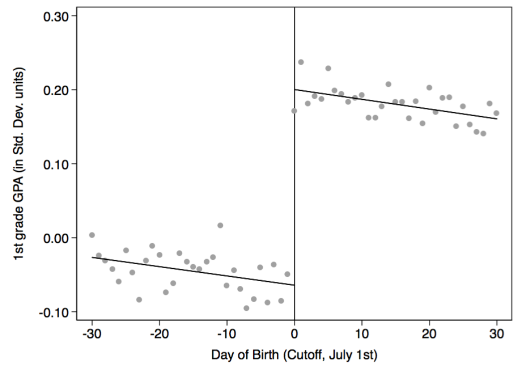
[7] Public schools are both publicly funded and administered. Voucher schools receive public funding but are privately managed, similar to charter schools in the US.

Figure 4: In-School Effects

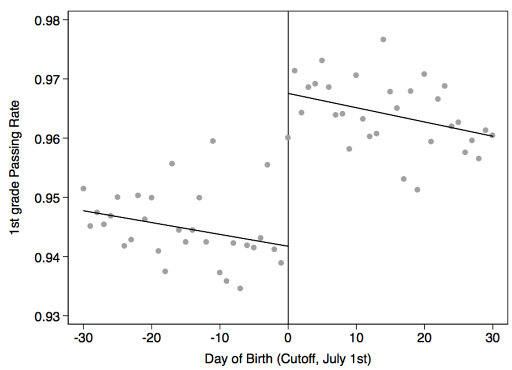
(a) School Starting Age



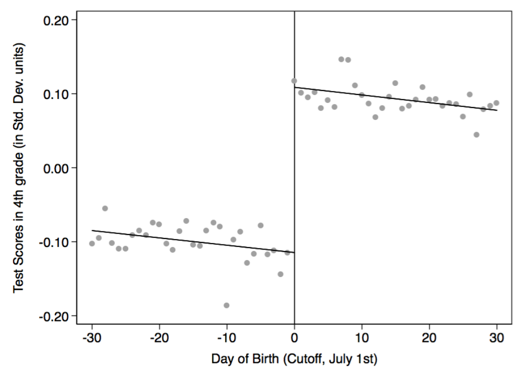
(b) GPA in 1st Grade



(c) Passing Rate in 1st Grade



(d) Test Scores in 4th Grade



Note: The graphs in Figure 4 plot the mean of the y-axis variable within day of birth, and fit estimated lines using all the underlying data, allowing for different slopes on each side of the cutoff. Each day of birth contains about 2K observations. Table 2 shows the results from the estimation of equation (1) for each of these outcomes.

Table 2: In-School Effects

	(1) Age at Entry	(2) GPA in 1st Grade	(3) Pass Rate in 1st Grade	(4) Test Scores in 4th Grade
$\widehat{\alpha_1}$	0.482***	0.263***	0.024***	0.207***
	(0.003)	(0.012)	(0.003)	(0.010)
June Mean	6.169	-0.059	0.944	-0.111
Observations	117,709	117,709	117,709	117,709

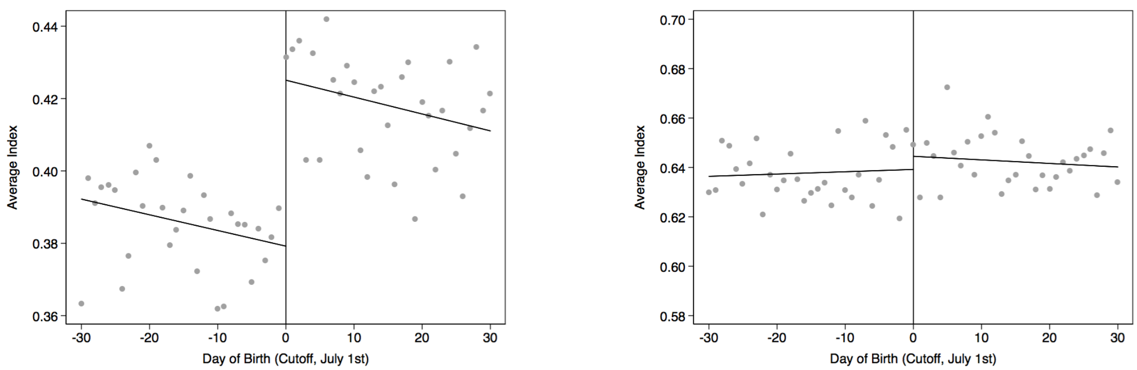
Notes: Table 2 shows the coefficient $\widehat{\alpha_1}$ estimated from the equation (1) for first graders on age at school entry, and in-school outcomes. The 'June Mean' is the mean of the dependent variable just below the threshold. The outcomes are first grade GPA (standardized within school and grade) and pass rate, and the Language-Math score in 4th grade. All estimations include cohort fixed effects and control for child gender, class size, school vulnerability, school rurality and type of school (public or voucher). Robust standard errors (in parentheses) are clustered by day of birth.

4.2 Effects on Parental Investments and Beliefs

Investments. Figure 5 summarizes our main findings on parental investments. Figure 5a shows that July-born children receive 3.4pp ($\sim 0.11\sigma$) of additional financial investments, while Figure 5b shows no effects on time investments. Going back to our conceptual framework, these findings suggest that parents reinforce skill gaps using financial investments, but do not use time investments to respond to differences in school performance.

Figure 5: Effects on Parental Investments

(a) Financial Investments **(b)** Time Investments



Note: The graphs in Figure 5 plot the mean of the y-axis variable within day of birth, and fit estimated lines using all the underlying data, allowing for different slopes on each side of the cutoff.

We present the corresponding point estimates in Table 3. In Appendix B we provide details on sample sizes and robustness checks. Panel A in Table 3 shows the results for financial investments. Column (1) shows the effect on the average index, plotted in Figure 5a. Parents increase average financial investments by 3.4 pp, which represents an effect of 0.11σ. In column (2) we show the effect on the factor index which shows the same effect of 0.11σ. Next, columns (3)-(6) show the effects for each investment variable separately. July-born children are 10 percent (4.3 pp over 42%) more likely to have a computer at home, 20 percent (3.2 pp over 16%) more likely to have an Internet connection, and 8 percent (3.7 pp over 48%) more likely to have ten or more books at home. In addition to investing in more educational assets, parents are 5 percent (2.2 pp over 49%) more likely to spend above the median of monthly expenditures for school items in our sample.

We examine parental responses in terms of time investments in Panel B of Table 3. Column (1) shows that the estimate on the average index, plotted in Figure 5b, is a precise zero (0.4 pp over a mean of 64%). The factor index effect is of 0.01σ and not different from zero, as shown in column (2). When revising effects for each variable in particular, we also find precise zero effects in whether parents congratulate their child for good grades (mean of 81%) and whether they know their grades (75%). This is despite the strong positive correlation between these variables and test scores shown in Figure 2b. We also find precise zeros in whether parents demand good grades (mean of 53% for both groups) and whether they help with study (56%). Finally, we find a small positive effect on whether parents help with homework (2pp over a mean of 57%).

Table 3: Effects on Parental Investments

Panel A: Financial Investments

	(1) Average Index	(2) Factor Index	(3) Computer at Home	(4) Internet at Home	(5) More 10 Books	(6) High Spending
$\widehat{\alpha_1}$	0.034***	0.108***	0.043***	0.032***	0.037***	0.022***
	(0.005)	(0.016)	(0.008)	(0.006)	(0.008)	(0.007)
June Mean	0.379	-0.070	0.404	0.158	0.478	0.478
Effect Size	0.107	0.110	0.107	0.204	0.078	0.047
Observations	51,818	51,818	51,818	51,818	51,818	51,818

Panel B: Time Investments: 'My parent always...'

	(1) Average Index	(2) Factor Index	(3) Congrats for good grades	(4) Knows my grades	(5) Demands good grades	(6) Helps to study	(7) Helps with homework
$\widehat{\alpha_1}$	0.004	0.013	-0.010	0.010	0.001	0.001	0.018**
	(0.006)	(0.022)	(0.006)	(0.010)	(0.012)	(0.009)	(0.009)
June Mean	0.643	0.011	0.810	0.752	0.526	0.559	0.569
Effect Size	0.015	0.014	-0.012	0.013	0.002	0.001	0.032
Observations	47,646	47,646	47,646	47,646	47,646	47,646	47,646

Notes: Table 3 shows the coefficient $\widehat{\alpha_1}$ estimated from the equation (1) on measures of parental financial investment (Panel A), and parental time investments (Panel B), previously presented in Figure 2a and Figure 2b. The 'June Mean' is the mean of the dependent variable just below the threshold, and the 'Effect Size' corresponds to $\widehat{\alpha_1}$/(June Mean) for binary outcomes, and to $\widehat{\alpha_1}$/ (standard deviation below the threshold) for non-binary outcomes. The dependent variables are described in subsection 3.2. All estimations include cohort fixed effects and controls for child gender, class size, school vulnerability, school rurality and type of school (public or voucher). Robust standard errors (in parentheses) are clustered by day of birth.

Beliefs. Parents of July-born children also have higher expectations about their child's educational future, driven by beliefs on children completing postsecondary degrees as we show in Table 4. Parents are more likely to believe that their child will complete college (4 percent, or 2.1 pp over 53%), grad school (13 percent, or 1.2pp over 8.7%), and a 4-year degree at a technical institute (2 percent, or 1.3 pp over 67%). We found no effects on beliefs on high school completion or technical high school completion (columns (5) and (6)).

This finding is consistent with beliefs responding positively to signals about the child's ability, which parents have been receiving between 1st and 4th grade. Survey data from SIMCE indicates that 75 percent of students report that their parents know their grades (Panel B, column (4) in Table 3). Therefore parents of July-born children are mostly aware that they perform relatively well, while parents of June-born children also know that their child is performing relatively worse.

Table 4: Effects on Parental Beliefs

	(1) College Expectation	(2) Grad School Expectation	(3) Institute Expectation	(4) High School Expectation	(5) Tech HSchool Expectation
$\widehat{\alpha_1}$	0.021***	0.012**	0.013*	0.013	0.003
	(0.007)	(0.005)	(0.008)	(0.008)	(0.003)
June Mean	0.530	0.087	0.666	0.709	0.963
Effect Size	0.039	0.133	0.020	0.019	0.003
Observations	51,818	51,818	51,818	51,818	51,818

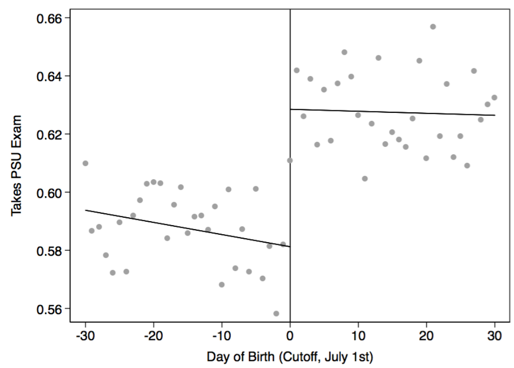
Notes: Table 4 shows the coefficient $\widehat{\alpha_1}$ estimated from the equation (1) on measures of parental expectations described in Figure 2c. The 'June Mean' is the mean of the dependent variable just below the threshold, and the 'Effect Size' corresponds to $\widehat{\alpha_1}$/(June Mean) for binary outcomes. The dependent variables are described in subsection 3.2. All estimations include cohort fixed effects and controls for child gender, class size, school vulnerability, school rurality and type of school (public or voucher). Robust standard errors (in parentheses) are clustered by day of birth.

4.3 Long-Run Outcomes

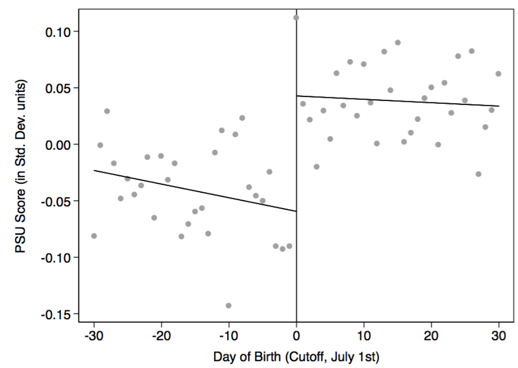
Figure 6 summarizes our main effects on long run outcomes. Following up children until they are 20 years old, we find that the early skill gaps lead to higher college entrance exam take-up (4 percentage points over a mean of 58%), scores (0.08σ), and higher college enrollment rates: overall (4 pp. over 25%) and at selective programs (3 pp. over 14%).

Figure 6: Effects on Long Run Outcomes

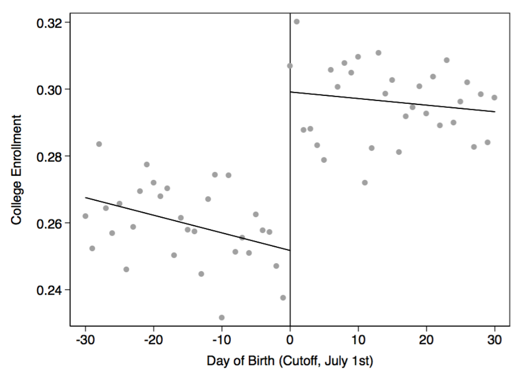
(a) Taking College Entrance Exam



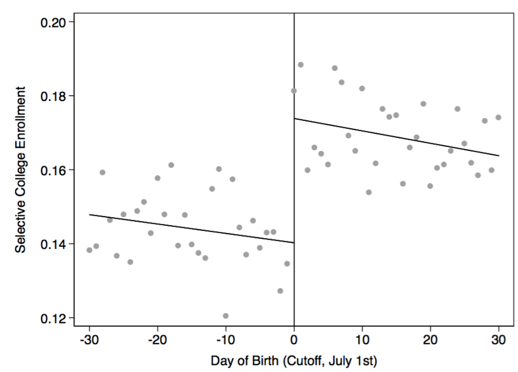
(b) Scores in the College Entrance Exam



(c) Overall College Enrollment



(d) Selective College Enrollment



Note: The graphs in Figure 6 plot the mean of the y-axis variable within day of birth, and fit estimated lines using all the underlying data, allowing for different slopes on each side of the cutoff. Each day of birth contains about 2K observations.

We show our point estimates in Table 5. Column (3) presents the estimate corresponding to Figure 6a showing that July-born students are 6 percent (3.6 pp over a mean of 58 percent) more likely to take the national college entrance exam.

Conditional on taking the test, students with an early skill advantage score 0.08σ higher, as shown in Figure 6b and in column (4) of Table 5. We interpret this effect as a lower bound because, among non-test-takers, those who start school with a skill advantage would arguably have performed better had they taken the test. In any case, if there are positive effects on the college entrance exam (taking the exam or scoring higher), these should translate into effects on college enrollment.

The next results show indeed effects on multiple measures of college enrollment, reported in columns (5) to (8) in Table 5. We find a 15 percent increase (3.7 pp over a mean of 25 percent) in college enrollment (see Figure 6c) for July-born students, which is consistent with the positive effects on the likelihood of being a test-taker. Meanwhile, the effects on enrollment at more selective universities (20 percent; 2.8 pp over a mean of 14 percent, see Figure 6d) and STEM programs (14 percent; 1 pp over a mean of 7 percent) are consistent with higher scores on the college entrance exam.

Table 5: Effects on Long Run Outcomes

	(1) Primary Grad	(2) High-School Grad	(3) PSU Exam	(4) PSU Score	(5) College Enroll	(6) Selective Enroll	(7) STEM Enroll	(8) College Grad
$\widehat{\alpha_1}$	0.013***	0.004	0.036***	0.084***	0.037***	0.028***	0.010***	0.007*
	(0.003)	(0.005)	(0.006)	(0.017)	(0.005)	(0.004)	(0.003)	(0.004)
June Mean	0.912	0.684	0.580	-0.067	0.253	0.138	0.069	0.165
Effect Size	0.014	0.006	0.063	.	0.147	0.203	0.141	0.042
Observations	117,709	117,709	117,709	71,509	117,709	117,709	117,709	111,664

Notes: Table 5 show the coefficient $\widehat{\alpha_1}$ estimated from the equation (1) on a host of outcomes. The 'June Mean' is the mean of the dependent variable just below the threshold, and the 'Effect Size' corresponds to $\widehat{\alpha_1}$/(June Mean) for binary outcomes. Table 5 presents results for our sample of first graders in columns (1) to (7) and for eight graders in column (8). All the dependent variables are outcomes measured at age 20, except the last outcome, measured at 27. 'Primary Grad' and 'High-School Grad' are primary and high-school graduation rates, respectively; 'PSU Exam' and 'PSU Score' measure whether children took the college entrance exam and their scores if they did. 'College Enrollment', 'Selective Enroll' and 'STEM Enroll' measure whether children enrolled at any college, at selective institutions and at STEM majors, respectively. 'College grad' indicates whether the children graduated from college. All estimations include cohort fixed effects and control for child gender, class size, school vulnerability, school rurality and type of school (public or voucher). Robust standard errors (in parentheses) are clustered by day of birth.

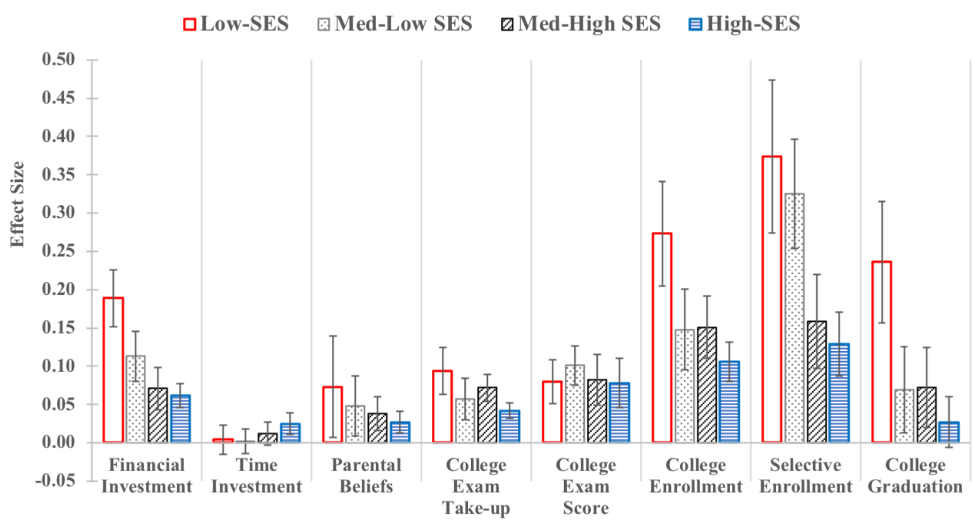
We finally use our sample of eighth graders to estimate effects on college graduation later on. The last column of Table 5 shows a precisely estimated effect of 0.7pp., which represents an effect size of 4 percent (0.7 pp over a mean of 16.5%). On average, a share of 16.5 percent of both June- and July-born students obtain a college degree by age 27. These rates are similar to back of the envelope computations from official reports by MINEDUC on higher education completion rates (MINEDUC, 2019).

4.4 Effects by Socioeconomic Background

Figure 7 and Table 6 summarize our main results on parental investments, beliefs, and children long run outcomes by socioeconomic status (SES). The main take away is that effect sizes on financial investments and college related outcomes are larger for lower SES children with a clear negative SES gradient in most outcomes.

We divide the sample of students into quartiles of the national vulnerability index and label each quartile as low SES, med-low SES, med-high SES, and high SES, accordingly. In this section we present our findings using effect sizes as they are informative of the relative importance of the effects for each SES group. We show results for financial and time investments using the respective average index described in subsection 3.2, and using college completion expectations for parental beliefs. In Appendix C we include results for each individual measure of investment and beliefs, that behave similarly to the summary measures presented here. Long run outcomes consist on the college entrance exam take-up and scores, college enrollment (overall and at selective institutions), and college graduation.

Figure 7: Effect Sizes by Socioeconomic Status



Note: Figure 7 plots effect sizes (with their standard errors) on parental investments, beliefs, and children's long run outcomes by quartiles of the national vulnerability index. Financial and time investments are each measured by the average index described in subsection 3.2. Parental beliefs are measured as college completion expectations. The long run outcomes are the college entrance exam take-up and scores, college enrollment (overall and at selective institutions), and college graduation.

Parental Investments and Beliefs. The effects on parental financial investments display a steep, negative SES gradient. The effect sizes are three times as large for the low-SES versus the high-SES group (19 vs 6 percent). Our estimates on time investments are close to zero with no distinguishable differences across groups. This result suggests that the average null effect shown in Table 3 did not hide differential effects by socioeconomic status.

Effects on parental beliefs, measured as college expectations, are decreasing by socioeconomic status but measured noisily. There is a difference of 4.6 percentage points between the low and high

SES groups that we can not distinguish from zero.

Long-Run Outcomes. Our estimates on the college exam take-up show an effect of 9 percent for low SES students, about 6 percent for students in the med-low and med-high group and 4 percent for high SES students. Test scores increase similarly by about 0.08σ for each group, but from a lower baseline score for the more disadvantaged SES groups (i.e., -0.42σ for low-SES versus 0.37σ for high-SES). These results suggests that we should see a relatively larger effect on college enrollment for lower SES students.

Consistently, for college enrollment, we find effect sizes of 27 percent for the low SES group, 15 percent for the med-low and med-high SES and 11 percent for the high SES group. This effect pattern is also present in our measures of enrollment at selective programs, with effects of 37% for low SES and 13% for high SES students. Finally, the results also show that the null average effect on college graduation in Table 5 was masking large SES differences. Low SES students with a skill advantage are 23% more likely to graduate from college, versus a noisily estimated 3% for the higher SES group .

Table 6: Effect Sizes by Socioeconomic Status

Outcome	(1) Low SES	(2) Med-Low SES	(3) Med-High SES	(4) High SES	(5) Low vs High Difference
Financial Investments	0.189***	0.113***	0.071***	0.062***	0.127***
Average Index	(0.0372)	(0.0330)	(0.0274)	(0.0155)	(0.040)
Time Investments	0.004	0.002	0.012	0.025*	-0.021
Average Index	(0.0194)	(0.0157)	(0.0151)	(0.0144)	(0.024)
Parental Beliefs	0.073	0.048	0.038*	0.027*	0.046
College Expectation	(0.0663)	(0.0393)	(0.0217)	(0.0144)	(0.068)
Long Run Outcomes					
Takes PSU Exam	0.094***	0.057**	0.072***	0.042***	0.052
	(0.0307)	(0.0273)	(0.0174)	(0.0100)	(0.032)
PSU Exam Scores	0.080***	0.101***	0.082**	0.078**	0.002
	(0.0286)	(0.0256)	(0.0333)	(0.0322)	(0.043)
College Enrollment	0.273***	0.148***	0.151***	0.106***	0.167**
	(0.0680)	(0.0525)	(0.0403)	(0.0255)	(0.073)
Selective College Enrollment	0.374***	0.325***	0.159***	0.129***	0.245**
	(0.1001)	(0.0713)	(0.0612)	(0.0419)	(0.109)
College Graduation	0.236***	0.069	0.072	0.027	0.209**
	(0.0792)	(0.0563)	(0.0522)	(0.0335)	(0.086)

Notes: Table 6 shows the effect sizes on parental investments, beliefs, and children long run outcomes by quartiles of the national vulnerability index. We label each quartile as low SES, med-low SES, med-high SES, and high SES, in columns (1)-(4). Column (5) show the difference between the low and high SES effect size. Standard errors for the effect sizes are computed using the delta method.

4.5 Discussion

Our empirical results show that the long-run effects of an early skill advantage, represented by A in Equation 6, are positive and large. For example, the probability of college enrollment increases by 15 percent overall. In our conceptual framework, we decompose this global effect into a direct effect component and two components related to money-intensive and time-intensive investments, C and D respectively. Our findings show that monetary investments reinforce the shock (i.e., C is positive) while time-intensive investments are neutral (i.e., D is zero). This suggests that the overall effect of a skill advantage at school entry, A, is explained by a direct effect and by reinforcing financial investments from parents, $B + C$.

In our model, optimal investments of parents are a function of beliefs, $\widetilde{\theta}_0$, about their child's abilities. We find that parents have higher beliefs about their children's future human capital. This result is consistent with parents interpreting in-school performance as signals of their child ability, and adjusting their investments according to their perceived productivity. Such mechanism is in line with Attanasio et al. (2020) who show that investments vary according to parental beliefs on heterogeneity of returns to such investments. In our context, parents may perceive that investing in books, computer, or other school related materials can be productive to complement the skills of children with higher grades. Our null results on time-intensive investments suggests that parents perceive their productivity to be similar for children with different skills at school entry. Precisely identifying the interactions between shocks and investments would require an exogenous variation on investments. Examples of recent papers studying these type of interactions are Duque et al. (2018), Johnson and Jackson (2017), Malamud et al. (2016), and Rossin-Slater and Wüst (2019).[8]

In addition, we find informative to compare our main effect on college enrollment (15 percent) to some of the popular early childhood interventions in the literature. Elango et al. (2015) provide an excellent review of these programs. For Head Start, Currie et al. (2002) show that there is a marginally significant increase in nine percent in the probability of attending college when they compare Head Starters to non-Head Starters. Likewise, Ludwig and Miller (2007) find that Head Start increases the likelihood of attending college by five percent. Deming (2009) shows that the same program increases the probability of attending college by ten percent approximately.[9] Anderson (2008) shows that participants of the Perry Preschool Program are 21 percent more likely to attend any college.[10] Our estimated effect on college enrollment is within those found by the early childhood interventions literature, suggesting that policy shocks on early skills in developing countries can be as important as programs designed to bolster children's abilities.

Our results have direct policy implications. Age-at-entry rules may generate noisy signals of skill because young starters of high ability signal low ability (due to age) as they perform worse in

[8] In the absence of an additional instrument to correct for endogenous parental behavior, other papers jointly model parental behavior and interventions or shocks in structural models used to isolate parameters of parental behavior from the human capital production function. Examples of recent papers in this area are Attanasio et al. (2019b) and Jervis (2017).

[9] Deming (2009) report their results in percentage points. We compute this relative to the control mean reported by Ludwig and Miller (2007) who study the same program and outcome using Census data.

[10] Anderson (2008) reports results separately by gender, which we average to compute the overall effect.

tests, and hence receive less investments than their older, high-ability equivalents. These signals would disappear if adjusted by age, but parents and children themselves observe and react based on the unadjusted in-school differences. From a policy perspective, there are not many alternatives for admissions that are not age based. Admission on a rolling basis throughout the year could smooth out these effects, but it may be impracticable. The most direct policy recommendation would be to make parents aware of the age-differences, providing them with both raw and age-adjusted measures of children's performance. In addition to this information, schools could train teachers to communicate and interpret the results to parents, particularly during the early school years when age-performance gaps are larger.

5 Conclusion

This paper studies the effects of having an early skill advantage at school entry on different types of parental investments and long-term outcomes. We combine rich administrative records with a regression discontinuity design in a middle-income country, and interpret our results within a human capital accumulation model. We find differential effects by type of investment, and lasting effects on educational outcomes after twenty years.

In particular, we find that parents do not invest time differentially, but do reinforce the skill gaps investing additional financial resources in children with higher ability at entry. Moreover, the reinforcing financial investments are more pronounced for children with lower socioeconomic backgrounds.

We further document that the early skill advantage translates into higher rates of college enrollment and at selective programs, with larger effects for low socioeconomic status students as well. These results add to the literature by emphasizing that early skill gaps can have sizable and heterogeneous consequences in adulthood in developing countries.

Our findings underscore the importance of developing a research agenda that studies several types of parental investments and connects them with long run outcomes. Such agenda would deepen our understanding of mediators of early shocks on inequality in adulthood outcomes.

References

Almond, D., Currie, J., and Duque, V. (2018). Childhood circumstances and adult outcomes: Act ii. *Journal of Economic Literature*, 56(4):1360–1446.

Almond, D. and Mazumder, B. (2013). Fetal origins and parental responses. *Annu. Rev. Econ.*, 5(1):37–56.

Anderson, M. L. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the abecedarian, perry preschool, and early training projects. *Journal of the American statistical Association*, 103(484):1481–1495.

Andrews, I. and Oster, E. (2019). A simple approximation for evaluating external validity bias. *Economics Letters*, 178:58–62. Working Paper.

Attanasio, O., Boneva, T., and Rauh, C. (2019a). Parental beliefs about returns to different types of investments in school children. Working Paper 25513, National Bureau of Economic Research.

Attanasio, O., Cattan, S., Fitzsimons, E., Meghir, C., and Rubio-Codina, M. (2019b). Estimating the production function for human capital: Results from a randomized control trial in colombia. *American Economic Review*, forthcoming.

Attanasio, O., Cattan, S., Fitzsimons, E., Meghir, C., and Rubio-Codina, M. (2020). Estimating the production function for human capital: results from a randomized controlled trial in colombia. *American Economic Review*, 110(1):48–85.

Barua, R. and Lang, K. (2016). School Entry, Educational Attainment, and Quarter of Birth: A Cautionary Tale of a Local Average Treatment Effect. *Journal of Human Capital*, 10(3):347–376.

Becker, G. S. and Tomes, N. (1976). Child endowments and the quantity and quality of children. *Journal of political Economy*, 84(4, Part 2):S143–S162.

Behrman, J. R., Pollak, R. A., and Taubman, P. (1982). Parental preferences and provision for progeny. *Journal of Political Economy*, 90(1):52–73.

Beuermann, D. W. and Jackson, C. K. (2018). Do parents know best? the short and long-run effects of attending the schools that parents prefer. Technical report, National Bureau of Economic Research.

Bharadwaj, P., Eberhard, J. P., and Neilson, C. A. (2018). Health at birth, parental investments, and academic outcomes. *Journal of Labor Economics*, 36(2):349–394.

Bharadwaj, P., Løken, K. V., and Neilson, C. (2013). Early life health interventions and academic achievement. *The American Economic Review*, 103(5):1862–1891.

Biroli, P., Boneva, T., Raja, A., and Rauh, C. (2018). Parental beliefs about returns to child health investments. Working Paper 11336, IZA Institute of Labor Economics.

Black, S. E., Devereux, P. J., and Salvanes, K. G. (2011). Too young to leave the nest? the effects of school starting age. *The Review of Economics and Statistics*, 93(2):455–467.

Boneva, T. and Rauh, C. (2018). Parental beliefs about returns to educational investments – the later the better? *Journal of the European Economic Association*, 16(6):1669–1711.

Bono, E. D., Francesconi, M., Kelly, Y., and Sacker, A. (2016). Early maternal time investment and early child outcomes. *The Economic Journal*, 126(596):F96–F135.

Carneiro, P., Galasso, E., Garcia, I. L., Bedregal, P., and Cordero, M. (2019). Parental Beliefs, Investments, and Child Development: Evidence from a Large-Scale Experiment. Working Papers 2019-051, Human Capital and Economic Opportunity Working Group.

Cattaneo, M. D., Jansson, M., and Ma, X. (2018). Manipulation testing based on density discontinuity. *The Stata Journal*, 18(1):234–261.

Cunha, F., Elo, I., and Culhane, J. (2013). Eliciting maternal expectations about the technology of cognitive skill formation. Working Paper 19144, National Bureau of Economic Research.

Cunha, F., Heckman, J. J., and Schennach, S. M. (2010). Estimating the technology of cognitive and noncognitive skill formation. *Econometrica*, 78(3):883–931.

Currie, J., Thomas, D., and Garces, E. (2002). Longer-term effects of head start. *American economic review*, 92(4):999–1012.

Dahl, G. B. and Lochner, L. (2012). The impact of family income on child achievement: Evidence from the earned income tax credit. *American Economic Review*, 102(5):1927–56.

Deming, D. (2009). Early childhood intervention and life-cycle skill development: Evidence from head start. *American Economic Journal: Applied Economics*, 1(3):111–34.

Dhuey, E., Figlio, D., Karbownik, K., and Roth, J. (2019). School starting age and cognitive development. *Journal of Policy Analysis and Management*, 38(3):538–578.

Dizon-Ross, R. (2019). Parents' beliefs about their children's academic ability: Implications for educational investments. *American Economic Review*, 109(8):2728–65.

Doepke, M. and Zilibotti, F. (2017). Parenting with style: Altruism and paternalism in intergenerational preference transmission. *Econometrica*, 85(5):1331–1371.

Duque, V., Rosales-Rueda, M., and Sanchez, F. S. (2018). How do early-life shocks interact with subsequent human-capital investments? evidence from administrative data. Technical report, Mimeo.

Elango, S., García, J. L., Heckman, J. J., and Hojman, A. (2015). Early childhood education. In *Economics of Means-Tested Transfer Programs in the United States, Volume 2*, pages 235–297. University of Chicago Press.

Francesconi, M. and Heckman, J. J. (2016). Child development and parental investment: Introduction. *The Economic Journal*, 126(596):F1–F27.

Heckman, J. J., Stixrud, J., and Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor economics*, 24(3):411–482.

Jervis, P. (2017). A dynamic model of early parental investments in children's human capital. Technical report, Mimeo.

Johnson, R. C. and Jackson, C. K. (2017). Reducing inequality through dynamic complementarity: Evidence from head start and public school spending. Technical report, National Bureau of Economic Research.

Jones, D. (2015). The Economics of Exclusion Restrictions in IV Models. NBER Working Papers 21391, National Bureau of Economic Research, Inc.

Lubotsky, D. and Kaestner, R. (2016). Do skills beget skills? evidence on the effect of kindergarten entrance age on the evolution of cognitive and non-cognitive skill gaps in childhood. *Economics of Education Review*, 53:194–206.

Ludwig, J. and Miller, D. L. (2007). Does head start improve children's life chances? evidence from a regression discontinuity design. *The Quarterly journal of economics*, 122(1):159–208.

Malamud, O., Pop-Eleches, C., and Urquiola, M. (2016). Interactions between family and school environments: Evidence on dynamic complementarities? Technical report, National Bureau of Economic Research.

Rossin-Slater, M. and Wüst, M. (2019). What is the added value of preschool for poor children? long-term and intergenerational impacts and interactions with an infant health intervention. *American Economic Journal: Applied Economics*.

Rubio-Codina, M., Araujo, M. C., Attanasio, O., Muñoz, P., and Grantham-McGregor, S. (2016). Concurrent validity and feasibility of short tests currently used to measure early childhood development in large scale studies. *PLoS One*, 11(8):e0160962.

Appendices

A Appendix: Robustness

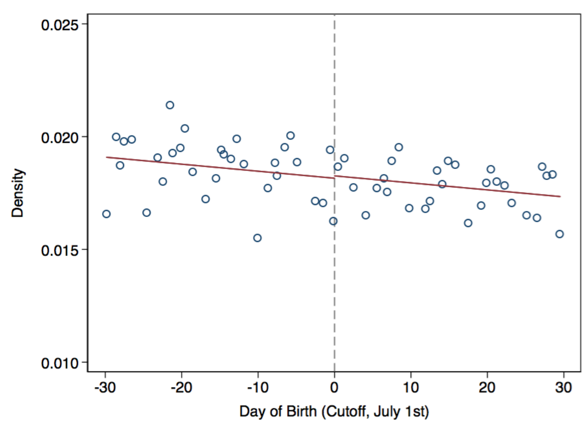
A.1 Density of Running Variable

We test for differences in unobserved characteristics by examining whether there is manipulation of birth dates near the cutoffs in our data. For example, it could be the case that more motivated parents planned the timing of their children's birth in order for them to be older when enrolling in primary school. If parents consider school starting age rules when timing conceptions or births dates by scheduling C-sections, for instance, our results would be subject to manipulation and sample selection bias.

In addition, we observe children once they are in first grade and ideally we would like to have birth records to perform our analysis. Data from official vital statistics (MINSAL 1996, 1997, 1998) show that the number of births is about 21K each month for the years we study. If we exclude the 7% of those children (who enroll in the private schools), then we get very close to our sample of 19K per month. In addition, the same source indicates that the number of births was evenly distributed by month of birth (taking into account the different number of days each month has), as we also find in our data with first grade enrollment.

We test for manipulation using a nonparametric test of discontinuity in the density of students born at each side of the eligibility rule, provided by Cattaneo et al. (2018). The manipulation test is -0.3668, with a p-value of 0.7138, which indicates that there is no statistical evidence of systematic manipulation of the running variable.

Figure A.1: Birth-Density per Day



Note: Figure A.1 plot the density of observations by each day in our data, and fits estimated lines using all the underlying data, allowing for different slopes on each side of the cutoff. The sample size is N=117,709.

Figure A.1 provides a graphical representation of the continuity in density test approach, plotting the density of observations by each day in our data. As we describe in the main text, we have on average about 2K observations per day. Dividing those observations over the total in our working sample for first graders (117K), we get a density value of about 0.017 each day, which is exactly what Figure A.1 shows. The density varies by holidays or weekends, and the fitted lines on both sides of the cutoff in Figure A.1 take that into account. This plot is consistent with the results from the formal test from Cattaneo et al. (2018), as the density estimates above and below the the cutoff (the two intercepts in the figure) are very near each other.

In addition to the nonparametric test by Cattaneo et al. (2018) we also test parametrically whether the density changes at the cutoff in Table A.1. Columns (1) to (7) show the coefficient $\widehat{\alpha_1}$ estimated from the equation (1) for first graders using the density of observations per day as dependent variable. The different columns add controls for weekends, holidays and birth year, and also vary the days near the cutoff used to run our regressions. The results are again consistent with both the graphical representation of the data and the nonparametric test, indicating no statistical evidence of systematic manipulation.

Table A.1: Testing Manipulation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\widehat{\alpha_1}$	0.001	0.000	0.000	0.000	0.001	-0.000	0.002
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Days near the Cutoff	30 days	30 days	30 days	30 days	20 days	10 days	3 days
Weekends	No	Yes	Yes	Yes	Yes	Yes	Yes
Holidays	No	No	Yes	Yes	Yes	Yes	Yes
Birth Year	No	No	No	Yes	Yes	Yes	Yes
Observations	117,709	117,709	117,709	117,709	79,007	40,303	13,167

Notes: Table A.1 show the coefficient $\widehat{\alpha_1}$ estimated from the equation (1) for first graders using the density of observations per day as dependent variable. Robust standard errors (in parentheses) are clustered by day of birth.

A.2 Covariates Smoothness

Our research design mimics a local experiment where children are exogenously (to potential outcomes) allocated to either being born in June or July. In this section we show that there are no other changes in our observable covariates occurring at the birth date threshold that could confound our analysis. Table A.2 shows the results of estimating equation (1) using each covariate in Table 1 as dependent variable. One important point to keep in mind is that we have much statistical power given our big sample sizes and therefore some point estimates are significant. However, the magnitudes are small enough to interpret those coefficient as precise zeros (effect sizes are never higher than 0.02) and moreover, our main estimates remain practically unchanged when we add covariates. In Table A.3 we show that point estimates change in the third decimal point.

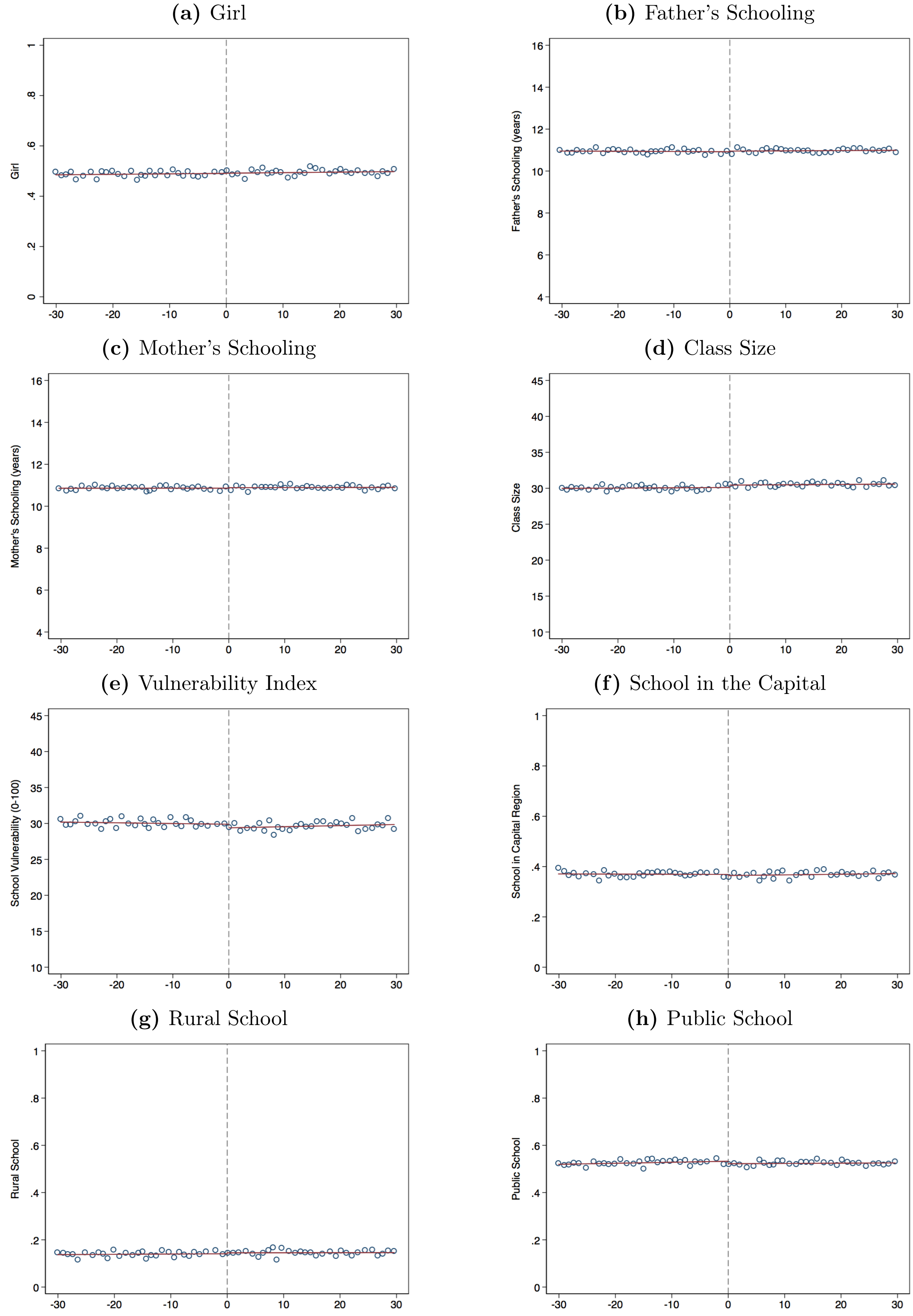
We complement these results with a graphical illustration for every covariate in Figure A.2, which provide further evidence of a smooth behavior at the July 1 cutoff.

Table A.2: Covariates Smoothness

	(1) Girl	(2) Father's Schooling	(3) Mother's Schooling	(4) Class Size	(5) IVE Index	(6) Capital Region	(7) Rural School	(8) Public School
$\widehat{\alpha_1}$	0.002	0.052	0.009	0.531	-0.003	-0.005	0.003	-0.011
	(0.006)	(0.041)	(0.041)	(0.102)	(0.002)	(0.005)	(0.003)	(0.004)
June Mean	0.487	10.897	10.823	29.920	0.303	0.366	0.146	0.531
$\widehat{\alpha_1}$/(June Mean)	0.005	0.005	0.001	0.018	0.011	0.015	0.019	0.021
Observations	117,709	85,753	89,404	117,709	117,709	117,709	117,709	117,709

Notes: Table A.2 show the coefficient $\widehat{\alpha_1}$ estimated from the equation (1) on covariates. Robust standard errors (in parentheses) are clustered by day of birth.

Figure A.2: Covariates Smoothness



Note: The graphs in Figure A.2 plot the mean of the y-axis variable within day of birth, and fit estimated lines using all the underlying data, allowing for different slopes on each side of the cutoff. Each day of birth contains about 2K observations. The y-axis variables are described in Table 1.

Table A.3: Robustness

		Primary Grad	High-School Grad	PSU Exam	PSU Score	College Enroll	Selective Enroll	STEM Enroll	College Grad
(A)	$\widehat{\alpha_1}$	0.01	0.01	0.04	0.08	0.05	0.03	0.01	0.01
		(0.00)	(0.01)	(0.01)	(0.02)	(0.01)	(0.00)	(0.00)	(0.00)
(B)	$\widehat{\alpha_1}$	0.01	0.01	0.04	0.08	0.04	0.03	0.01	0.01
		(0.00)	(0.01)	(0.01)	(0.02)	(0.01)	(0.00)	(0.00)	(0.00)
(C)	$\widehat{\alpha_1}$	0.01	0.00	0.04	0.07	0.04	0.03	0.01	0.00
		(0.00)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)
(D)	$\widehat{\alpha_1}$	0.01	0.00	0.03	0.07	0.04	0.03	0.01	0.00
		(0.00)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)

Notes: In row (A) we include cohort fixed effects, weekends,and holidays; in row (B) we include controls in (A) plus the Demographics (Female, Class Size, Rural School, School in Capital Region); in row (C) we include controls in (B) plus the Vulnerability Index; in row (D) we include controls in (C) plus the Public School

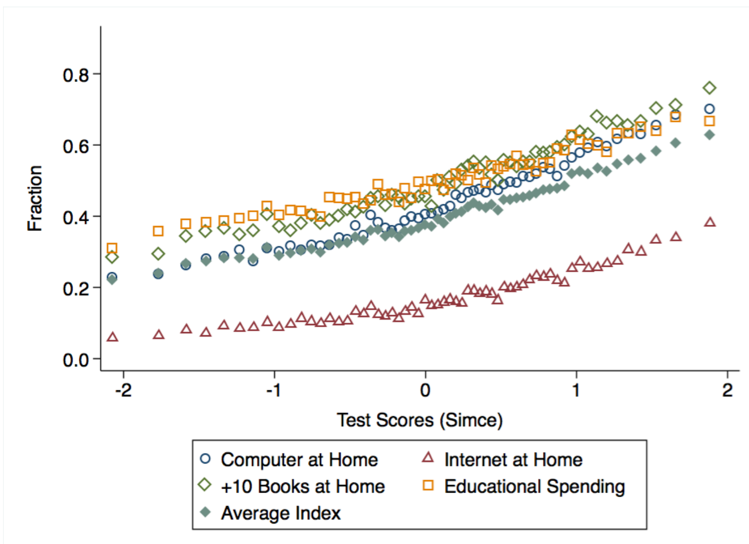
B Appendix: Investments

Financial Investments

We use data for children who were in first grade in years 2002, 2003 and 2004, whose parents were surveyed by SIMCE in years 2005, 2006 and 2007. We use these survey-years because they ask students about the time parents spend with them on educational activities and this survey was not implemented the previous years, and because questions change in the next years. The sample size is of abut 500K observations.

We first we show in Figure B.1 the raw data relating financial investments and college expectations to SIMCE test scores (in standard deviation units). Figure B.1 is a non-parametric plotillustrative of the positive correlation between parental financial investments and children's test scores.

Figure B.1: Financial Investments and Test Scores

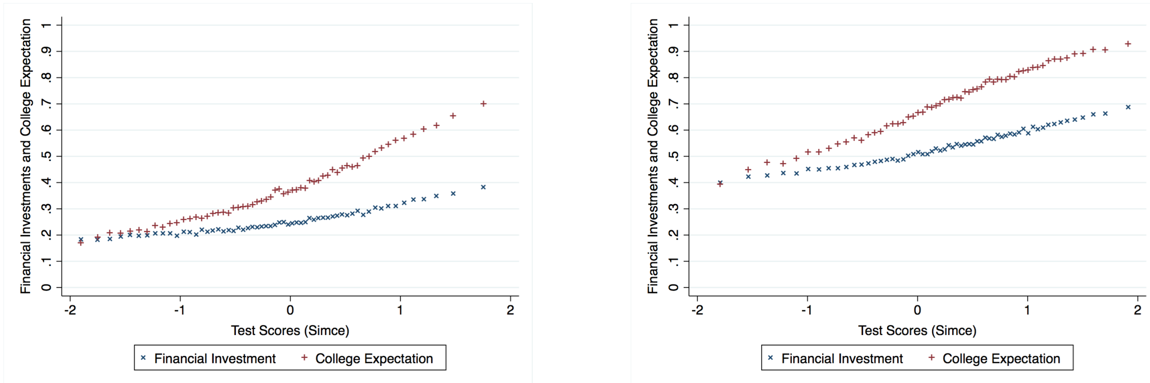


Note: The graphs in Figure B.1 plot the mean of the y-axis variables within equal sized bins of SIMCE test scores (in standard deviation units) in 4th grade, with a sample of approximately 500K observations. The y-axis variables are our measures of parental financial investments (having computer, internet connection and more than 10 books at home, spending above the median in educational items and an the average of those four variables in an Index).

Figure B.2 plots the average index of financial investments and parental beliefs (measured by college expectations), by socioeconomic status. The graphs show a positive correlation by socioeconomic status, showing even a steeper gradient for college expectations in the lower SES group. Overall, graphs in Figure B.1 and Figure B.2 show that in the raw data, parental financial investments and college expectations are correlated with test scores, and that the correlation also exists by socioeconomic group.

Figure B.2: Financial Investments, College Expectations and Test Scores by SES

(i) Low SES **(j)** High SES



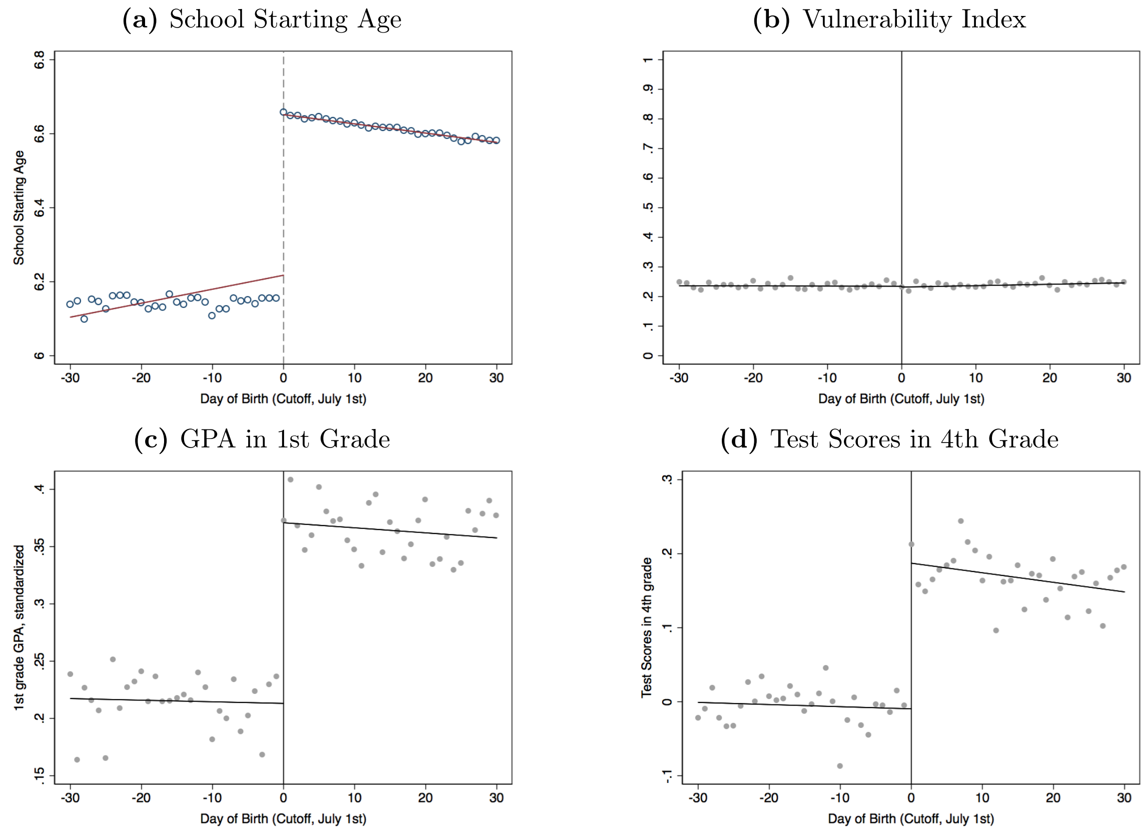
Note: The graphs in Figure B.2 plot the mean of the y-axis variables within equal sized bins of SIMCE test scores (in standard deviation units) in 4th grade, with approximately 250K observations each. The y-axis variables are the average of our measures of parental financial investments (having computer, internet connection and more than 10 books at home, spending above the median in educational items) and whether parents think their child will attend college in the future.

Effects on the Survey Sample

Having data on first graders in years 2002, 2003 and 2004 allows us to exploit two discontinuities, using data for children born in June and July in 1996 and 1997 (as explained in our research design in Figure 3). Therefore we use two thirds of our original sample of 117K, and then given that survey response is about 75% we are left with approximately 50K observations to test effects on financial investments and college expectations.

Figure B.3 and Table B.1 show that results for this sample are similar to our working sample in the main text. There is a jump of about half a year in school starting age near the July 1 threshold, while mothers' schooling is smooth. In terms of outcomes, July-born students have higher GPAs (0.16σ) in first grade and higher test scores (0.18σ) in the fourth grade, very similar to what we found in our working samples. As an additional robustness check we also show results for the vulnerability index, which behaves smoothly near the cutoff as shown in Figure B.3 and Table B.1.

Figure B.3: School Starting Age, In-school Results and Vulnerability Index



Note: The graphs in Figure B.3 plot the mean of the y-axis variable within day of birth, and fit estimated lines using all the underlying data, allowing for different slopes on each side of the cutoff. Each day of birth contains about 2K observations.

Table B.1: Results for the Sample with Financial Investments

	(1) Age at Entry	(2) Vulnerability Index	(3) GPA in 1st Grade	(4) Test Scores in 4th Grade
$\widehat{\alpha_1}$	0.501***	0.000	0.162***	0.178***
	(0.005)	(0.000)	(0.012)	(0.013)
June Mean	6.143	0.286	0.215	-0.005
Effect Size	0.082	0.000	0.753	35.055
Observations	51,818	51,818	51,818	51,818

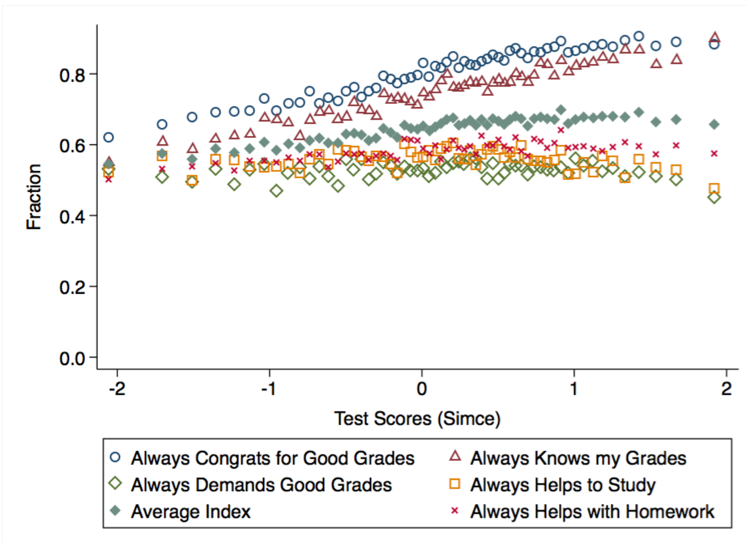
Notes: Table B.1 show the coefficient $\widehat{\alpha_1}$ estimated from the equation (1) for first graders on age at school entry, the vulnerability index and in-school outcomes. These are first grade GPA (GPA1, standardized within school and grade) and the Language-Math score in 4th grade. Robust standard errors (in parentheses) are clustered by day of birth.

Parental Time Investments

We use data for children who were in first grade in years 2008 to 2010, and were surveyed by SIMCE in years 2011 to 2013. We use these surveys because they ask students about the time parents spend with them on educational activities and this survey was not implemented the previous years, and because questions change in the next years. For each cohort we have approximately 233K students and a survey response rate of two-thirds, leaving us with a the dataset of 460K observations

We show in Figure B.4 the raw data relating parental time investments to SIMCE test scores (in standard deviation units). The non-parametric plot in Figure B.4 shows a positive correlation for two measures of parental involvement (whether parents always know and congratulate children for their grades), and a flat, slightly u-shaped correlation for three other measures (whether parents demand good grades, help with homework and help to study).

Figure B.4: Time Investments and Test Scores

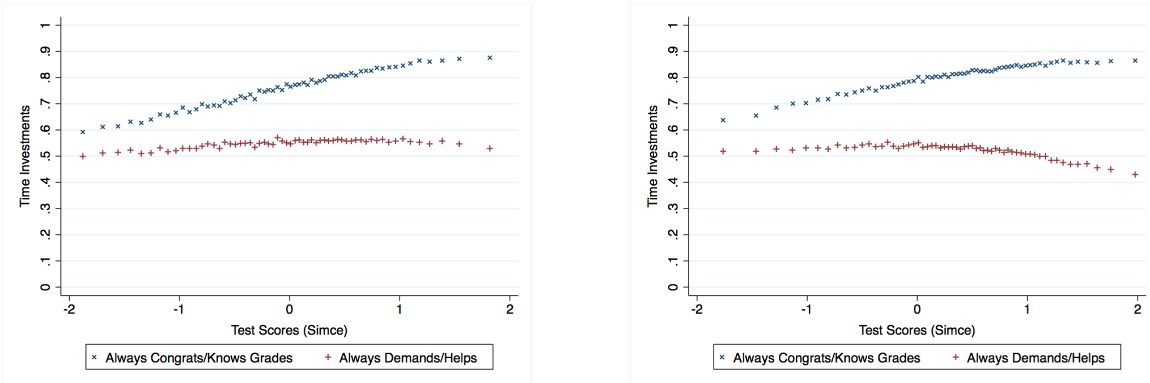


Note: The graphs in Figure B.4 plot the mean of the y-axis variables within equal sized bins of SIMCE test scores (in standard deviation units) in 4th grade, with a sample of approximately 460K observations. The y-axis variables are our measures of parental time investments.

In Figure B.5 we show the same correlations after grouping these two groups of parental investments, by socioeconomic status. The graphs in Figure B.5 display a similar correlation by socioeconomic status. Overall, graphs in Figure B.4 and Figure B.5 show that in the raw data, some parental time investments are correlated with test scores and others not, and that the patterns behaves similarly by socioeconomic group.

Figure B.5: Time Investments and Test Scores by SES

(a) Low SES **(b)** High SES



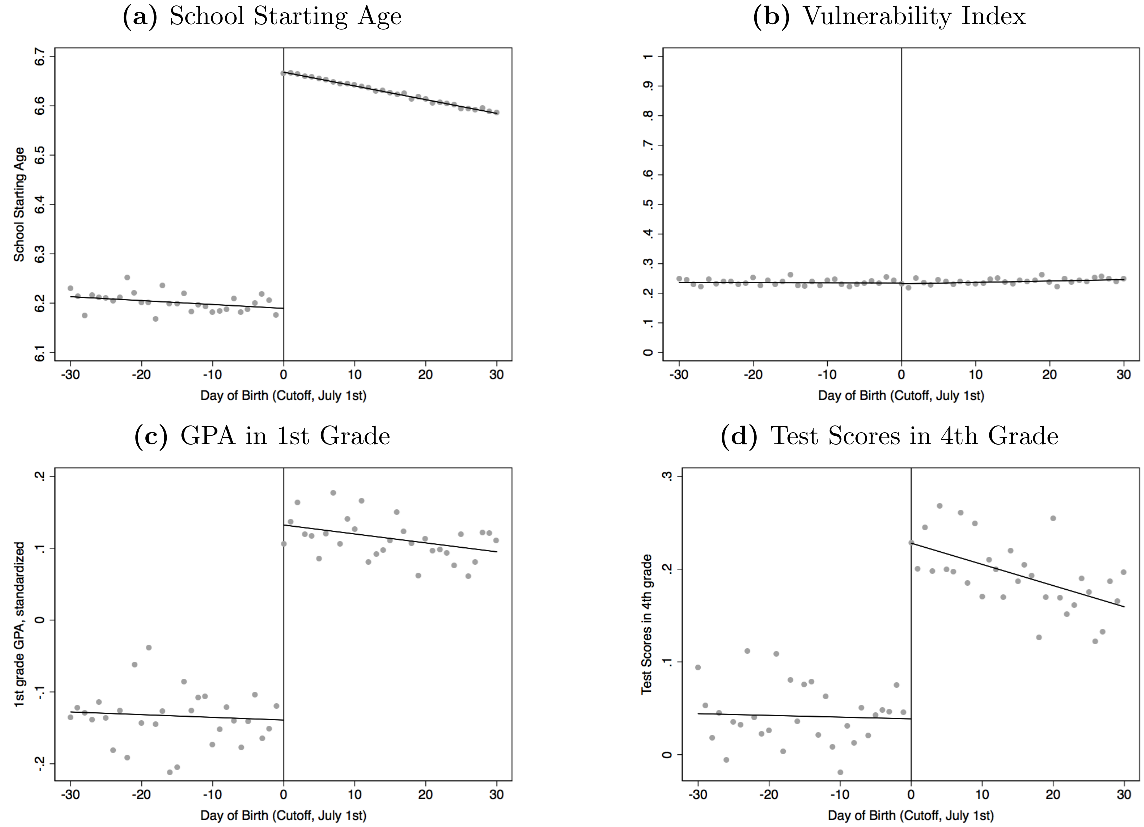
Note: The graphs in Figure B.5 plot the mean of the y-axis variables within equal sized bins of SIMCE test scores (in standard deviation units) in 4th grade, with approximately 250K observations each. The y-axis variables are the average of our measures of parental time investments.

Effects on the Survey Sample

Having data on first graders in years 2008 to 2010 allows us to exploit two discontinuities, using data for children born in June and July in 2002 and 2003 (analogously as explained in our research design in Figure 3). Similarly to our sample for financial investments we are left with approximately 50K observations to test effects on time and teacher investments.

Figure B.6 and Table B.2 show that results for this sample are similar to our working sample in the main text. There is a jump of about half a year in school starting age near the July 1 threshold, while the behavior of the vulnerability index near the cutoff is smooth. In terms of outcomes, July-born students have higher GPAs (0.26σ) in first grade and higher test scores (0.18σ) in the fourth grade, very similar to what we found in our working samples.

Figure B.6: School Starting Age, In-school Results and Vulnerability Index



Note: The graphs in Figure B.6 plot the mean of the y-axis variable within day of birth, and fit estimated lines using all the underlying data, allowing for different slopes on each side of the cutoff. Each day of birth contains about 2K observations.

Table B.2: Results for the Sample with Time Investments

	(1) Age at Entry	(2) Vulnerability Index	(3) GPA in 1st Grade	(4) Test Scores in 4th Grade
$\widehat{\alpha_1}$	0.480***	-0.000	0.263***	0.181***
	(0.006)	(0.000)	(0.014)	(0.013)
June Mean	6.202	0.235	-0.133	0.040
Observations	47,646	47,646	47,646	47,646

Notes: Table B.2 show the coefficient $\widehat{\alpha_1}$ estimated from the equation (1) for first graders on age at school entry, vulnerability index and in-school outcomes. These are first grade GPA (GPA1, standardized within school and grade) and the Language-Math score in 4th grade. Robust standard errors (in parentheses) are clustered by day of birth.

Factor Indexes

In this subsection we provide details of the composite score computation for financial and time investments. We compute a 'factor index' for each type of investment using principal components, which reduces the dimensionality of the investment measures to one composite score. For each index we performed the rotation of the loading matrix using the varimax method to produce the orthogonal factor.

Figure B.7 summarizes the results for each type of investment. The eigenvalues indicate the total variance accounted by each factor. According to the Kaiser criterion we retain the factor with eigenvalues equal or higher than 1. The percentage of variance explained by the factors is 48% and 34% for financial and time investments, respectively.

The factor loadings show the importance of each variable contributing the the factor, while the uniqueness indicates the variance of each variable not shared with other variables in the overall factor model. The higher the loading (and the lower the uniqueness) the more relevant is the variable in defining the factor dimensionality.

Figure B.7: Factor Analyses

(a) Factor Eigenvalues, Financial Investments

Factor analysis/correlation — Number of obs = 51,818
Method: principal-component factors — Retained factors = 1
Rotation: orthogonal varimax (Kaiser off) — Number of params = 4

Factor	Eigenvalue	Difference	Proportion	Cumulative
Factor1	1.91139	1.07443	0.4778	0.4778
Factor2	0.83696	0.06580	0.2092	0.6871
Factor3	0.77117	0.29069	0.1928	0.8799
Factor4	0.48048	.	0.1201	1.0000

LR test: independent vs. saturated: chi2(6) = 2.7e+04 Prob>chi2 = 0.0000

(b) Factor Loadings, Financial Investments

Factor loadings (pattern matrix) and unique variances

Variable	Factor1	Uniqueness
Computer	0.8052	0.3517
Internet	0.7530	0.4331
High Spending	0.5482	0.6995
More 10 books	0.6290	0.6043

(c) Factor Eigenvalues, Time Investments

Factor analysis/correlation — Number of obs = 47,646
Method: principal-component factors — Retained factors = 1
Rotation: orthogonal varimax (Kaiser off) — Number of params = 5

Factor	Eigenvalue	Difference	Proportion	Cumulative
Factor1	1.71764	0.77870	0.3435	0.3435
Factor2	0.93894	0.00129	0.1878	0.5313
Factor3	0.93766	0.19941	0.1875	0.7188
Factor4	0.73824	0.07073	0.1476	0.8665
Factor5	0.66752	.	0.1335	1.0000

LR test: independent vs. saturated: chi2(10) = 1.4e+04 Prob>chi2 = 0.0000

(d) Factor Loadings, Time Investments

Factor loadings (pattern matrix) and unique variances

Variable	Factor1	Uniqueness
Congrats Gr.	0.6357	0.5959
Knows Grades	0.5467	0.7012
Demands Gr.	0.3687	0.8641
Helps Study	0.6564	0.5691
Helps Homewo	0.6693	0.5521

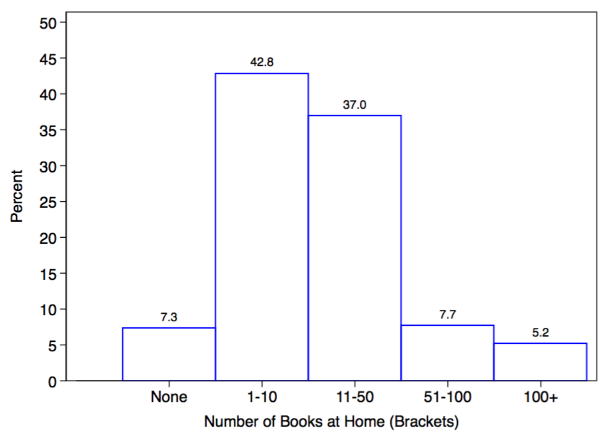
Robustness to the grouping of answers

Number of Books

Parents provide information on the number of books they own by brackets. Figure B.8 shows the distribution of their answers by bracket. A 7% of parents report having no books, 43% report 1-10, 37% say 11-50, and then 7.7% report to have 51-100 books, with a 5.2% reporting more than 100 books. About half of parents report to have more than 10 books, which is the variable we use in the main text as a proxy of having books.

Table B.3 shows that our results are robust to a number of ways of defining the *books* variable. Column (1) shows results on the number of books imputed as the middle value within each category (we used 120 books for the last category). Column (2) presents the results on the variable as it comes in the data (ranging from 0 to 4). Columns (3), (4) and (5) show the results on a dummy for more than 0, 10 and 50 books. The only variable where there are no effects is the 50 books category, but 90% of parents report to have less than 50 books. Other than than, all other variables show an effect on having books.

Figure B.8: Number of Books at Home



Note: The graphs in Figure B.8 shows the distribution of the categorical variable 'Number of Books at Home', reported by parents in SIMCE surveys. The sample size is N=51,818.

Table B.3: Results for Books Grouping

	(1) Imp. Number of Books	(2) Bracket of Books	(3) More than 0 Books	(4) More than 10 Books	(5) More than 50 Books
$\widehat{\alpha_1}$	1.298**	0.042***	0.013***	0.037***	-0.006
	(0.510)	(0.014)	(0.004)	(0.008)	(0.005)
June Mean	36.248	1.588	0.917	0.478	0.137
Effect Size	0.036	0.027	0.014	0.078	-0.041
Observations	51,818	51,818	51,818	51,818	51,818

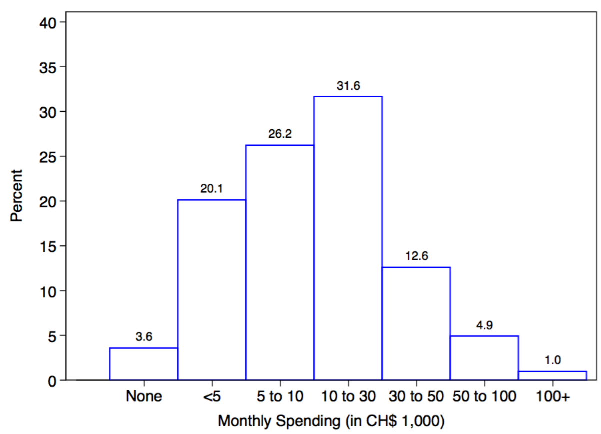
Notes: Table B.3 show the coefficient $\widehat{\alpha_1}$ estimated from the equation (1) for first graders on different ways of presenting the 'Books at Home' variable. Robust standard errors (in parentheses) are clustered by day of birth.

Spending Categories

Parents provide information on their monthly spending on educational inputs, by brackets. Figure B.9 shows the distribution of their answers by bracket. A 4% of parents report no spending, 20% report less than 5K Chilean pesos ($CH), 26% say 5-10, 32% report 10-30, 13% report 30-50, and then 5% report to spend 50-100, with a 1% reporting more than 100$CH. About half of parents report to spend than 10 $CH, which is the variable we use in the main text as a proxy for spending.

Table B.4 shows that our results are robust to a number of ways of defining the *spending* variable. Column (1) shows results on spending with values imputed as the middle value within each category (we used 120 $CH for the last category). Column (2) presents the results on the variable as it comes in the data (ranging from 0 to 6). Columns (3) to (6) show the results on a dummy for more than 5K, 10K, 30K and 50K of $CH. With the exception of the first dummy, all other variables show an effect on spending, with sizable effect sizes.

Figure B.9: Spending Categories



Note: The graphs in Figure B.9 shows the distribution of the categorical variable 'Monthly Spending on Educational Inputs', reported by parents in SIMCE surveys. The sample size is N=51,818.

Table B.4: Results for Spending Grouping

	(1) Imp. Amount of Spending	(2) Bracket of Spending	(3) More 5K Sp	(4) More 10K Sp	(5) More 30K Sp	(6) More 50K Sp
$\widehat{\alpha_1}$	723.596**	0.049**	0.009	0.022***	0.013**	0.006*
	(294.659)	(0.019)	(0.007)	(0.007)	(0.006)	(0.003)
June Mean	15793.682	2.424	0.757	0.478	0.165	0.052
Effect Size	0.046	0.020	0.011	0.047	0.077	0.111
Observations	51,818	51,818	51,818	51,818	51,818	51,818

Notes: Table B.4 show the coefficient $\widehat{\alpha_1}$ estimated from the equation (1) for first graders on different ways of presenting the 'Monthly Spending on Educational Inputs' variable. Robust standard errors (in parentheses) are clustered by day of birth.

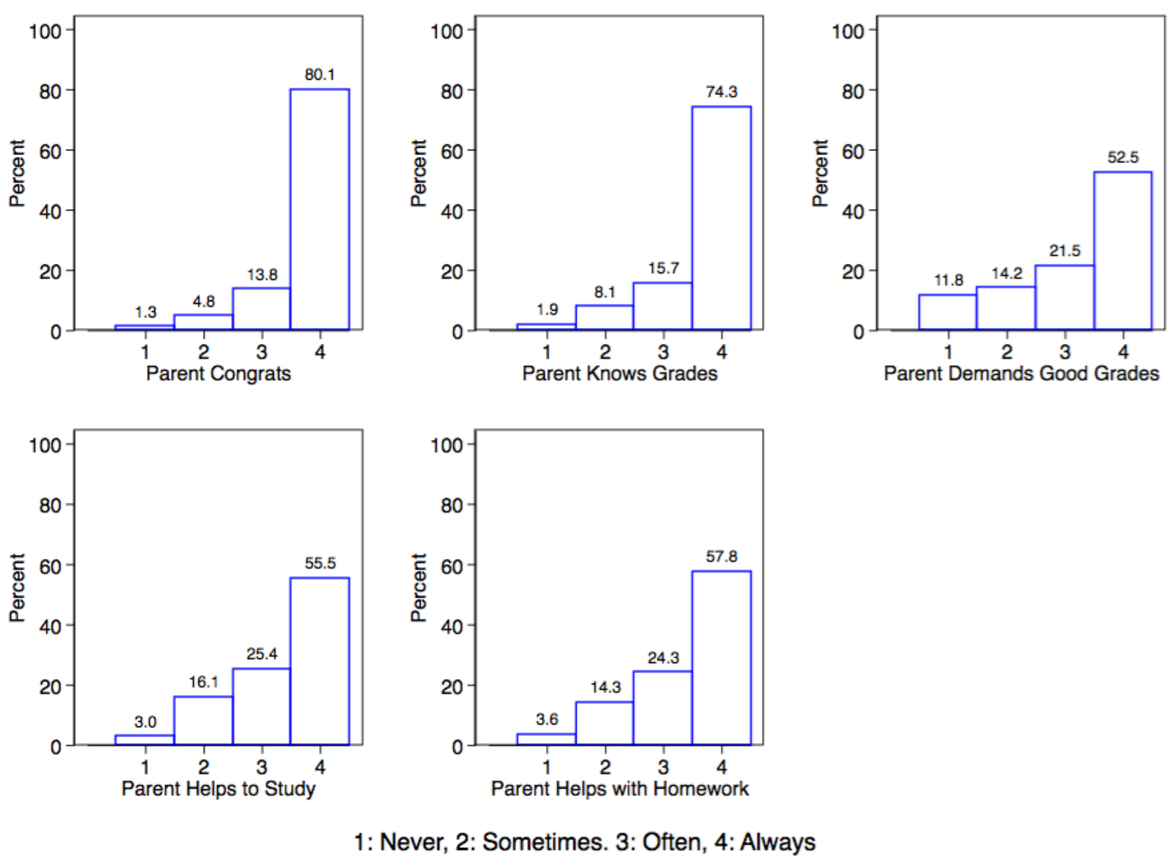
Time Investment Categories

Students are asked about the time spent with their parents on educational activities. In particular, children report in a 1-4 Likert scale whether their parents help them study or with their homework, help them understand difficult subjects, whether parents know their grades, and whether parents demand improving grades. Available answers for each item are on the scale of 'Never', 'Sometimes', 'Most of the time', and 'Always'.

Figure B.10 shows the distribution of their answers by question. Most of the students answer always for each question, ranging from 80% to 56%. This is why generate variables that equal one if the child answered that her parent does each activity 'Always', and zero otherwise.

Table B.5 shows that our results are robust to the way we use students answers. The first Panel shows the results for the answers in a 1-4 range, the second Panel shows the results grouping variables in the categories Often and Always, and the last panel shows results as in the main text. Effect sizes in almost all cases never higher than 2% and similar across tables.

Figure B.10: Time Investment Categories



Note: The graphs in Figure B.10 shows the distribution of the categorical variables of time investment reported by parents in SIMCE surveys. The sample size is N=47,646.

Table B.5: Results for Time Investment Grouping

	(1) Congrats for good grades (from 1-4)	(2) Knows my grades (from 1-4)	(3) Demands good grades (from 1-4)	(4) Helps to study (from 1-4)	(5) Helps with homework (from 1-4)
$\widehat{\alpha_1}$	0.001	0.036**	0.028	0.010	0.078***
	(0.010)	(0.015)	(0.025)	(0.015)	(0.017)
June Mean	3.735	3.628	3.139	3.339	3.341
Effect Size	0.000	0.010	0.009	0.003	0.023
Observations	47,646	47,646	47,646	47,646	47,646

	(1) Congrats for good grades (Often & Always)	(2) Knows my grades (Often & Always)	(3) Demands good grades (Often & Always)	(4) Helps to study (Often & Always)	(5) Helps with homework (Often & Always)
$\widehat{\alpha_1}$	0.010**	0.021***	0.012	0.005	0.046***
	(0.004)	(0.006)	(0.008)	(0.006)	(0.009)
June Mean	0.938	0.896	0.734	0.811	0.808
Effect Size	0.010	0.024	0.017	0.006	0.057
Observations	47,646	47,646	47,646	47,646	47,646

	(1) Congrats for good grades (Always)	(2) Knows my grades (Always)	(3) Demands good grades (Always)	(4) Helps to study (Always)	(5) Helps with homework (Always)
$\widehat{\alpha_1}$	-0.010	0.010	0.001	0.001	0.018**
	(0.006)	(0.010)	(0.012)	(0.009)	(0.009)
June Mean	0.810	0.752	0.526	0.559	0.569
Effect Size	-0.012	0.013	0.002	0.001	0.032
Observations	47,646	47,646	47,646	47,646	47,646

Notes: Table B.5 show the coefficient $\widehat{\alpha_1}$ estimated from the equation (1) for first graders on different ways of presenting the ‘Monthly Spending on Educational Inputs’ variable. Robust standard errors (in parentheses) are clustered by day of birth.

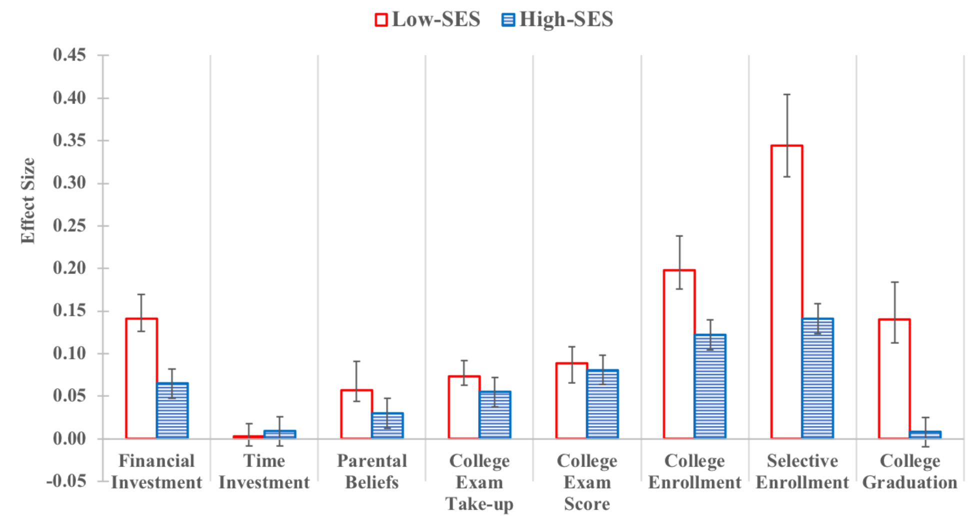
C Appendix: Additional Results

Table C.1: Effect Sizes, by Quartiles of Socioeconomic Status

	(1) Low SES	(2) Med-Low SES	(3) Med-High SES	(4) High SES	(5) Low vs High Difference
Financial Investments					
Average Index	0.189***	0.113***	0.071***	0.062***	0.127***
	(0.0372)	(0.0330)	(0.0274)	(0.0155)	(0.040)
Computer at Home	0.305***	0.157***	0.123***	0.033*	0.272***
	(0.0882)	(0.0573)	(0.0442)	(0.0176)	(0.090)
Internet at Home	0.597**	0.454***	0.131	0.155***	0.442*
	(0.2472)	(0.1480)	(0.0920)	(0.0407)	(0.251)
+10 Books	0.143***	0.097*	0.043	0.066***	0.077*
	(0.0405)	(0.0506)	(0.0352)	(0.0222)	(0.046)
High Spending	0.128**	0.034	0.036	0.039*	0.089
	(0.0595)	(0.0374)	(0.0274)	(0.0218)	(0.063)
Time Investments					
Average Index	0.004	0.002	0.012	0.025*	-0.021
	(0.0194)	(0.0157)	(0.0151)	(0.0144)	(0.024)
Congrats Grades	0.047***	0.006	0.026*	0.013	0.034
	(0.0170)	(0.0194)	(0.0133)	(0.0150)	(0.023)
Knows Grades	0.019	0.019	0.026	0.022	-0.003
	(0.0244)	(0.0221)	(0.0175)	(0.0176)	(0.030)
Demands Good Grades	0.087**	0.048	0.025	0.008	0.079
	(0.0432)	(0.0333)	(0.0347)	(0.0408)	(0.059)
Helps to Study	0.007	0.001	0.045	0.048	-0.041
	(0.0401)	(0.0299)	(0.0395)	(0.0434)	(0.059)
Helps with Homework	0.042	0.025	0.001	0.056**	-0.014
	(0.0372)	(0.0247)	(0.0316)	(0.0253)	(0.045)
Parental Beliefs					
College Expectation	0.073	0.048	0.038*	0.027*	0.046
	(0.0663)	(0.0393)	(0.0217)	(0.0144)	(0.068)
Grad School Expectation	0.429**	0.167	0.165*	0.038	0.391*
	(0.2086)	(0.1174)	(0.0957)	(0.0762)	(0.222)
Institute Expectation	0.043	0.057**	0.005	0.009	0.034
	(0.0440)	(0.0265)	(0.0164)	(0.0116)	(0.046)
High School Expectation	0.039	0.042*	0.004	0.007	0.032
	(0.0404)	(0.0219)	(0.0142)	(0.0098)	(0.042)
Long Run Outcomes					
Takes PSU Exam	0.094***	0.057**	0.072***	0.042***	0.052
	(0.0307)	(0.0273)	(0.0174)	(0.0100)	(0.032)
PSU Exam Scores	0.080***	0.101***	0.082**	0.078**	0.002
	(0.0286)	(0.0256)	(0.0333)	(0.0322)	(0.043)
College Enrollment	0.273***	0.148***	0.151***	0.106***	0.167**
	(0.0680)	(0.0525)	(0.0403)	(0.0255)	(0.073)
Selective College Enrollment	0.374***	0.325***	0.159***	0.129***	0.245**
	(0.1001)	(0.0713)	(0.0612)	(0.0419)	(0.109)
College Graduation	0.236***	0.069	0.072	0.027	0.209**
	(0.0792)	(0.0563)	(0.0522)	(0.0335)	(0.086)

Notes: Table C.1 shows the effect sizes on parental investments, beliefs, and children long run outcomes by quartiles of the national vulnerability index. We label each quartile as low SES, med-low SES, med-high SES, and high SES, in columns (1)-(4). Column (5) show the difference between the low and high SES effect size. Standard errors for the effect sizes are computed using the delta method.

Figure C.1: Effect Sizes by Low-High Socioeconomic Status



Note: Figure C.1 plots the effect sizes on parental investments, beliefs and children long run outcomes. Financial and time investments are each measured by the average index described in subsection 3.2. Parental beliefs are measured as college completion expectations. Long run outcomes consist on the college entrance exam take-up and scores, college enrollment (overall and at selective institutions), and college graduation.

Table C.2: Effect Sizes, by Low-High Socioeconomic Status

Variable	(1) Low SES	(2) High SES	(3) Difference
Financial Investments			
Average Index	0.141***	0.065***	0.076**
	(0.0285)	(0.0151)	(0.0323)
Computer at Home	0.207***	0.068***	0.139**
	(0.0512)	(0.0214)	(0.0555)
Internet at Home	0.482***	0.148***	0.334**
	(0.1336)	(0.0429)	(0.1403)
+10 Books	0.116***	0.056***	0.060
	(0.0363)	(0.0150)	(0.0393)
High Spending	0.070**	0.036**	0.034
	(0.0308)	(0.0172)	(0.0353)
Time Investments			
Average Index	0.003	0.009	-0.006
	(0.0144)	(0.0110)	(0.0181)
Congrats Grades	0.022*	0.003	0.019
	(0.0126)	(0.0099)	(0.0160)
Knows Grades	0.001	0.024*	-0.023
	(0.0185)	(0.0141)	(0.0233)
Demands Good Grades	0.018	0.016	0.002
	(0.0280)	(0.0303)	(0.0413)
Helps to Study	0.004	0.005	-0.001
	(0.0275)	(0.0271)	(0.0386)
Helps with Homework	0.033	0.033*	0.000
	(0.0213)	(0.0191)	(0.0286)
Parental Beliefs			
Grad School Expectation	0.267***	0.087	0.180
	(0.0997)	(0.0616)	(0.1172)
College Expectation	0.057*	0.030**	0.027
	(0.0337)	(0.0128)	(0.0360)
Institute Expectation	0.051**	0.002	0.049*
	(0.0249)	(0.0086)	(0.0263)
Long Run Outcomes			
Takes PSU Exam	0.073***	0.055***	0.018
	(0.0187)	(0.0101)	(0.0213)
College Enrollment	0.198***	0.122***	0.076*
	(0.0398)	(0.0225)	(0.0457)
Selective College Enrollment	0.344***	0.141***	0.203***
	(0.0605)	(0.0365)	(0.0707)
College Graduation	0.140***	0.008	0.132**
	(0.0442)	(0.0272)	(0.0519)

Notes: Table C.2 shows the effect sizes on parental investments, beliefs and children long run outcomes, by low/high socioeconomic status (columns 1 and 2 respectively). Column 3 shows the difference between the low and high SES effect size. Standard errors for the effect sizes are computed using the delta method.