Online Appendix for Early Skill Effects on Parental Beliefs, Investments and Children Long-Run Outcomes

Pablo Celhay[a] and Sebastian Gallegos[b]

[a]Pontificia Universidad Católica de Chile
[b]UAI Business School and HCEO-UChicago

June 2, 2022

Appendices

A Appendix: Robustness

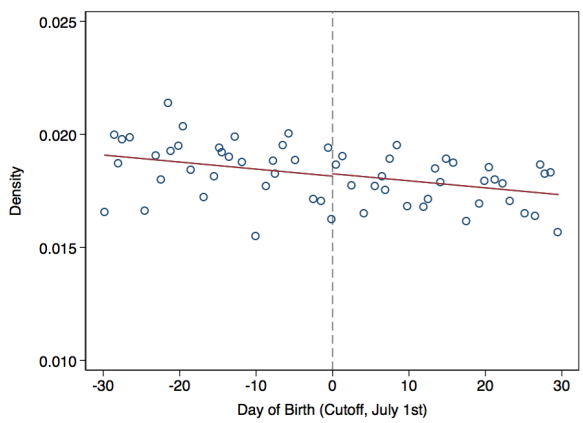
A.1 Density of Running Variable

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 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

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. Our research design mimics a local experiment where children are exogenously (to potential outcomes) allocated to either being born in June or July.

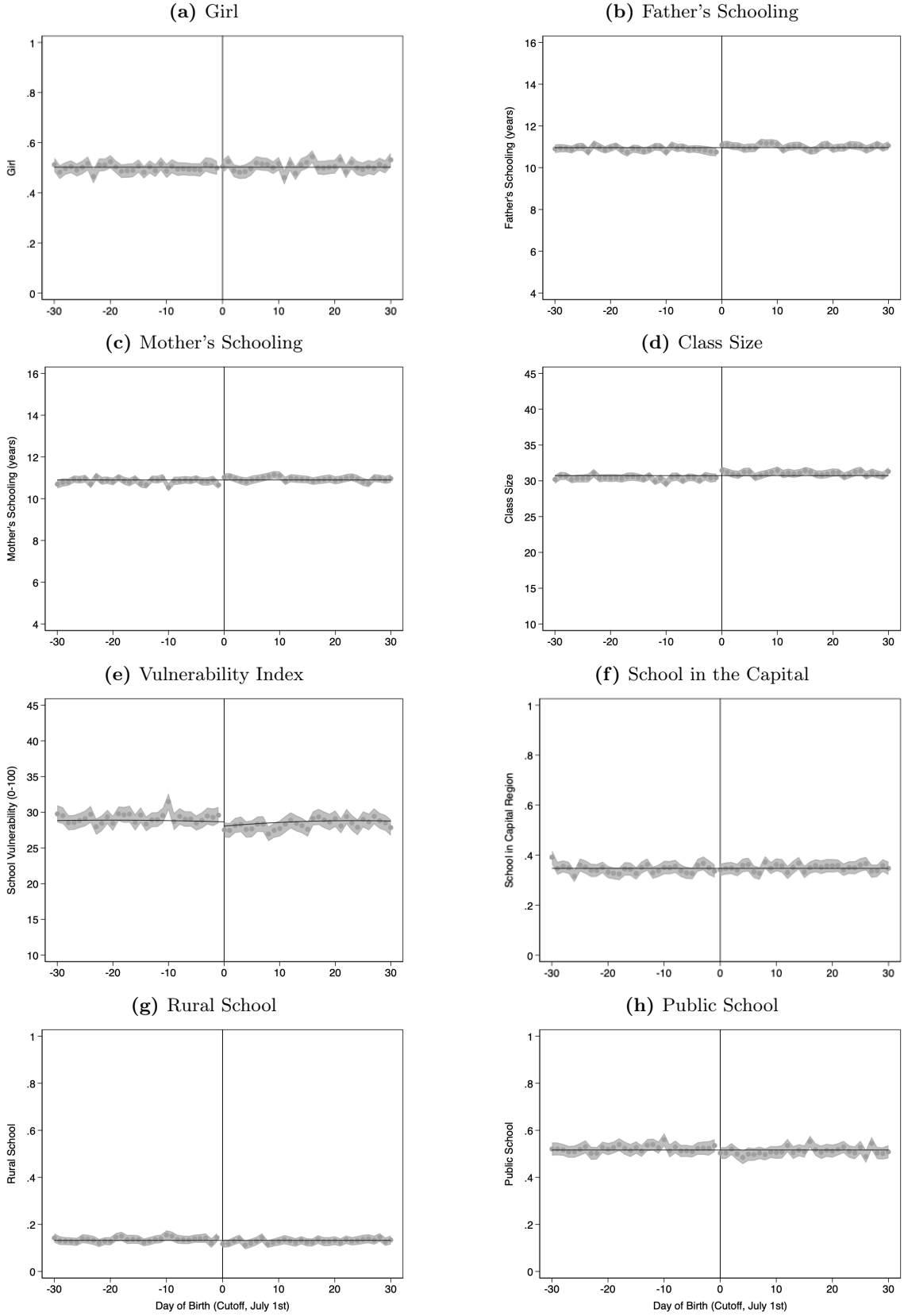
Accordingly, we find precise zeros when estimating equation (1) using each covariate in Table 1 as dependent variable. The effect sizes reported in Table A.2 are never higher than 0.02. and are precisely estimated (small standard errors). Some of these tiny estimates (e.g., class size and probability of going to a public school) are statistically significant at conventional levels, because we are highly powered to detect them (regressions in Table A.2. have almost 118K observations) and not because of their magnitude. We complement these results with a graphical illustration for every covariate in Figure A.2 and Figure A.3 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) Vulnerability 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.

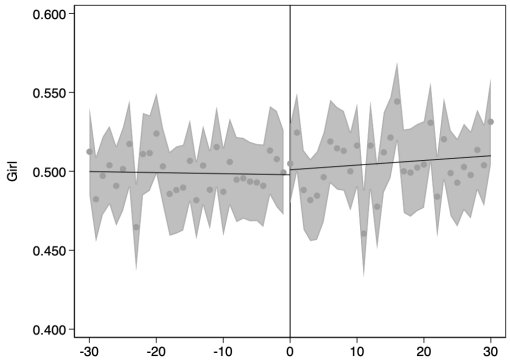
Figure A.2: Covariates Smoothness with Confidence Intervals



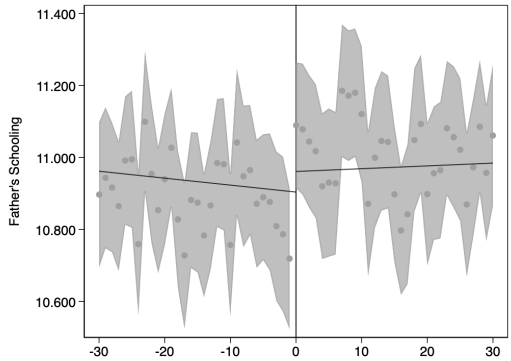
Note: The graphs in Figure A.2 plot the mean of the y-axis variable within day of birth, and fit estimated lines using all the underlying data, allowing for different slopes on each side of the cutoff. Each day of birth contains about 2K observations.

Figure A.3: Covariates Smoothness with Confidence Intervals and Re-scaled Y-axes

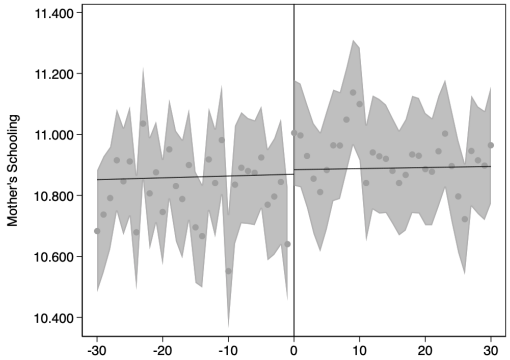
(a) Girl: $\widehat{\alpha_1} = 0.002$ (0.006)



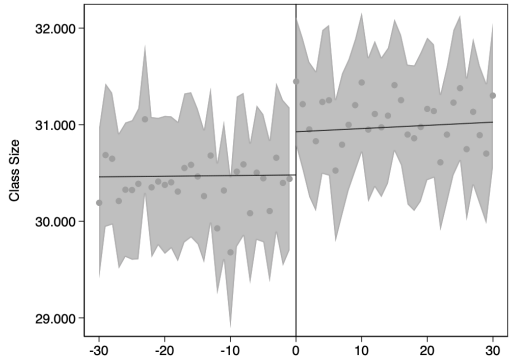
(b) Father's Schooling: $\widehat{\alpha_1} = 0.052$ (0.041)



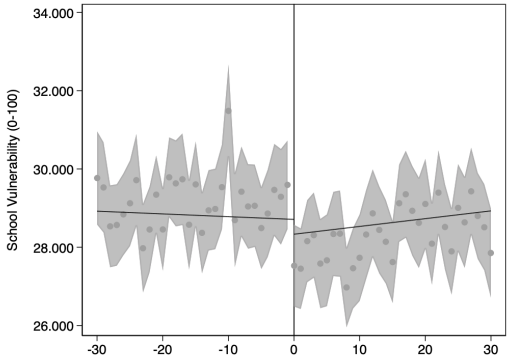
(c) Mother's Schooling: $\widehat{\alpha_1} = 0.009$ (0.041)



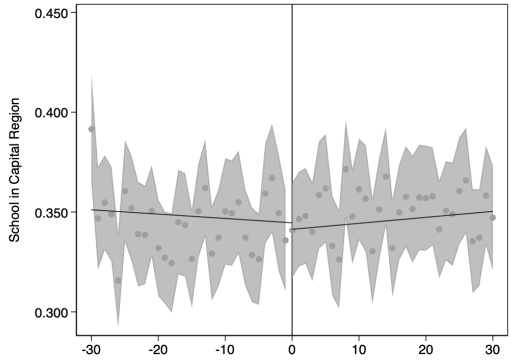
(d) Class Size: $\widehat{\alpha_1} = 0.531$ (0.102)



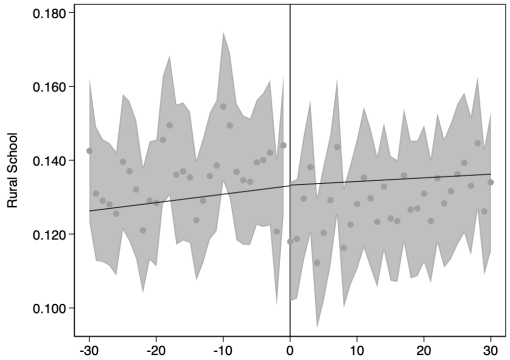
(e) Vulnerability Index: $\widehat{\alpha_1} = -0.003$ (0.002)



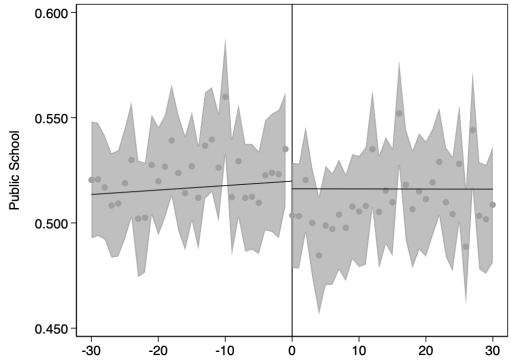
(f) School in the Capital: $\widehat{\alpha_1} = -0.005$ (0.005)



(g) Rural School: $\widehat{\alpha_1} = 0.003$ (0.003)



(h) Public School: $\widehat{\alpha_1} = -0.011$ (0.004)



Note: The graphs in Figure A.3 replicate those in Figure A.2 with re-scaled y-axes.

A.3 Robustness to Including Different Sets of Covariates

As yet another robustness check, Table A.3 reports the coefficients $\widehat{\alpha_1}$ estimated from the equation (1) on long run outcomes using different set of controls.

In Panel (A) we include cohort fixed effects, weekends,and holidays; in Panel (B) we include the controls in (A) and add a set of demographics (female, class size, rural school, school in capital region); in Panel (C) we add the school vulnerability Index; in Panel (D) we add public school attendance

Our main estimates remain practically unchanged across panels when we add controls, and the same compared with Table 5 in the main text.

Table A.3: Robustness

	(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
Panel (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)
June Mean	0.912	0.684	0.580	-0.067	0.253	0.138	0.069	0.165
Observations	117,709	117,709	117,709	71,509	117,709	117,709	117,709	111,664
Panel (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)
June Mean	0.912	0.684	0.580	-0.067	0.253	0.138	0.069	0.165
Observations	117,709	117,709	117,709	71,509	117,709	117,709	117,709	111,664
Panel (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)
June Mean	0.912	0.684	0.580	-0.067	0.253	0.138	0.069	0.165
Observations	117,709	117,709	117,709	71,509	117,709	117,709	117,709	111,664
Panel (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)
June Mean	0.912	0.684	0.580	-0.067	0.253	0.138	0.069	0.165
Observations	117,709	117,709	117,709	71,509	117,709	117,709	117,709	111,664

Notes: Table A.3 reports the coefficients $\widehat{\alpha_1}$ estimated from the equation (1) on long run outcomes using different set of controls. In Panel (A) we include cohort fixed effects, weekends,and holidays; in Panel (B) we add set of demographics (female, class size, rural school, school in capital region); in Panel (C) we add the school vulnerability Index; in Panel (D) we add public school attendance. Robust standard errors (in parentheses) are clustered by day of birth.

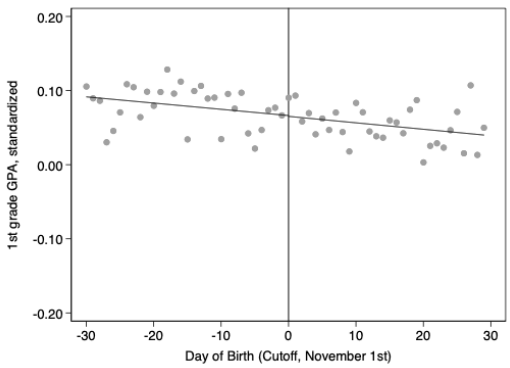
A.4 Placebo Tests

We implemented a host of placebo tests following your suggestion and confirm that there is no gap on in-school outcomes, parental investments and long term outcomes. We use students born in October and November and estimate the coefficient at the arbitrary November 1st threshold.

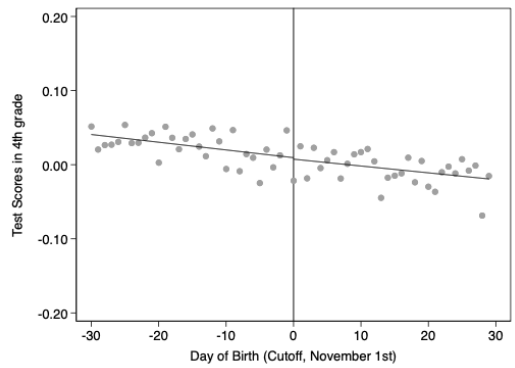
We find precise zero effects on GPA in 1st grade and test scores in fourth grade, as shown in Figure A.4 and Table A.4.

Figure A.4: Placebo: Figures for In-school Effects

(a) GPA in 1st Grade



(b) Test Scores in 4th Grade



Note: The graphs in Figure A.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 A.4: Placebo: Table with In-school Effects

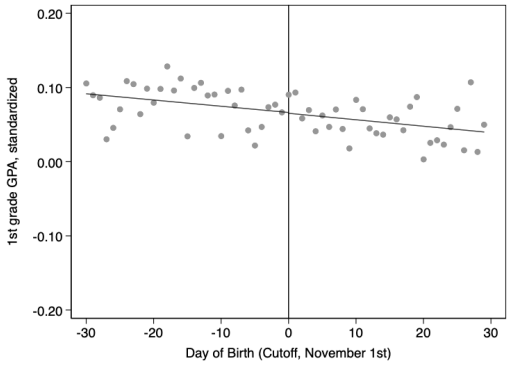
	(1) GPA in 1st Grade	(2) Test Scores in 4th Grade
$\widehat{\alpha_1}$	-0.003	-0.002
	(0.012)	(0.009)
October Mean	0.057	0.009
Observations	121,331	121,331

Notes: Table A.4 shows the coefficient $\widehat{\alpha_1}$ estimated from the equation (1) for first graders on age at school entry, and in-school outcomes. The 'October 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.

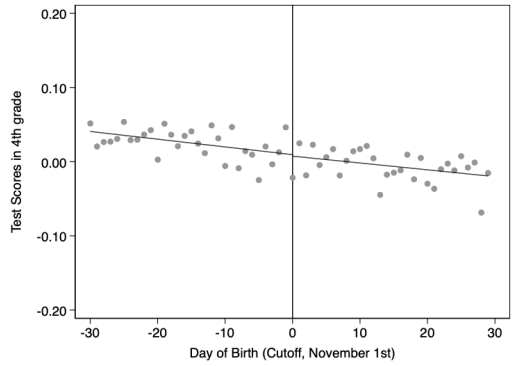
We also find precise zero effects on parental investments and long term outcomes. Figure A.5 shows results for all outcomes, while Table A.5 and Table A.6 show the coefficient $\widehat{\alpha_1}$ estimated from the equation (1) on parental investments and long run outcomes.

Figure A.5: Placebo Figures

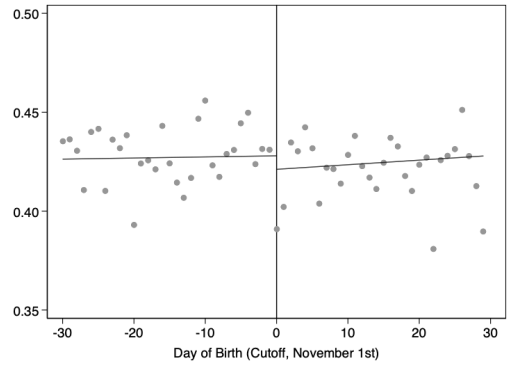
(a) GPA in 1st Grade



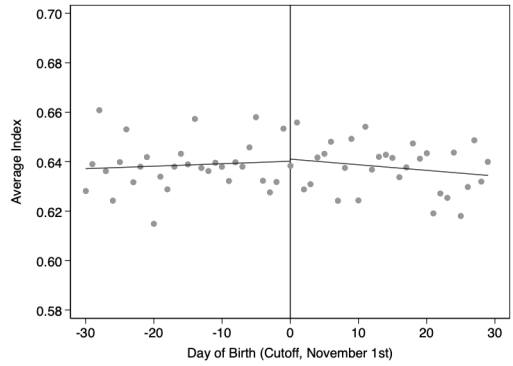
(b) Test Scores in 4th Grade



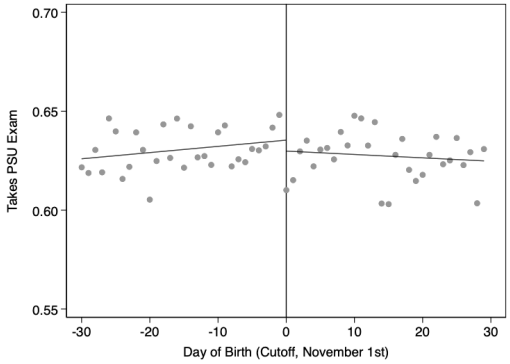
(c) Financial Investments



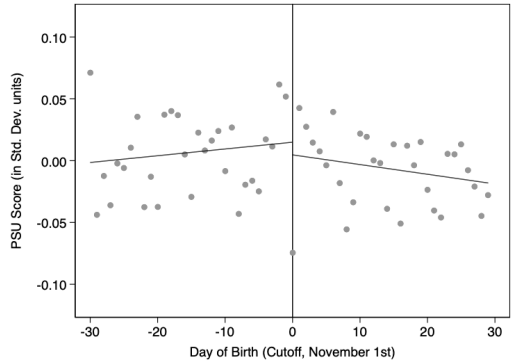
(d) Time Investments



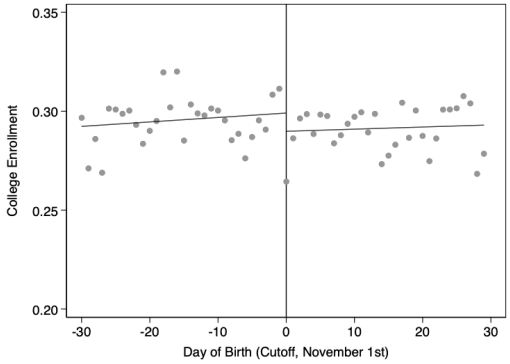
(e) Taking College Entrance Exam



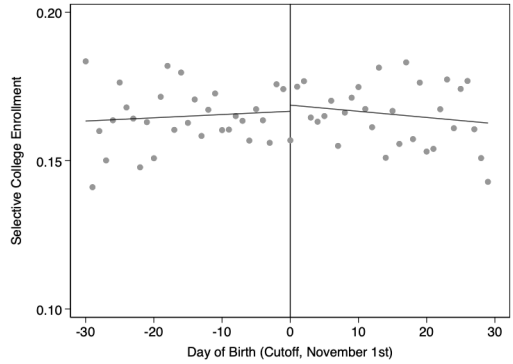
(f) Scores in the College Entrance Exam



(g) Overall College Enrollment



(h) Selective College Enrollment



Note: The graphs in Figure A.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. Each day of birth contains about 2K observations.

Table A.5: Placebo: Effects on Parental Investments

Panel A: Financial Investments

November 1 threshold

	(1) Average Index	(2) Factor Index	(3) Computer at Home	(4) Internet at Home	(5) More 10 Books	(6) High Spending
$\widehat{\alpha_1}$	-0.007	-0.024	-0.017	-0.006	-0.006	-0.000
	(0.005)	(0.017)	(0.008)	(0.006)	(0.009)	(0.008)
October Mean	0.431	0.033	0.476	0.200	0.523	0.527
Effect Size	-0.022	-0.024	-0.036	-0.028	-0.011	-0.000
Observations	52,775	52,775	52,775	52,775	52,775	52,775

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

November 1 threshold

	(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
$\widehat{\alpha_1}$	0.001	0.003	0.002	0.009	-0.002	0.007	-0.013
	(0.005)	(0.019)	(0.008)	(0.009)	(0.010)	(0.009)	(0.008)
October Mean	0.640	0.007	0.795	0.730	0.519	0.561	0.596
Effect Size	0.003	0.003	0.002	0.013	-0.003	0.013	-0.021
Observations	49,611	49,611	49,611	49,611	49,611	49,611	49,611

Notes: Table A.5 shows the coefficient $\widehat{\alpha_1}$ estimated from the equation (1) on parental investments. The 'October Mean' is the mean of the dependent variable just below the threshold. The outcomes are defined in the main text. Robust standard errors (in parentheses) are clustered by day of birth.

Table A.6: Placebo Effects on Long Run Outcomes

November 1 threshold

	(1) Primary Grad	(2) High-School Grad	(3) PSU Exam	(4) PSU Score	(5) College Enroll	(6) Selective Enroll	(7) STEM Enroll
$\widehat{\alpha_1}$	0.001	-0.001	-0.006	-0.009	-0.009	0.002	-0.001
	(0.003)	(0.005)	(0.005)	(0.015)	(0.005)	(0.004)	(0.003)
October Mean	0.930	0.711	0.636	-0.104	0.298	0.167	0.081
Effect Size	0.005	0.001	0.011	0.010	0.020	0.006	0.003
Observations	119,006	119,006	119,006	74,849	119,006	119,006	119,006

Notes: Table A.6 shows the coefficient $\widehat{\alpha_1}$ estimated from the equation (1) on long run outcomes. The 'October Mean' is the mean of the dependent variable just below the threshold. The outcomes are defined in the main text. Robust standard errors (in parentheses) are clustered by day of birth.

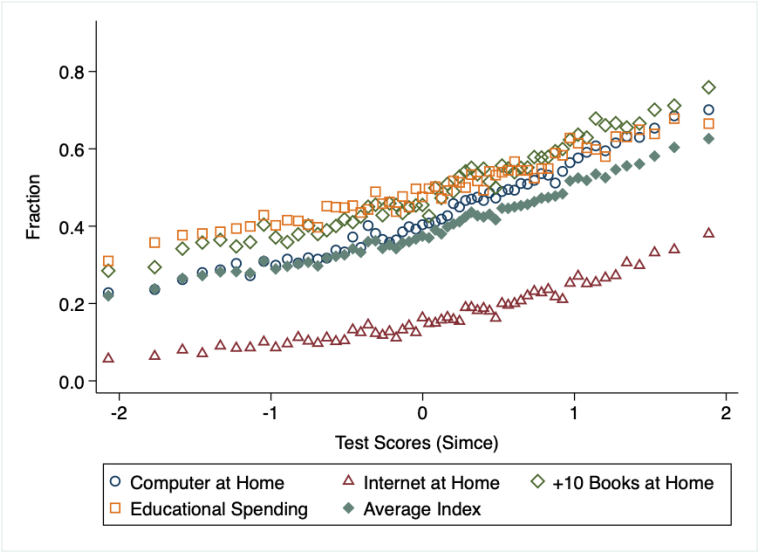
B Appendix: Investments

B.1 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 plot illustrative of the positive correlation between parental financial investments and children's test scores.

Figure B.1: Financial Investments and Test Scores

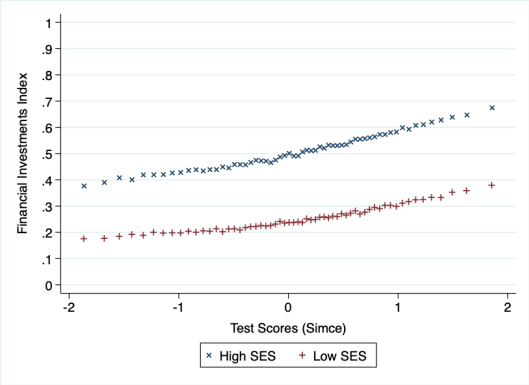


Note: The graphs in Figure B.1 plot the mean of the y-axis variables within equal sized bins of SIMCE test scores (in standard deviation units) in 4th grade, with a sample of approximately 500K observations. The y-axis variables are our measures of parental financial investments (having computer, internet 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).

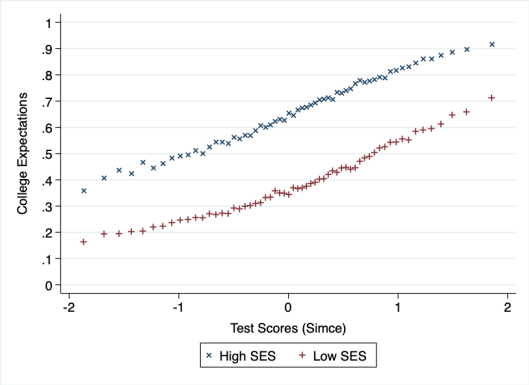
Figure B.2 plots the average index of financial investments and parental beliefs (measured by college expectations), by socioeconomic status. The graphs shows that the positive correlation holds for both high and low socioeconomic status, with a gap that remains constant along the test score distribution. 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

(a) Financial Investments



(b) College Expectations



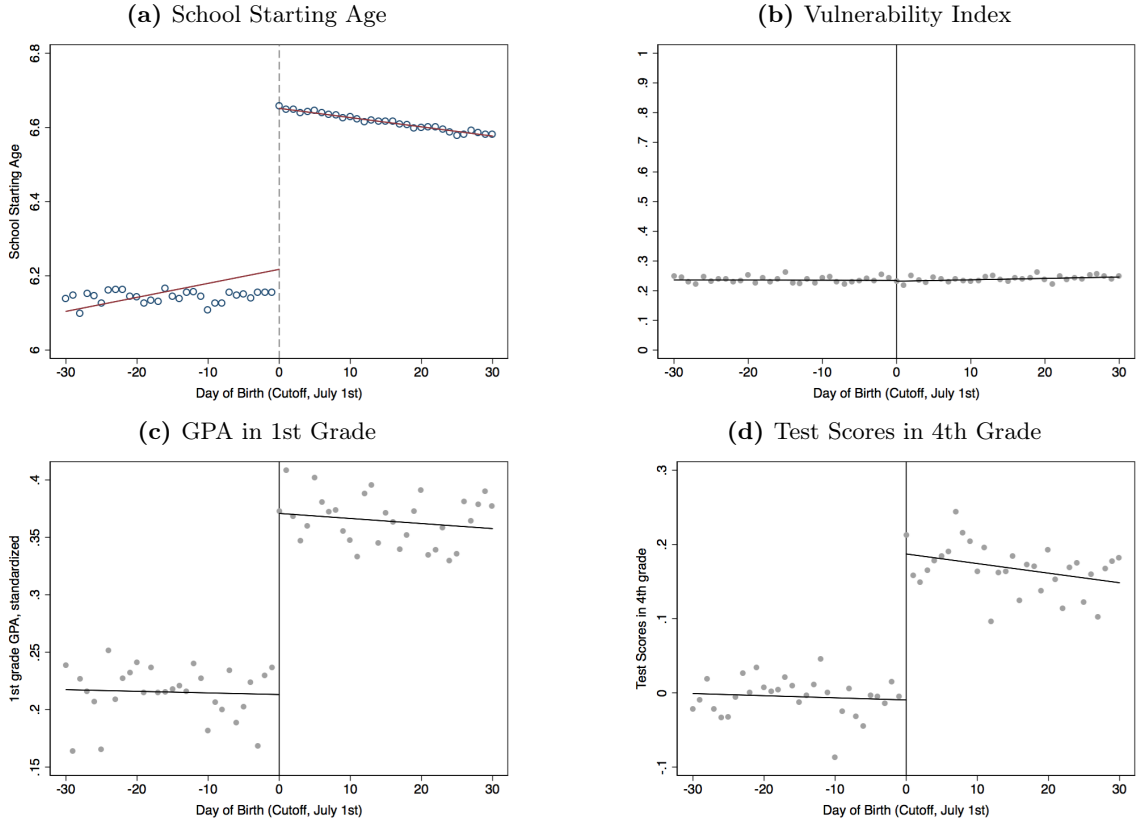
Note: The graphs in Figure B.2 plot the mean of the y-axis variables within equal sized bins of SIMCE test scores (in standard deviation units) in 4th grade, with approximately 250K observations each. The y-axis variables are the average of our measures of parental financial investments (having computer, 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

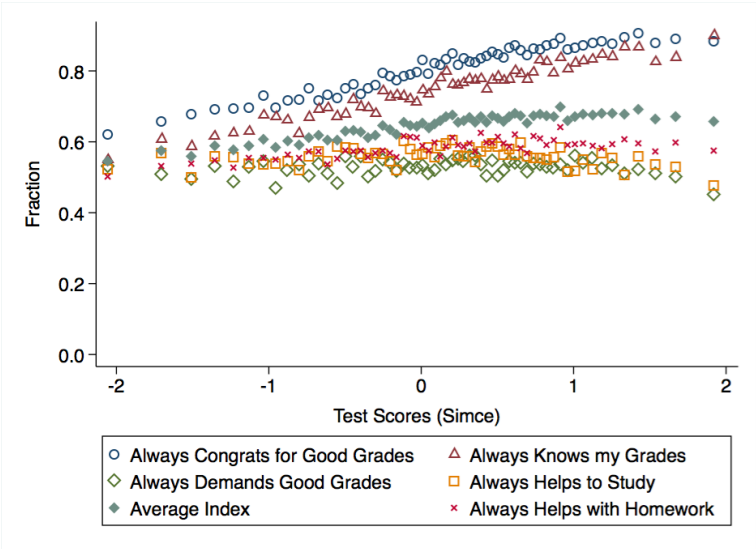
Notes: Table B.1 show the coefficient $\widehat{\alpha_1}$ estimated from the equation (1) for first graders on age at school entry, the vulnerability index and in-school outcomes. These are first grade GPA (GPA1, standardized within school and grade) and the Language-Math score in 4th grade. Robust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

B.2 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

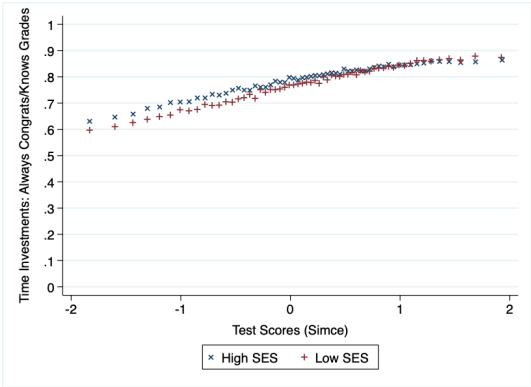


Note: The graphs in Figure B.4 plot the mean of the y-axis variables within equal sized bins of SIMCE test scores (in standard deviation units) in 4th grade, with a sample of approximately 460K observations. The y-axis variables are our measures of parental time investments.

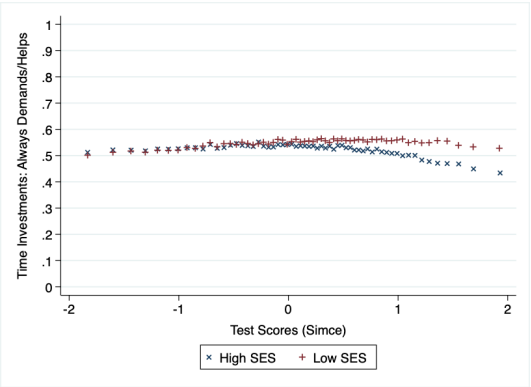
In Figure B.5 we show the same correlations after grouping these two groups of parental investments, by socioeconomic status. The graphs in Figure B.5 show that the general pattern of correlations holds for both high and low socioeconomic status, with a slighly more marked u-shaped profile for high SES parents and students.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) Always Congrats/Knows



(b) Always Demands/Helps



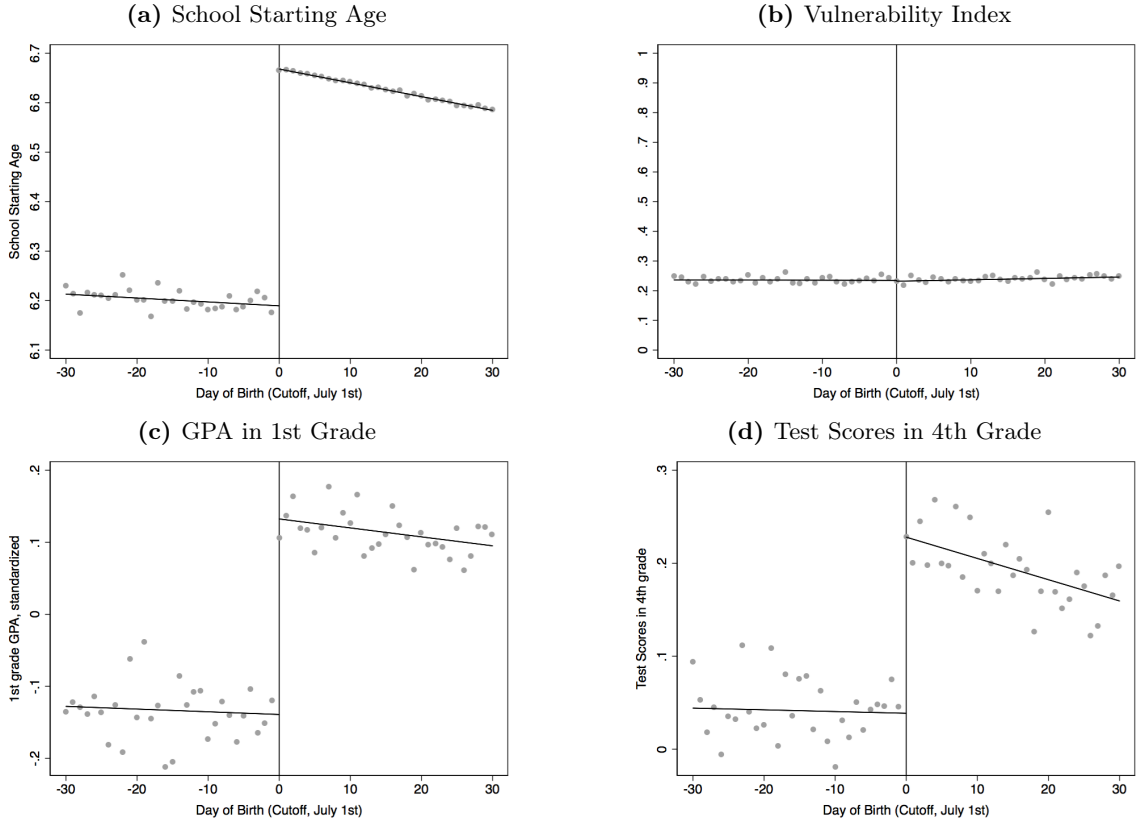
Note: The graphs in Figure B.5 plot the mean of the y-axis variables within equal sized bins of SIMCE test scores (in standard deviation units) in 4th grade, with approximately 250K observations each. The y-axis variables are the average of our measures of parental time investments.

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.

B.3 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.

Tables B.3 and B.4 summarize the results for each type of investment. The eigenvalues reported in Table B.3 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 in Table B.4 show the importance of each variable contributing the retained 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.

Table B.3: Factor Analyses: Eigenvalues

Factor	Financial Investments	Time Investments
Factor 1	1.911	1.718
Factor 2	0.837	0.939
Factor 3	0.771	0.938
Factor 4	0.480	0.738
Factor 5	.	0.668

Notes: Table B.3 reports the eigenvalues of factor analyses conducted using principal components, with a rotation of the loading matrix performed with the varimax Kaiser criterion (Kaiser 1958). Column 1 shows the results for financial investments and column 2 does the same for time investments. According to the Kaiser criterion we retain the factor with eigenvalues equal or higher than 1 for each case.

Table B.4: Factor Analyses: Loadings and Uniqueness

Financial Investments			Time Investments		
Variable	Loading	Uniqueness	Variable	Loading	Uniqueness
Computer	0.805	0.352	Congrats Grades	0.636	0.596
Internet	0.753	0.433	Knows Grades	0.547	0.701
High Spending	0.548	0.700	Demands Grades	0.369	0.864
More 10 Books	0.629	0.604	Helps Study	0.656	0.569
	.	.	Helps Homework	0.669	0.552

Notes: Table B.4 reports the factor loadings and uniqueness each variable contributing the retained factor 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.

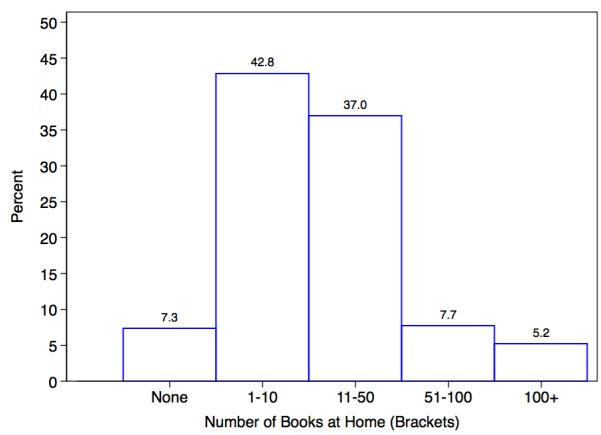
B.4 Robustness to the Grouping of Answers

Number of Books

Parents provide information on the number of books they own by brackets. Figure B.7 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.5 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.7: Number of Books at Home



Note: The graphs in Figure B.7 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.5: 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

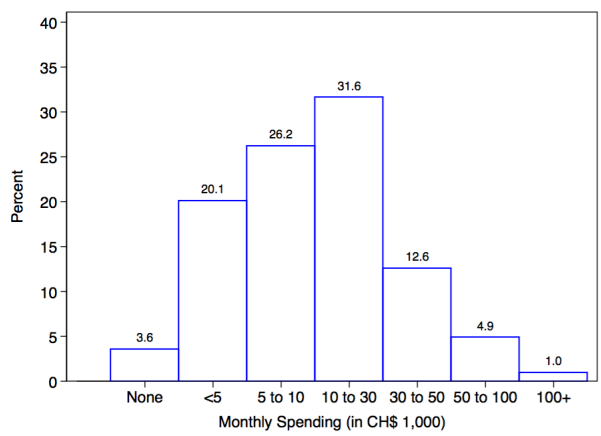
Notes: Table B.5 show the coefficient $\widehat{\alpha_1}$ estimated from the equation (1) for first graders on different ways of presenting the 'Books at Home' variable. Robust standard errors (in parentheses) are clustered by day of birth.

Spending Categories

Parents provide information on their monthly spending on educational inputs, by brackets. Figure B.8 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.6 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.8: Spending Categories



Note: The graphs in Figure B.8 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.6: 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.6 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.

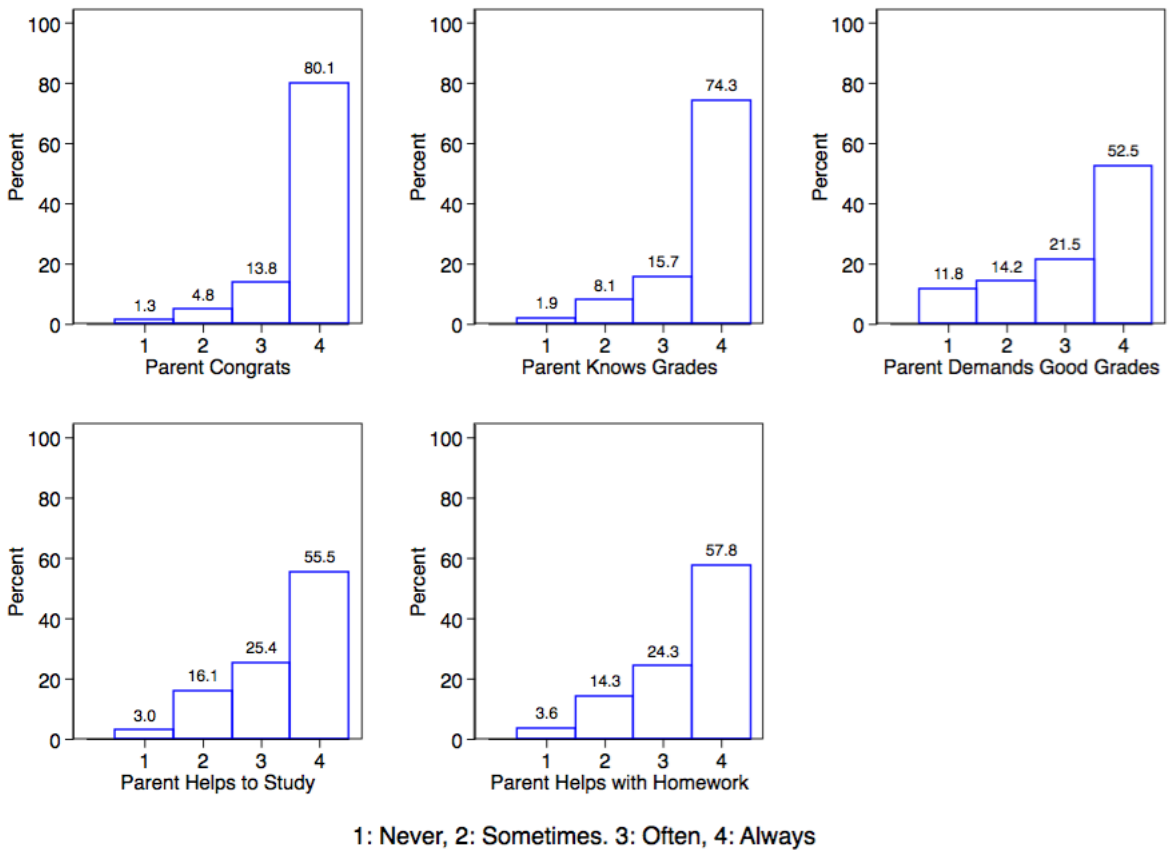
Time Investment Categories

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.9 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.7 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.9: Time Investment Categories



Note: The graphs in Figure B.9 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.7: Results for Time Investment Grouping

Panel 1	(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

Panel 2	(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

Panel 3	(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.7 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.

B.5 Effects on Raw Measures of Parental Investments

We also implemented a robustness exercise using the measures of parental time and material investments by themselves.

Time Investments: The estimates on raw measures are precise zeros, i.e., coefficients of tiny magnitude and precisely estimated. Table B.8 shows the results using the raw time investments (answers with four possibilities). Figure B.10 shows the graphs for the raw measures of time investments.

Table B.8: Results for Time Investments Measured in a 1 to 4 Scale

Panel 1	(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

Notes: The table shows the coefficient $\widehat{\alpha_1}$ estimated from the equation (1) for first graders on measures of time investments, in a 1 to 4 scale. Robust standard errors (in parentheses) are clustered by day of birth.

Material Investments: The estimates on raw measures of material investments are statistically significant and of sizable magnitude. Table B.9 shows the results using the individual measures of material investments in columns 3 to 6. In Appendix B.4 we also show that the effects are robust to different ways of definining the books and spending variables. Figure B.11 show the graphs for the raw measures of material investments.

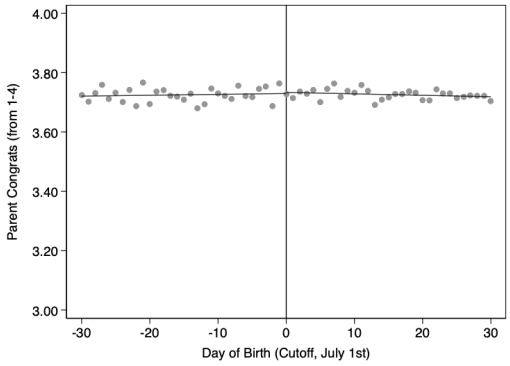
Table B.9: Results for Material Investments

	(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

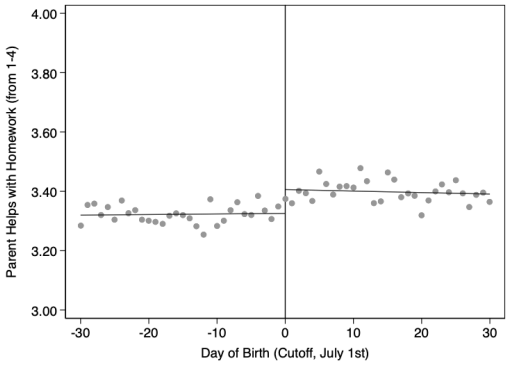
Notes: Table B.9 show the coefficient $\widehat{\alpha_1}$ estimated from the equation (1) for first graders on measures of material investments. Robust standard errors (in parentheses) are clustered by day of birth.

Figure B.10: Effects on Raw Time Investment Measures

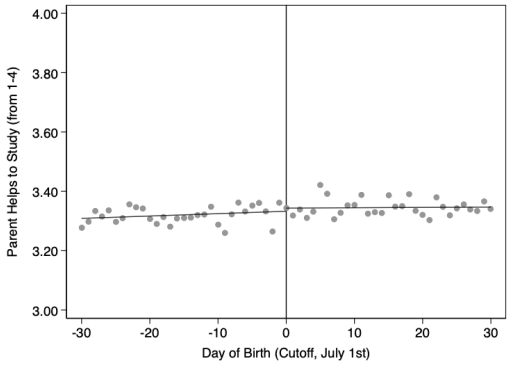
(a) Congrats for good grades (from1-4): 0.001 (0.010)



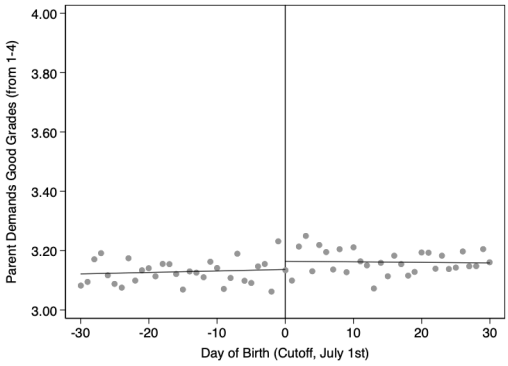
(b) Knows my grades (from1-4): 0.036 (0.015)



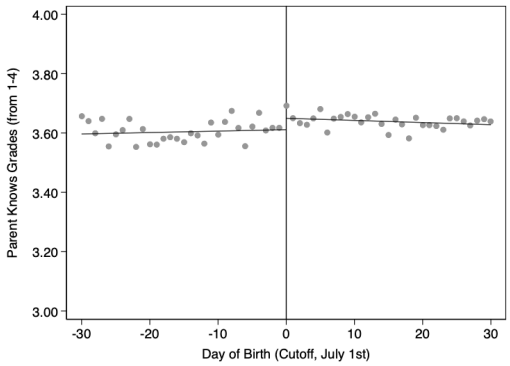
(c) Demands good grades (from1-4): 0.028 (0.025)



(d) Helps to study (from1-4): 0.010 (0.015)



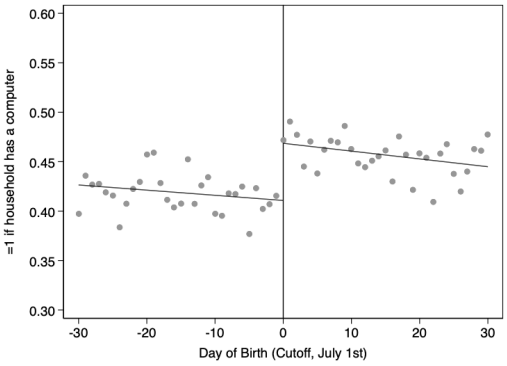
(e) Helps with homework: 0.078 (0.017)



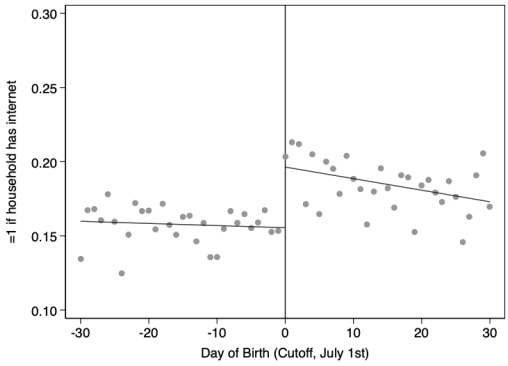
Note: The graphs in Figure B.10 plot the mean of the y-axis variable within day of birth, and fit estimated lines using all the underlying data, allowing for different slopes on each side of the cutoff. Each day of birth contains about 2K observations. The heading of each graph reports the estimated coefficient at the July 1st threshold, with corresponding standard errors in parentheses.

Figure B.11: Effects on Raw Financial Investment Measures

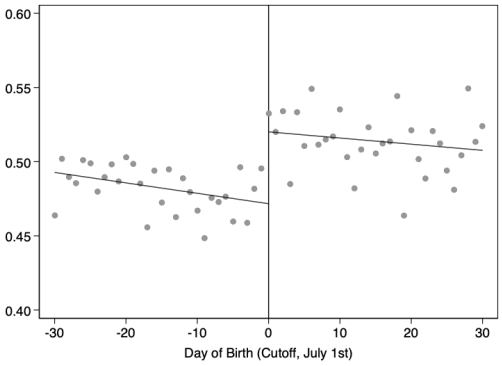
(a) Computer at Home: 0.043 (0.008)



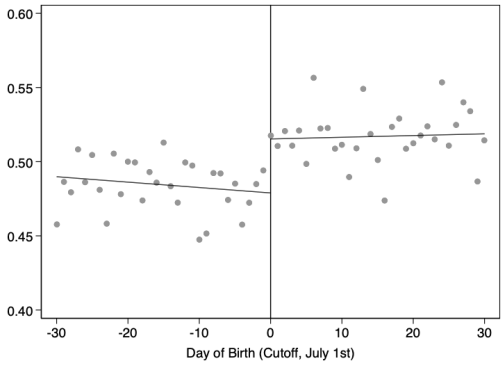
(b) Internet at Home: 0.032 (0.006)



(c) More than 10 Books at Home: 0.037 (0.008)



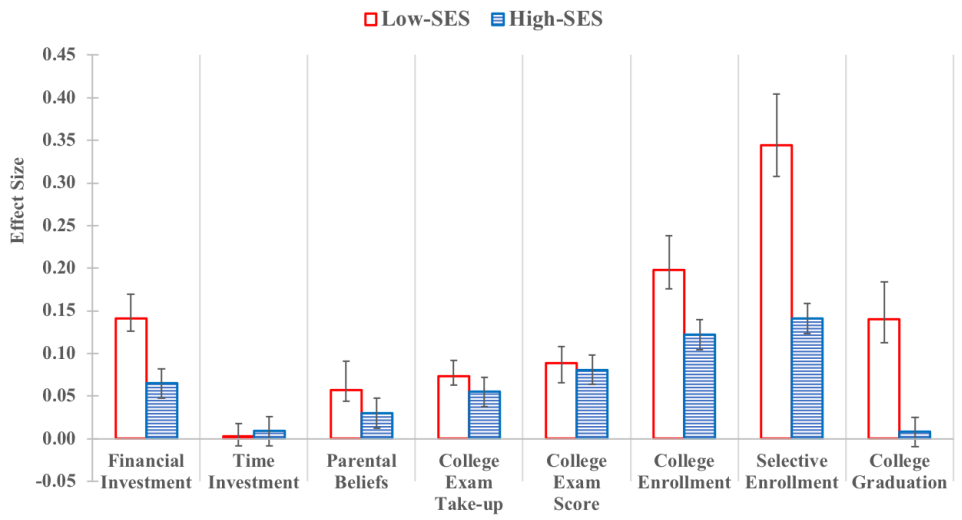
(d) Higher Spending: 0.022 (0.007)



Note: The graphs in Figure B.11 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 heading of each graph reports the estimated coefficient at the July 1st threshold, with corresponding standard errors in parentheses.

C Appendix: Additional Results

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.1: Effect Sizes, by Quartiles of Socioeconomic Status

	(1) Low SES	(2) Med-Low SES	(3) Med-High SES	(4) High SES	(5) Difference (1)-(4)
Panel 1: 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)
More than 10 Books at Home	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)
Panel 2: 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)
Panel 3: 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)
Panel 4: 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)
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 presents separate estimates of Equation (1) reported as effect sizes, on parental investments, beliefs, and children long run outcomes (in rows) by quartiles of the school vulnerability index (in columns 1-4). We label each quartile as low SES, med-low SES, med-high SES, and high SES, in columns (1)-(4). Column (5) reports the difference in effect size between the low and high SES groups. Standard errors are computed using the delta method.

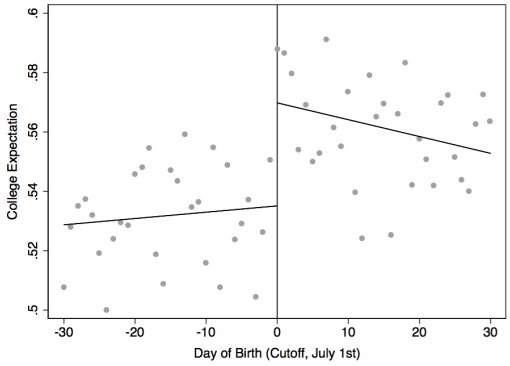
Table C.2: Effect Sizes, by Low-High Socioeconomic Status

Variable	(1) Low SES	(2) High SES	(3) Difference (1) - (2)
Panel 1: 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)
More than 10 Books at Home	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)
Panel 2: 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)
Panel 3: 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)
Panel 4: 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)

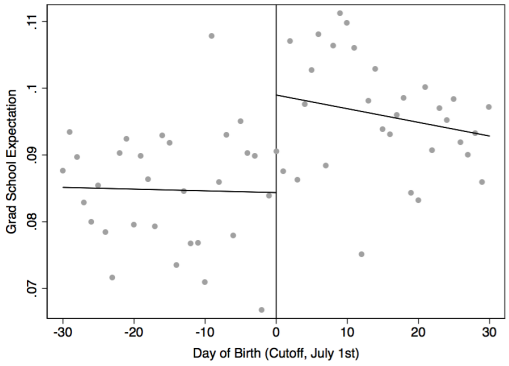
Notes: Table C.2 presents separate estimates of Equation (1) reported as effect sizes, on parental investments, beliefs, and children long run outcomes (in rows) by categories of school vulnerability index (in columns). We classify all observations below the median of the index as low SES, and all above as high SES). Column (3) reports the difference in effect size between the low and high SES groups. Standard errors are computed using the delta method.

Figure C.2: Effects on Parental Beliefs

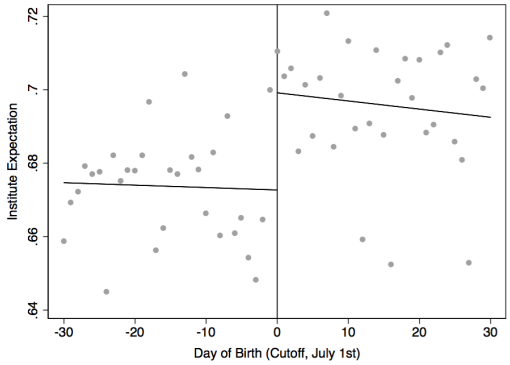
(a) College Expectations



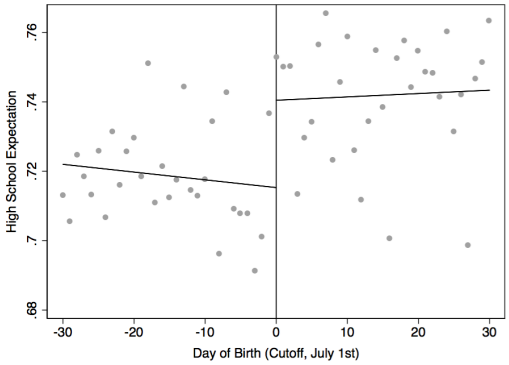
(b) Grad School Expectations



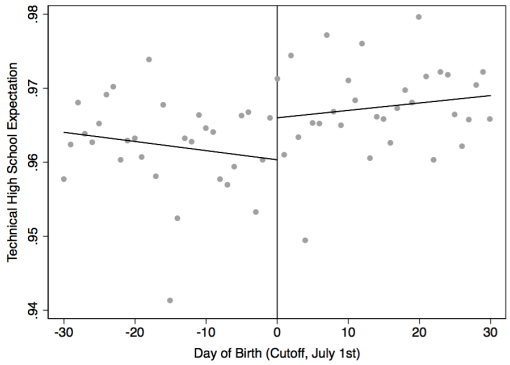
(c) Institute Expectations



(d) High School Expectations

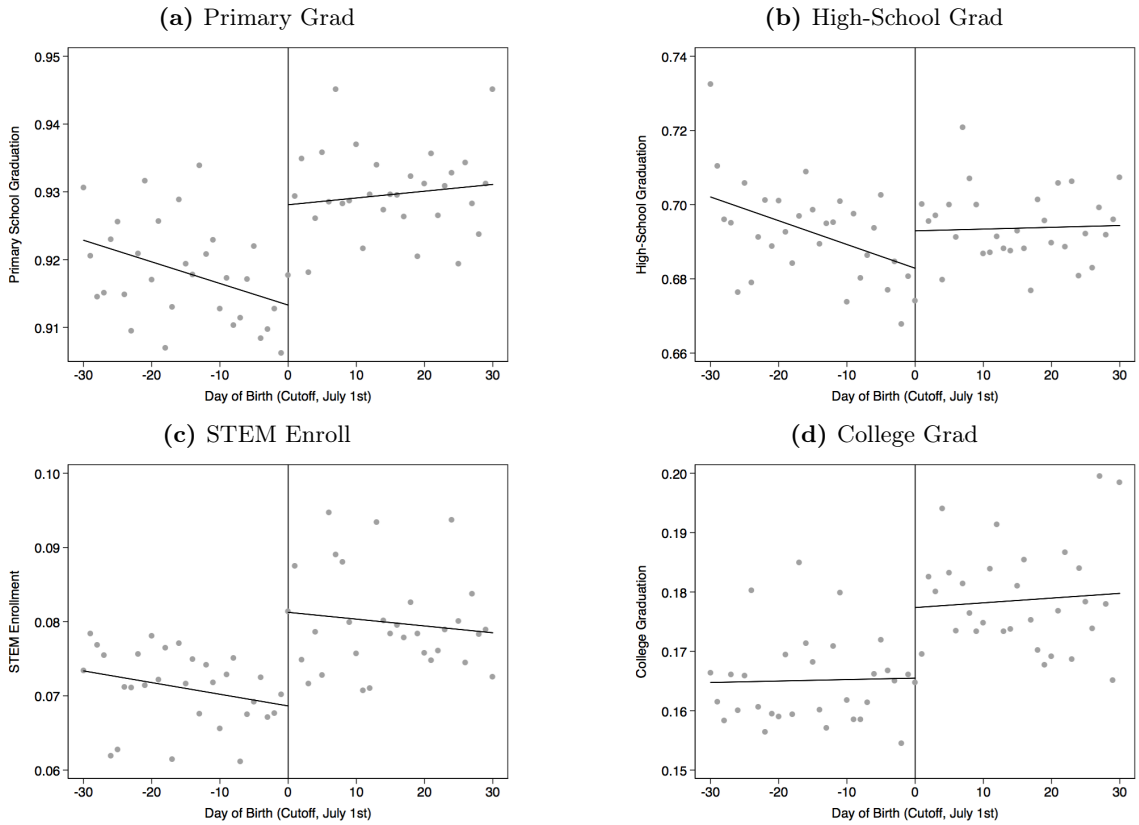


(e) Tech Hhigh School Expectations



Note: The figures in Figure C.2 are graphical analog to estimates in Table 4. Each graph plots 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 cutoffs. The dependent variables are described in subsection 3.2.

Figure C.3: Additional Effects on Long Run Outcomes



Note: The figures in Figure C.3 are graphical analog to estimates in Table 5 . Each graph plots 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 cutoffs. The dependent variables are described in the main text and in Table 5.

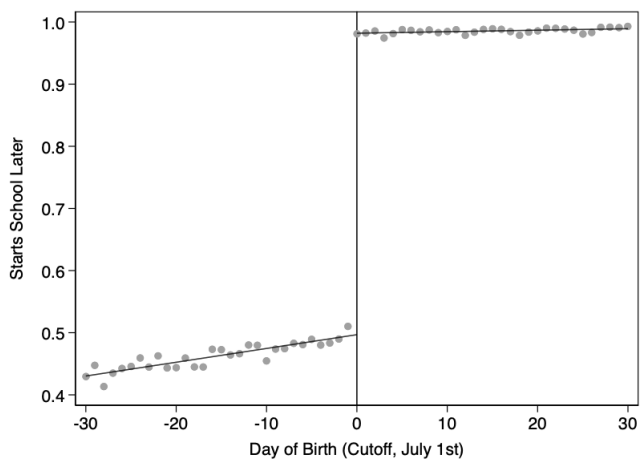
C.1 On Non-compliance and Interpretation of our Results

In the main text we focus on intent-to-treat estimates, which by definition are not adjusted by non-compliance rates. We interpret our results as the effect arising from the existence of the school entry rule (analog to being 'offered' a treatment) rather than to receiving the treatment itself. As we explain in the main text, we have several reasons supporting this methological decision.[20]

With that caveat in mind, we think that describing non-compliance help to understand how the school entry rules work in practice in our setup. Below, we document the fraction of students complying (or not) with the cutoff, classifying them as on-time entrants, red-shirters and green-shirters. We also characterize the students in these groups using the covariates in our sample.

Non-compliance : The first takeaway is that most of the non-compliance is due to red-shirting. We classify students as on-time students, redshirters or greenshirters. The on-time students are those born in June who started early, or those born in July who started late. The redshirters are those born in June who started school later, and the greenshirters were born in July but enrolled early.
Figure C.4 shows the fraction of students who start later by date of birth.To the left of the cutoff, almost half of June born children do not comply with the rule and delay enrollment (they are the redshirters). To the right of the cutoff, almost all July born children comply with the rule and delay enrollment. The fraction of those born in July who enroll early (greenshirters) is extremely small (about 2%), which is consistent with the enforcement of the rule, which we detail next.

Figure C.4: Starting School Later



[20] First, taking non-compliance into account would involve computing a LATE, for which we need additional assumptions to hold (e.g., exclusion restriction and monotonicity). Dhuey et al. (2019) provide an eloquent description of these topics. We expand on these issues in the Methods Section. Second, our ITT estimates approximate a lower bound of the effects. The reason is that with a binary instrument (being assigned to start early or late due to the rule), the denominator in the Local Average Treatment effect is always lower than 1. Then the LATE would be mechanically higher than the reduced-form estimate, and thus we see our ITT estimates as conservative.

Table C.3 characterizes on-time students, redshirters and greenshirters. The main takeaway is that redshirters tend to come from better socioeconomic backgrounds, which is similar to what has been found in the U.S. [21]

Table C.3: Characterizing On-time Students, Redshirters and Greenshirters

	(1)	(2)	(3)	(4)
	June Born		July Born	
Variable	On-Time	Redshirter	On-Time	Greenshirter
Father's Schooling	10.77	11.05	11.01	10.80
Mother's Schooling	10.72	10.94	10.93	10.75
Girl	0.51	0.46	0.50	0.47
Class Size	28.52	31.55	30.66	28.83
School Vulnerability (0-100)	32.66	27.57	29.32	32.13
School in Capital Region	0.31	0.44	0.37	0.41
Rural School	0.17	0.11	0.14	0.18
Public School	0.55	0.51	0.52	0.50
Voucher School	0.45	0.49	0.48	0.50
Observations	31,858	27,363	57,646	842
Fraction	0.27	0.23	0.49	0.01

Enforcement of the School-entry Rules

How are school-entry rules enforced in practice? Principals at both at voucher and public schools rely on an important 129 page document called Circular 1. The Circular 1 is essentialy a handbook that summarizes all duties, rules to follow, and deadlines that are necessary to be recognized by the National Law as schools and receive the funds.

The government makes no difference between voucher and public schools when overseeing and auditing schools through a National Education Board (NBE or *Superintendencia de Educacion*). The NBE implements some random checks to enforce the Law. This procedure happens in both public and voucher schools because both receive funding in the form of a per-student amount.

Enforcing the school-entry is one of the easiest tasks for the NBE because schools are required to have all children's birth certificates and indicate at which grade they are enrolled. Principals are fined and eventually laid off if they enroll children that are too young to start school (greenshirters) but there is no penalty for delaying school (redshirting). This is consistent with the data showing almost no greenshirting but important fraction of redshirters in schools.

[21] See, e.g, Bassok and Reardon (2013).

D Appendix: Mediation Analysis

D.1. Model of Mediation

We consider the following linear model, closely following the notation in Fagereng et al (2021):

$$Y_z = \kappa_z + \underbrace{\sum_{j \in J_p} \alpha_z^j \theta_z^j}_{measured\ mediators} + \underbrace{\sum_{j \in J/J_p} \alpha_z^j \theta_z^j}_{unmeasured\ mediators} + X'\beta_z + \tilde{\varepsilon}_z \tag{D.1}$$

where z indexes the child's school starting status (which in our case takes two values, early or late). The Equation (D.1) specifies the potential outcome Y_z as a function of observed mediator factors (those belonging to set J_p), unobserved mediator factors (those not belonging to set J_p), a vector X of predetermined variables used in our main estimations (such as child gender, socioeconomic status, class size, etc. and a smooth function of date of birth) and an idiosyncratic error term $\tilde{\varepsilon}_z$.

With unmeasured mediators absorbed in the error term and the intercept, equation (D.1) becomes:

$$Y_z = \tau_z + \sum_{j \in J_p} \alpha_z^j \theta_z^j + X'\beta_z + \varepsilon_z \tag{D.2}$$

Where $\varepsilon_z = \tilde{\varepsilon}_z + \sum_{j \in J/J_p} \gamma_z^j (\theta_z^j - E(\theta_z^j))$ is a mean-zero error term, and $\tau_z = \kappa_z + \sum_{j \in J/J_p} \gamma_z^j E(\theta_z^j)$.

The coefficients associated to mediators and control characteristics are specified in the following way:

$$\alpha_z^p = \alpha_0^p + \alpha^p \cdot z \qquad \beta_z = \beta_0 + \beta \cdot z \qquad \tau_z = \tau_0 + \tau \cdot z \tag{D.3}$$

For identification of $(\alpha_0^p, \alpha^p, \beta_0, \beta)$ we need to impose the assumption that unmeasured mediators are uncorrelated with both X and the measured mediator variables, as shown in Heckman and Pinto (2015). To further simplify the model, Fagereng et al (2021) assume no interactions between treatment and mediators or control characteristics, i.e. $\alpha^p = \beta = 0$. This last assumption does not change their results, and does not change ours. We show estimates both imposing and relaxing this assumption in the next subsection.

The last component of the mediation model includes specifying the mediator variables as a function of the school starting status (z) and the vector of characteristics X as:

$$\theta_z^j = \mu_0^j + X'\mu_1^j + \mu_2^j \cdot z + \eta^j \quad j \in J_p \tag{D.4}$$

Where η^j is a mean zero error term. Therefore, the average causal effect of starting school later (z') versus early (z) can be decomposed into a direct and an indirect effect as:

$$E(Y_{z'} - Y_z) = (z' - z)\tau + \sum_{j \in J_p} \alpha_0^j E(\theta_{z'}^j - \theta_z^j) = \underbrace{(z' - z)\tau}_{Direct\ effect} + \underbrace{\sum_{j \in J_p} \alpha_0^j (z' - z)\mu_2^j}_{Indirect\ effect} \tag{D.5}$$

The computation of the direct and indirect effect involves estimating $(\tau, \alpha_0^j, \mu_2^j)$. To obtain estimates of τ and α_0^j we estimate:

$$Y = \tau_0 + Z\tau + \sum_{j \in J_p} \alpha_0^j \theta^j + X'\beta_0 + \varepsilon \tag{D.6}$$

where the variable Z is equal to one if child is born in July and is equal to zero if child is born in June of the same year. To estimate μ_2^j we estimate the linear model for the observed mediator variables with each mediator as dependent variable and X and Z as regressors. We present our results below.

D.2. Mediation Results

Main Results. Panel A in Table D.1 presents the estimates from equation (D.6), assuming $\alpha^p = \beta = 0$. We use college enrollment as our outcome, and parental financial investments and beliefs (defined as in our main text) as observed mediators.[1]

The results show that the mediator variables are strong predictors of college enrollment, ($\widehat{\alpha_0^i}$ and $\widehat{\alpha_0^b}$ in the first two rows of Table D.1) and that the association between school entry and college enrollment remains significant, after conditioning on the mediator variables ($\widehat{\tau}$, in the third row).

Panel B in Table D.1 presents the estimates for the μ_2^j's. These estimates are consistent with financial investments and beliefs responding positively to signals about the child's ability.

Table D.1: Coefficients from Linear Potential Outcome Equation

	Coefficient	Standard Error
Panel A		
$\widehat{\alpha_0^i}$	0.0548	(0.00235)
$\widehat{\alpha_0^b}$	0.1540	(0.00387)
$\widehat{\tau}$	0.0168	(0.00547)
Panel B		
$\widehat{\mu_2^i}$	0.108	(0.016)
$\widehat{\mu_2^b}$	0.021	(0.007)
Observations	51,818	

Notes: Panel A presents the estimates from equation (D.6), assuming $\alpha^p = \beta = 0$. Panel B reports the estimates for the μ_2^j's in the linear model for the observed mediators variables, with each mediator as dependent variable. We use timely college enrollment as our outcome, and parental financial investments and beliefs (defined as in our main text) as observed mediators.

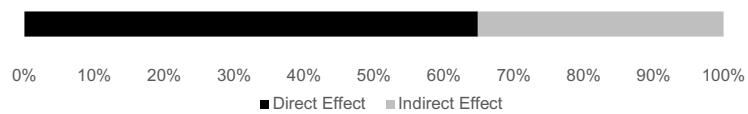
[1] Parental time investments have zero effect on our long run outcomes as we showed in the main text, and are therefore ruled out as mediators.

We use these coefficients to decompose the long run effect into direct and indirect effects as described in equation (D.5). We plot the results in Figure D.1.

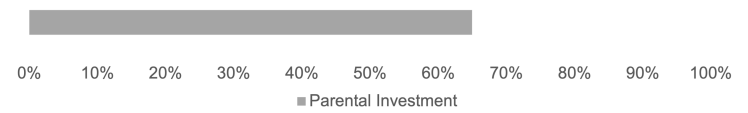
Panel (a) shows that 66% of the total effect is explained by direct effects of starting school later, while 34% of treatment effects are mediated by beliefs and financial investments. As a benchmark, Fagereng et al (2021) find that their mediators (parental wealth transfers being the most important) explain about 37% of the average causal effect from assignment to wealthier parents on children's accumulation of wealth. In panel (b) we show that approximately two thirds of the mediated (indirect) effect in our setup is explained by financial investment of parents.

Figure D.1: Decomposition of Causal Effect on College Enrollment

(a) Direct and Indirect (Mediated) Effects as a Share of the Causal Effects



(b) Parental Investments as a Share of the Indirect (Mediated) Effect

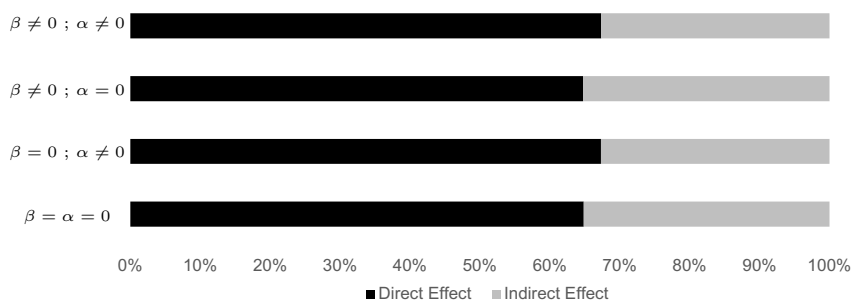


Note: Panel (a) in Figure D.1 decomposes the causal effect of starting school later on college enrollment into shares of indirect and direct effects, as illustrated in equation (D.5). Panel (b) shows how much of the indirect effect can be attributed to financial investments by parents.

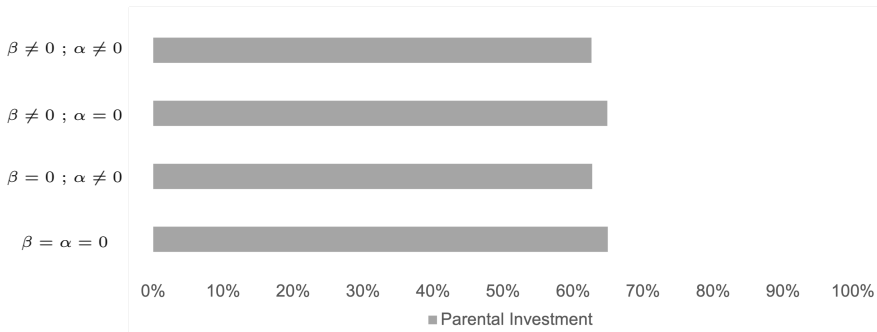
Robustness. We now show that the previous results are robust to relaxing the $\alpha^p = \beta = 0$ assumption. Panel (a) in Figure D.2 shows that the shares of direct and indirect effects remain stable under different assumptions for α^p and β. Panel (b) shows that the share of parental investments as a share of the indirect effect is also robust to relaxing the $\alpha^p = \beta = 0$ assumption. Consistently, Table D.2 reports that the estimates of $(\tau, \alpha_0^j, \mu_2^j)$ remain stable across different restrictions for α^p and β.

Figure D.2: Decomposition of Causal Effect on College Enrollment

(a) Direct and Indirect (Mediated) Effects as a Share of the Causal Effects



(b) Parental Investments as a Share of the Indirect (Mediated) Effect



Note: Figure D.2 replicates Figure D.1 imposing different restrictions on α^p and β.

Table D.2: Robustness: Results under Different Restrictions on α^p and β

	(1)	(2)	(3)	(4)
$\widehat{\alpha_0^i}$	0.0548^{***}	0.0548^{***}	0.0492^{***}	0.0490^{***}
	(0.00235)	(0.00235)	(0.00298)	(0.00312)
$\widehat{\alpha_0^b}$	0.154^{***}	0.154^{***}	0.152^{***}	0.152^{***}
	(0.00387)	(0.00389)	(0.00493)	(0.00509)
$\widehat{\tau}$	0.0168^{**}	0.0275	0.0142^{*}	0.0166
	(0.00547)	(0.0154)	(0.00614)	(0.0160)
Observations	51,818	51,818	51,818	51,818
Restrictions	$\alpha^p = \beta = 0$	$\alpha^p = 0$; $\beta \neq 0$	$\alpha^p \neq 0$; $\beta = 0$	$\alpha^p \neq 0$; $\beta \neq 0$

Notes: Table D.2 replicates Panel (A) of Table D.1 in column (1) and imposes different restrictions on α^p and β in columns (2) to (4).

E Appendix: Children within Same Families

In this section we study multiple children within the same families. We implement a family fixed effects strategy, along the lines of Dhuey, Figlio, Karbownik and Roth (JPAM, 2019), restricting the sample to families where we observe at least two siblings in our data. Then, we further require that these siblings are the first two in the family and that both are born in either June or July.

Data

The Ministry of Education (MINEDUC) shared with us a dataset with family ID for all students in school between years 2005 to 2014.

Analitycal Sample. We keep the first two siblings within each family, born in either July or June in years 1996 to 1998, which correspond to the birth cohorts in our main analysis. The estimating samples remain large, varying in sizeby the availability of the outcomes.

From survey data, we have information on parental investments for 4,604 children (2,302 families) and on beliefs and test scores for 5,320 children (2,660 families)[2]. From administative records, we have data on college attendance for 18,632 (9,316 families).

Specification

With these data we estimate the same sibling fixed effects specification as in Dhuey, Figlio, Karbownik and Roth (JPAM, 2019):

$$y_{ij} = \alpha_j + \alpha_1 Z_{ij} + \alpha_2 X_{ij} + \mu_{ij} \tag{E.1}$$

Where all X variables are the same as in the main text, and are now also indexed by a family subscript j. The Z variable is an indicator for being born in July, and the family fixed effect α_j accounts for observed and unobserved characteristics that are the same across different siblings within families. We also include order of birth, sex, and year of birth in X_{ij}. As such, the identifying variation now compares two siblings born in months at each side of the cutoff. We cluster standard errors by family.

Results

Overall, the results show that controlling for family characteristics that are fixed across siblings does not change our point estimates in a substantial way. These findings suggest that our main results are not driven by family characteritics. Table E.1 reports the sibling fixed effects results (in even columns) compared to our main estimates (in odd columns).

The results from the sibling fixed effects strategy show that July-born children are 4.8 pp more likely to have a computer at home (4.3 pp in our main results), and 4.2 pp more likely to have an Internet connection (3.7 in our main results). In addition, results show that parents are 2.6 pp more

[2]The availability depends on whether the exam was implemented in the year each sibling attended 4th grade.

likely to believe that their child will complete college (2.1 in our main results). In this case the magnitudes are similar but we lose some precision when using sibling fixed effects.[3]

The last columns show that when using sibling fixed effects July-born children score 0.201σ in tests scores at fourth grade (0.207σ in our main results) and are 2.5 pp more likely to attend college on time (3.7 pp in our main results).

Table E.1: Robustness of Main Results to Sibling Fixed Effects

	Computer at Home		Internet at Home		College Expectation		Test Scores in 4th Grade		College Enroll	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$\widehat{\alpha_1}$	0.043***	0.048***	0.032***	0.042**	0.021***	0.026	0.207***	0.201***	0.037***	0.025***
	(0.008)	(0.017)	(0.006)	(0.017)	(0.007)	(0.018)	(0.010)	(0.030)	(0.005)	(0.009)
Observations	51,818	4,604	51,818	4,604	51,818	5,320	117,709	5,320	117,709	18,632
Siblings FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes

Notes: Even columns in Table E.1 show the coefficient $\widehat{\alpha_1}$ estimated from equation (E.1) on measures of parental financial investment and expectations (beliefs). Financial investments and expectations are having computer and Internet connection at home, and expectations refer to whether the caregiver believes the child would complete college in the future. Odd columns report the same estimate but for our main specifications. Robust standard errors (in parentheses) are clustered at the family level. All specifications control for order of birth, sex, and cohort of birth.

[3]This is in part due to lower sample size, and also to lower 'within-unit' variability, as written by Lee and Lemieux (JEL, 2010), page 338: "*The inclusion of individual dummy variables may lead to an increase in the variance of the RD estimator for another reason. If there is little 'within-unit' variability in treatment status, then the variation in the main variable of interest (treatment after partialling out the individual heterogeneity) may be quite small.*"

F Appendix: Human Capital Accumulation Model

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

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

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

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

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

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

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

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

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

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

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

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

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

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.