School Starting Age and the College Application Margin: Evidence from Chile

Cristhian Molina*

November 6, 2025

[Preliminary draft, do not cite]

[Click Here for the Latest Version]

Abstract

This study examines the impact of delayed school enrollment on college application and enrollment outcomes in Chile using a quasi-experimental design based on first-grade enrollment age cutoffs. Our findings reveal that students who enroll in first grade one year later are approximately 2% more likely to apply to and enroll in college. Additionally, a one-year delay in enrollment is associated with a higher high school GPA by 0.08 standard deviations. The local average treatment effect estimates on standardized test scores show no statistical difference in math scores and an increase of 0.05 standard deviations in verbal scores, particularly for students born after the July 1^{st} cutoff. These results are more pronounced among female students. These findings suggest that policy adjustments to school entry age can enhance educational outcomes and equity, with a particular emphasis on supporting younger and male students to close achievement gaps.

Keywords: Relative Age Effect, School Entry Age, Academic Performance, Standardized Test Scores, Education Policy

JEL: I21, J01, J13

*Department of Economics, University of Illinois at Urbana-Champaign. cmolina3@illinois.edu

I Introduction

The timing of formal school entry has been a focal point in educational research, as it significantly influences academic performance, educational pathways, and long-term outcomes. The relative age effect (RAE) in education refers to the academic and developmental disparities that emerge among classmates due to variations in age at the time of school entry. This effect is especially pronounced in educational systems with rigid cutoff dates for school entry, leading to potential age differences of up to a year within the same grade. Such disparities can profoundly impact early academic performance, college enrollment, and labor market success.

Research on the relative age effect has produced mixed results. Studies from developed countries, such as those by Cascio & Schanzenbach (2013); Crawford et al. (2010); Datar (2006); Elder & Lubotsky (2009) have shown that older students tend to perform better in standardized tests, with observed effects ranging from 0.11σ to 0.7σ, in the case of U.S. and England. Bedard & Dhuey (2006) analyzed data from 19 developed countries[1], finding even larger effects, up to 0.8σ during the initial school years. However, these advantages often diminish over time, as evidenced by research conducted using data from Norway and the U.S. indicating a reduction to a tenth of the original effect by secondary education (Elder & Lubotsky, 2009; Grenet, 2011). Now, as suggested in Black et al. (2011); Grenet (2011) and Dobkin & Ferreira (2010), there are costs associated to the decision of postponing school enrollment since starting school at an older age can negatively affect long-term educational attainment, or wages before the age of 30.

In contrast, studies in developing countries, such as Peña (2017) and McEwan & Shapiro (2008), suggest that in places like Chile and Mexico, older students also outperform younger peers in early academic years, but these effects tend to persist over time. Despite extensive research, there remains a lack of consensus on how the relative age effect varies across different contexts and demographic groups and whether these effects extend into higher education and career outcomes.

[1] The list of countries included is Austria, Belgium, Canada, Czech Republic, Denmark, England, Finland, France, Greece, Iceland, Italy, Japan, New Zealand, Norway, Portugal, Slovak Republic, Spain, Sweden, and the United States.

This study is driven by ongoing debates regarding the optimal age for school entry and the potential policy implications of adjusting school entry cutoff dates. It aims to contribute to the understanding of the relative age effect in education by investigating the impact of delayed school enrollment on college application and enrollment outcomes, focusing specifically on the Chilean context. Building on McEwan & Shapiro (2008) findings, which showed that delaying school entry by a year reduces first-grade retention and enhances middle-school performance, this research explores the broader implications of these early educational decisions on subsequent academic achievements and college enrollment. By examining the long-term effects of the relative age effect, and how this effect varies with students' socioeconomic status (SES), this research aims to provide policymakers with insights into whether adjustments to these cutoff dates could imply more equitable educational outcomes.

The primary research question addressed in this study is how the relative age effect influences the likelihood of applying to and enrolling in college. We also examine how the age at school entry impacts students' performance on college admission tests, separating between effects in standardized tests and historical performance in high school. Finally, evidence suggests that gender is an important factor correlated with academic redshirting in parents, and therefore we evaluate whether the relative age effect in college application outcomes varies by students' gender.

This paper makes several significant contributions to the existing literature on the relative age effect. Firstly, it builds on the work of McEwan & Shapiro (2008) by analyzing how delayed school enrollment impacts college application and enrollment outcomes in Chile. Secondly, it offers empirical evidence from a developing country context, which has been underrepresented in previous studies. Thirdly, by utilizing exact birth dates and first-grade enrollment age cutoffs, we establish a quasi-experimental framework that generates estimates in our context. This approach addresses many issues highlighted in related literature regarding the failure of instruments to satisfy the monotonicity assumption (Barua & Lang, 2016). Additionally, we provide evidence supporting the assumptions behind our empirical strategy, demonstrating that parental characteristics are not linked to sorting behavior around the thresholds (Crawford et al., 2010; Shigeoka, 2015).

The empirical strategy employed in this study is based on an instrumental variable approach, in which schools' cutoff dates introduce a natural variation in students' first-grade enrollment ages. We leverage this discontinuity in school entry cutoff dates to estimate the impact of age at school entry on subsequent academic outcomes. This method allows for a comparison between students who are marginally eligible to start school and those who are marginally ineligible, thereby isolating the effect of relative age from other confounding factors. Data used for this analysis comes from comprehensive administrative records from Chile, including information on students' exact birthdates, test scores, school performance, and SES.

Key findings indicate that older students are more likely to apply to and enroll in college. Further analysis reveals that these students generally achieve higher grades in high school, positioning themselves better in grade distributions compared to their younger peers. Interestingly, older students show better performance in verbal tests, while no significant differences are found in math tests, suggesting that some effects identified in earlier studies may diminish over time. Additionally, our estimates reveal that the relative age effect appears more pronounced among female students, highlighting the role of gender in moderating these dynamics.

The remainder of the article is organized as follows. Section II briefly introduces the elements relevant to this research from the Chilean education system, Section III describes the empirical strategy and the possible threats to its validity, while Section IV describes the data. Sections V and VI discuss results for college application/enrollment and test score outcomes, respectively.

II Institutional Background

School enrollment

Chile has a comprehensive educational system that is mandatory from first grade until high school. Students can enroll in preschool at the age of three, and they start the first grade of school (already mandatory) at approximately age six. They are expected to attend eight years of *educación básica* (grades 1-8), and four years of *educación media* (grades 9-12). Chile's official first-grade enrollment cutoff is April 1, with the school year beginning on March 1, implying a minimum enrollment age of 5.92 years. However, until 2017, a Ministry of Education decree allowed schools to implement cutoff dates as late as July 1, resulting in lower minimum enrollment ages.[2] Following McEwan & Shapiro (2008), we present empirical evidence that sharp cutoffs appear on the first day of April, May, June, and July, though the last was the most common in Chilean schools.

College application/enrollment

In Chile, admission to college is decided by a centralized system that uses a student–proposing deferred acceptance algorithm (DAA) to match students with programs. Students can apply to 1,890 programs hosted by 41 universities. Admission to these programs requires standardized tests known as "*Prueba de Selección Universitaria*" (PSU). All entrance exam takers complete exams in mathematics and language, and many students also take optional tests in history and science. Scores are scaled to a distribution with a range of 150 to 850 and a mean and median of 500. Entrance exam scores, together with high–school GPA and class ranking, are the primary components of the composite scores used for admissions, scholarships, and student loan eligibility. While optional, about 95% of high school graduates annually participate in the entrance exam. After taking the entrance exam and receiving their scores, students can submit a rank–ordered list of preferred programs. The centralized system uses a DAA to optimize matchings based on student preferences and programs' capacity.

III Empirical strategy

We use data on Chilean students to estimate the effect of enrollment age on student outcomes. Considering the unique setting in Chile, we follow the strategy employed by McEwan & Shapiro

[2] In 2017, Chile published the Decrete 1718 that enforced April 1st as the only first-grade enrollment cutoff for all schools

(2008). Our initial approach involves a linear model estimated by ordinary least squares (OLS):

$$O_{ig} = \beta_0 + \beta_1 A_i + X_i \beta_2 + \varepsilon_{ig} \tag{1}$$

where O_{ig} is the outcome of student i at the end of grade g, A_i is the student's age in decimal years upon beginning the first grade, X_i is a vector of child and family variables determined before the child's birth, and the independent errors ε_{ig} are distributed $N\left(0, \sigma^2\right) .\beta_1$ represents the effect of delaying enrollment by one year. However, if $\operatorname{Cov}\left(A_i, \varepsilon_{ig}\right) \neq 0$, then $\hat{\beta}_1^{OLS} \neq \beta_1$. This seems plausible since children with lower (and unobserved) physical, cognitive, or social readiness-factors potentially correlated with outcomes-are more likely to delay enrollment.

To address this issue, we exploit exogenous variation in first-grade enrollment age created by Chile's enrollment cutoff dates. Students turning six on or after cutoff dates must delay enrollment by one year. Identification of enrollment age effects is based on comparing the outcomes of "treated" students born on or just to the right of cutoffs, with those of untreated students born just to the left of cutoffs. The interpretation of our coefficients relies on the assumption that birth dates are random near cutoffs, similar to a very local randomized experiment (Cattaneo et al., 2019). At a minimum, we assume that precise birth timing near cutoffs does not introduce sharp differences in unobserved variables affecting student outcomes.

To obtain estimates based on this variation, let B denote a student's day of birth in the calendar year, omitting the i subscript. Allowing for leap years, $B = 1$ for birthdays falling on January 1 and $B = 366$ for birthdays falling on December 31. Define four dummy variables, $D_j = 1\left(B \geq \bar{B}_j\right) \forall j \in \{1, 2, 3, 4\}$, indicating values of B that equal or exceed enrollment cutoffs. [3] To compare student outcomes around each discontinuity, we estimate the following equation via OLS:

$$O = \phi_0 + \phi_1 D_1 + \phi_2 D_2 + \phi_3 D_3 + \phi_4 D_4 + f(B) + u \tag{2}$$

[3] The enrollment cutoffs are $\overline{B_1} = 92$ (April 1), $\overline{B_2} = 122$ (May 1), $\overline{B_3} = 153$ (June 1), $\overline{B_4} = 183$ (July 1)

where $f(B)$ is a function of B that captures smooth, seasonal effects of birth dates on student outcomes (McCrary & Royer, 2011; Van Der Klaauw, 2002). We specify it as a piecewise quadratic polynomial (later, we visually assess the fit of this functional form and also use a higher-order polynomial):

$$f(B) = \sum_{k=1}^{2} \delta^k B^k + \sum_{j=1}^{4} \sum_{k=1}^{2} \delta_j^k D_j \left(B - \bar{B}_j\right)^k$$

where δ_j represent coefficients on polynomial terms.

In Equation 2, the ϕ_j terms summarize the sharp differences in outcomes between students born close to each enrollment cutoff, due to the enrollment delay treatment. However, three possible parental behaviors suggest these estimates capture only intent-to-treat effects. First, parents may voluntarily delay a child's enrollment beyond the legal minimum age. Second, parents may request that local school personnel allow the child to enroll before reaching the legal minimum age. Third, families may choose between schools with different enrollment cutoff dates.

Given these behaviors, we estimate the following equations via two-stage least squares (TSLS):

$$A = \alpha_0 + \alpha_1 D_1 + \alpha_2 D_2 + \alpha_3 D_3 + \alpha_4 D_4 + f(B) + v \tag{3}$$

$$O = \beta_0 + \beta_1 A + f(B) + \varepsilon \tag{4}$$

Estimates of Equation 3 reveal whether birthdays near enrollment cutoffs create sharp variation in enrollment age. Given partial compliance and the four discontinuities, $\alpha_j < 1 \forall j \in \{1, 2, 3, 4\}$. In Equation 4, β_1 is the effect on O of a one-year increase in enrollment age. If treatment effects vary across students, β_1 is the weighted average of four local average treatment effects (LATEs), with weights proportional to the ability of each D instrument to predict enrollment age (Angrist

& Krueger, 1992). We also can generate four estimates of the effect at each cutoff $\overline{\text{B}}_j$. [4] We interpret the four estimates as LATEs for students with birthdays near the respective cutoffs who are induced to delay enrollment.

For Equations 4 and 5 to provide a consistent estimator of β_1, the excluded instruments must be uncorrelated with unobserved variables that influence outcomes:

$$\text{cov}\left(D_j, \varepsilon\right) = 0 \quad \forall j \in \{1, 2, 3, 4\}$$

One potential threat to internal validity is the precise timing of births among families with unobserved characteristics that affect outcomes, which might create a correlation between D_j and the error ε. A plausible scenario is that higher-income parents, who have better access to medical services, may plan births to occur just before or after the cutoff dates. Another concern is that our administrative data might omit some students who attend private schools. If the likelihood of attending private schools changes significantly at the enrollment cutoff, this could lead to sample selection bias, thereby violating the assumption that D_j and ε are uncorrelated.

To assess whether precise birth timing and sample selection influence our estimates, we conduct two tests to verify the necessary conditions for this assumption. First, we re-estimate Equation 2, using each component of X as the dependent variable. In each iteration, we test the null hypothesis that $\phi_j = 0 \quad \forall j \in \{1, 2, 3, 4\}$. Failure to reject these hypotheses would indicate that observed covariates vary smoothly around the enrollment cutoffs. Second, we examine the impact of including predetermined X in Equations 4 and 5 changes $\hat{\beta}_1^{TSLS}$. If our assumption is true, then including predetermined variables will not change the magnitude of $\hat{\beta}_1^{TSLS}$.

Additionally, we use another strategy to provide suggestive, though not definitive, evidence supporting the validity of our instruments. We inspect histograms for any unusual breaks around the

[4] We estimate the first-stage Equation 3, and then four variants of Equation 4, each controlling for three of four D_j. For example, a second-stage regression controlling for D_1 to D_3 would identify an effect in the vicinity of the July 1 cutoff (also see Van Der Klaauw (2002)).

enrollment cutoffs, as such breaks would imply that birth timing might invalidate our assumption.

IV Data

In our study, we combine student–level data from two different sources, provided by different units depending on the Chilean Ministry of Education.

First-grade enrollment

First, we use data from six consecutive cohorts of first-graders between 2002 and 2007, which the schools reported to the Ministry of Education at the end of each academic year. This data reports students' exact birth date, gender, year of enrollment in first grade, final grade for the academic year, final situation (pass, retained, transferred), and a unique identification number similar to a U.S. Social Security number that we use to link different sources of data. We construct variables that measure students' exact enrollment age (A), type of school in first grade, and GPA. We denote a student's exact enrollment age (A) as the days elapsed between birth and March 1 of the first year in which the student appears enrolled in first grade, divided by 365.25.

To reach our estimation sample, we apply several exclusion restrictions to the first-grader data. Primarily, we exclude students attending rural schools since, in some small rural areas, several ages and levels are grouped into one to create sizeable cohorts. This could affect our outcomes of interest, and therefore, we do not include them in our analysis. Second, we include only the first observation of each student in first grade and drop subsequent observations if the student is retained. Third, we exclude observations with missing values in any of the variables used in our analysis.

Given that our main interest lies in college application rates, we augment this data with college–application test results and application records.

College application/enrollment

The Department of Evaluation, Measurement, and Educational Registry (DEMRE) provides a second data source. It encompasses the college admission scores of all test-takers, together with the rank order lists they submit when applying to college programs, if any. From this source, we extract several variables describing students' college application portfolios, including scores in math and verbal tests, GPA score, and GPA rank score (we normalize them to use their Z-scores in the analysis). We also get students' eligibility for the excellence scholarship, which is a fixed amount of money directly targeted towards paying students' tuition, and granted by the state to students in the top 10% of the GPA distribution of their cohort.[5]Finally since we observe the full universe of students taking the test, we collect whether they apply and or enroll in a college degree. As with our first grade enrollment sample, We only consider the first time that students take the college admission tests and drop subsequent attempts. DEMRE also collects a short survey given to students before the college admission tests, from which we construct variables reflecting the socioeconomic composition of the student's household, like family income and parents' education.

After processing the two sources of information, our estimation sample is limited to those students that can be traced starting their first–grade enrollment until they take the college admission test. This includes $759,990$ first-graders between the years 2002 and 2007, enrolled in $4,762$ urban schools, and taking the college admission test between years 2013 and 2018[6] Of the full universe of first graders considered in our sample, we observe college admission outcomes for nearly 91% of them[7]

[5]The student granted the scholarship must come from either a public or semi-public school.

[6]We exclude college admission tests from 2019 since college admission tests could not be conducted normally during that year, due to social protests happening in the country. We also exclude years after 2019 due to the COVID19 pandemic and a change in the evaluation test.

[7]We observe 7% out of the remaining 9% of unmatched first-graders in future college admission tests, being dropped due to our exclusion criteria. The remaining 2% are students that presumably did not take college admission tests.

Descriptive Statistics

The average student entered first grade at age 6.27. The latest possible enrollment cutoff is July 1^{st}, implying a minimum enrollment age of 5.67. Some schools use the April 1^{st} cutoff, suggesting that the oldest student enrolled in that year should be 6.92 years old, instead of 6.67 if July 1^{st} was the mandatory cutoff for every school. Table 1 shows some descriptive statistics of the students comprised in our analysis. Each column in the table groups students by their first-grade enrollment age, and cells display the mean of each variable for students contained in that group. Stars reflect if the mean in that cell is statistically different than the one in column (2). It can be seen that less than 1 percent of the sample violates this cutoff by enrolling earlier, while less than 2% violate it by postponing enrollment for one more year (last row of Table 1).

Students who voluntarily delay enrollment past 6.92 years are more likely to come from a high-income family, less likely to come from mid- or low-income families, and more likely to have college-educated parents compared to students in column (2). The same children who delay enrollment are also less likely to be retained in first grade, more likely to be male, to be enrolled in a private school in their first grade, and to have higher scores in all college admission tests. They also presumably differ in unobservable dimensions that affect outcomes, thus biasing enrollment age effects based on Equation 1.

V Effects on college application/enrollment

Figure 1 illustrates the relationship between the day of birth and two outcomes of interest. Panel A of Figure 1 shows the relationship between first-grade enrollment age and the day of birth. The points show means within day-of-birth cells, and they reveal a strong increase in enrollment age on the first days of April, May, June, and July. The first three discontinuities (April, May, and June) represent increases in the enrollment age of approximately 0.1 years in each case, while the one around July causes an increase in enrollment age of nearly 0.5 years. Panels B and C of

Figure 1 explore whether these sharp discontinuities in first-grade enrollment age are joined by discontinuities in the likelihood of either applying or enrolling into college, respectively. Students born on July 1^{st} or after are more likely to apply and enroll in college by approximately 2%. Tables 3 and 5 summarize the estimates of the first stage, reduced form, and 2SLS for the likelihood of enrolling into college, mirroring the results displayed on Panel C of Figure 1.

The first set of results, presented in Tables 2 and 3, include separate estimates of the first stage presented in Equation 3, and the reduced form of Equation 4 using college application and enrollment as the outcome, respectively. Columns (1) to (3) of Tables 2 and 3 demonstrate the strong relationship between first-grade enrollment age and birth date, even when adding year and day of the week fixed effects, as well as controlling for parents' characteristics. These results confirm that the day of birth is effective in predicting the first-grade enrollment age, crucial for the validity of our instrumental variable approach. Columns (4) to (6) of Tables 2 and 3 show an increase in the likelihood of both applying and enrolling into college only for students born on or after July. These estimates work as a reference point for our 2SLS estimates, highlighting the potential bias in OLS estimates from the reduced form.

The results of our 2SLS estimations are reported in Tables 4 and 5. As before, each column represents regressions using different sets of control variables, while different panels reflect different sets of excluded instruments. Panel A shows the results of conducting the 2SLS estimation using all four excluded instruments (D_{april}, D_{may}, D_{june}, D_{july}) to predict first-grade enrollment age. They suggest that a one-year delay in enrollment increases the probability of both applying and enrolling in college by 4.8 and 4.5%, respectively. In the over-identified regression, Hansen J tests produce p-values above 0.4 for both cases, providing no evidence against the instruments' validity. Durbin-Wu-Hausman tests reject the null hypothesis that enrollment age is exogenous. Panels B to E show exactly identified models using only one cutoff as the single excluded instrument. For both studied outcomes, the results that rely on the variation generated around May 1^{st} and June 1^{st} have a positive sign with larger coefficients and standard errors than the ones in Panel A. The estimates relying on the variation caused around April 1^{st} have a negative sign, but even larger standard errors. As depicted in Tables 4 and 5, July 1^{st} provides the strongest instrument, and therefore the estimates in Panel E are very similar to estimates in Panel A.

Birth dates and schools cutoffs

In Chile, like many other countries, scheduled births often result in fewer births on weekends and holidays (Borrescio-Higa & Valdés, 2019; McEwan & Shapiro, 2008). This pattern is evident in our data as well. Figure 2a, which shows birth data for 1997, clearly illustrates a decrease in births on weekends. This pattern is not observed When looking at a pooled sample across multiple years, as shown in Figure 2b. Here, we also observe a noticeable decline in births on three major national holidays, two of which are close to enrollment cutoffs. A further analysis of our sample indicates that parents of children born on Sundays are, on average, 6% less likely to hold a college degree compared to parents of children born on Mondays, and 5.5% less likely to be a high-income family. Similarly, parents of children born on holidays are 3% less likely to hold a college degree compared to parents of children born on other days. We address concerns that could arise from these facts by controlling for day of the week/holidays fixed effects, which do not significantly change our estimates in Tables 4, and 5.

In Chile, evidence shows a strong correlation between parents' SES and scheduled C-sections in the private system (Borrescio-Higa & Valdés, 2019), which raises a more critical concern of whether parents deliberately time births around enrollment cutoffs to influence school starting age (Shigeoka, 2015). Figure 2b shows the birth histogram for the entire sample and does not indicate any significant changes in birth date density around the enrollment cutoffs, suggesting that parents are unlikely to schedule births after the cutoff dates purposefully.

To further validate our results, we examined whether students' family characteristics vary sharply around the enrollment cutoffs. [8] Figure 3 plots average student characteristics by day-of-birth cells and shows no substantial breaks near the cutoffs nor significant seasonal patterns.

[8] Due to the nature of our data, the gender of the student and the type of school in which the student enrolls in first grade are to be considered baseline characteristics. Both family income and parents' education are collected from a survey before taking the college admission tests

Robustness and Heterogeneity

We analyze the robustness of the estimated effects by replicating the estimations performed using 2SLS with a more contemporary approach, the Regression Discontinuity (RD) design (Cattaneo et al., 2019). Panel A of Tables 6 and 7 present the same results obtained in Panel A of Tables 4 and 5, respectively. Panels B to E present the RD estimations around each enrollment cutoff.[9]

Table 6 shows results that are consistent in direction with those in Table 4 but the coefficients for each individual cutoff are notably smaller. Our RD estimates indicate that delaying first-grade enrollment by one year increases the probability of applying to college by 2.4%, approximately half the effect size estimated using the 2SLS approach. We find that the coefficients for the June and May cutoffs are positive but not statistically significant, the coefficient for the April cutoff is negative and not statistically significant, and the July 1^{st} cutoff is the only one showing significant positive effects. These results persist even after including controls for day-of-the-week and birth-month, as well as parents' characteristics.

Table 7 presents results for college enrollment, revealing significant differences from those observed in Table 5. The coefficient for the April cutoff, although now positive, remains not statistically significant. The coefficient for the May cutoff changes from positive to negative when including controls for day-of-the-week and birth-month, but it remains close to zero. The most noteworthy findings are associated with the July 1^{st} cutoff. For the initial specifications, the coefficient remains positive and statistically significant, although nearly half the size of the original 2SLS coefficient. However, after accounting for parents' characteristics, the coefficient approaches zero and loses its statistical significance.

These observations suggest that the initial positive effect of delaying school enrollment on college

[9] Specifically, we estimate the following equation:

$$O = \beta_0 + \beta_1 D_j + f(B - B_j)^k + \beta_2' D_{-j} + \beta_3' F + \epsilon \qquad \forall j$$

where D_j represents each discontinuity, $f(B - B_j)$ a local polynomial fit of order k on both sides of the discontinuity, D_{-j} all other discontinuities, F day-of-the-week, and birth-moth fixed effects.

enrollment is significantly influenced by family income and parental education. When these factors are controlled for, the effect diminishes, indicating that socioeconomic background plays a crucial role in driving the observed outcomes.

VI Effects on performance

Historical High-school performance

Figure 4 presents the relationship between the day of birth and high school performance outcomes. Figure 5a shows the relationship between first-grade enrollment age and the day of birth, demonstrating the same sharp discontinuities in enrollment age around the beginning of April, May, June, and July as seen previously. Figures 4b and 4c explore whether these discontinuities are mirrored in the students' GPA scores and class ranking, respectively.

For high school GPA, as illustrated in Figure 4b, students born after July 1^{st} exhibit a slight increase in GPA scores, indicating that a delay in first-grade enrollment age might positively impact high school performance. This pattern goes in line with related literature that suggests that relatively older students tend to perform better academically during their high school years, potentially due to their maturity, developed cognitive skills, or other unobserved factors (Black et al., 2011; Grenet, 2011).

Figure 4c shows the relationship between the day of birth and class ranking, revealing a similar trend. Students born on or after July 1^{st} are more likely to hold a higher relative position in their GPA distribution, which aligns with the results observed for GPA scores. These findings highlight the long-term benefits of delayed school enrollment age on academic performance throughout high school.

Tables 10 and 11 summarize the 2SLS and RD estimates for high school GPA and GPA ranking, respectively. The results of our 2SLS estimations are reported in Panel A of both Tables. Using all four excluded instruments (D_{april}, D_{may}, D_{june}, D_{july}) to predict first-grade enrollment age, the 2SLS results suggest that a one-year delay in enrollment increases the high school GPA and GPA ranking by 0.045 and 0.048 standard deviations, respectively. Panels B to E show RD results using each discontinuity as the running variable while controlling for the other cutoffs. The estimates for each discontinuity exhibit similar patterns, with July 1^{st} providing the strongest and most consistent instrument for predicting higher GPA scores and rankings.

Performance in Standardized Tests

Figure 5 illustrates the relationship between the day of birth and standardized test scores, specifically in math and verbal tests. Figures 5b and 5c explore the discontinuities in math and verbal scores, respectively. Figure 5b suggests that students born after July 1st tend to have higher math scores, while Figure 5c shows a similar but less pronounced effect for verbal scores. This difference may reflect varying cognitive demands of math versus verbal tasks or differences in how these skills develop with age.

Tables 12 and 13 summarize the 2SLS and RD estimates for math and verbal test scores, respectively. The results of our 2SLS estimations are reported in Panel A of both Tables. Using all four excluded instruments, the 2SLS results suggest that a one-year delay in enrollment increases math and verbal scores by 0.053 and 0.105 standard deviations, respectively. Panels B to E show RD results using each discontinuity as the running variable while controlling for the other cutoffs. The estimates indicate that the variation around July 1^{st} consistently predicts higher math and verbal scores, although the coefficient is statistically significant only for verbal scores.

Interpreting the effects of relative age on high school grades and standardized test scores requires acknowledging several potential limitations and mechanisms that may influence the results. Firstly, high school grades and standardized test scores are influenced by a multitude of factors beyond relative age. These include, but are not limited to, socioeconomic background, school quality, parental involvement, and individual student motivation. While our analysis controls for several observable characteristics, unobserved factors could still play a significant role. Therefore, the estimated effects should be interpreted as part of a broader context of educational outcomes influenced by diverse and complex mechanisms.

Secondly, the variation in retention and promotion policies across different schools and districts might impact the results. Schools may have differing criteria for advancing students, which could lead to discrepancies in how relative age effects manifest in academic performance and test scores. Recent research shows that when students are assigned a score based on their relative position within their class, they may switch to less demanding high schools in their final years to boost their scores (Concha-Arriagada, 2023). This behavior would be problematic if it were predominantly observed in students born around the cutoffs. However, further examination of our data does not reveal patterns between students transferring schools in their high school years and their date of birth.

Thirdly, peer effects could also play a significant role. Older students within a grade might benefit from social and academic interactions with relatively younger peers, potentially enhancing their performance. Conversely, younger students might face challenges in keeping up with older classmates, affecting their academic outcomes. Additionally, teacher effects could influence results if teachers tend to prefer and positively reinforce older students.

Finally, the long-term effects of relative age might interact with various educational interventions and support programs. For example, schools with robust tutoring and mentoring programs might mitigate the disadvantages faced by younger students, while schools lacking support may see more pronounced relative age effects.

VII Conclusions

In this study, we investigate the effects of delayed school enrollment on college application and enrollment outcomes, as well as high school academic performance and standardized test scores, within the context of the Chilean education system. Our analysis leverages first-grade enrollment age cutoffs to create a quasi-experimental framework, allowing us to address several of the issues described in related literature regarding the validity of instruments used to estimate the impact of delaying school in one year.

Our findings indicate that students who enroll in first grade at an older age are approximately 2% more likely to apply to and enroll in college. These students also achieve higher high school GPA scores by 0.045 standard deviations. Furthermore, while older students perform significantly better on verbal tests, the same effect is not observed for math tests, suggesting that the relative age effect may vary across different academic domains. Specifically, a one-year delay in enrollment does not have a statistically significant effect on math scores and increases verbal scores by 0.05 standard deviations.

Importantly, the relative age effect appears more pronounced among male students, emphasizing the need to consider gender differences when designing educational policies. These results contribute to the ongoing debate on the optimal school entry age and suggest that adjusting school entry cutoff dates could lead to more equitable educational outcomes. Policymakers should consider these findings when developing strategies to support younger students and to mitigate any disadvantages they may face due to being relatively younger than their peers.

However, it is crucial to interpret these findings with caution. High school grades and standardized test scores are influenced by a multitude of factors beyond relative age, such as socioeconomic background, school quality, parental involvement, and individual motivation. Additionally, variations in retention and promotion policies across schools and districts, as well as potential peer and teacher effects, may also play significant roles in shaping educational outcomes.

This research provides valuable insights for policymakers aiming to optimize school entry age policies to enhance educational equity and improve long-term academic and career outcomes for students. Future research should continue to explore the broader implications of school entry age on various aspects of educational and labor market success, with particular attention to strategies that support younger and male students in closing achievement gaps.

References

Angrist, J. D. & Krueger, A. B. (1992). The Effect of Age at School Entry on Educational Attainment: An Application of Instrumental Variables with Moments from Two Samples. *Journal of the American Statistical Association*, 87(418), 328–336. Publisher: Taylor & Francis _eprint: https://www.tandfonline.com/doi/pdf/10.1080/01621459.1992.10475212.

Barua, R. & 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.

Bedard, K. & Dhuey, E. (2006). The Persistence of Early Childhood Maturity: International Evidence of Long-Run Age Effects. *The Quarterly Journal of Economics*, 121(4), 1437–1472. Publisher: Oxford University Press.

Black, S. E., Devereux, P. J., & 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.

Borrescio-Higa, F. & Valdés, N. (2019). Publicly insured caesarean sections in private hospitals: a repeated cross-sectional analysis in Chile. *BMJ Open*, 9(4), e024241. Publisher: British Medical Journal Publishing Group Section: Health economics.

Cascio, E. U. & Schanzenbach, D. W. (2013). The Impacts of Expanding Access to High-Quality Preschool Education.

Cattaneo, M. D., Idrobo, N., & Titiunik, R. (2019). A Practical Introduction to Regression Discontinuity Designs: Foundations. *Elements in Quantitative and Computational Methods for the Social Sciences*. ISBN: 9781108684606 9781108710206 Publisher: Cambridge University Press.

Concha-Arriagada, C. (2023). Should I Stay, or Should I Go? Strategic Responses to Improve College Admission Chances.

Crawford, C., Dearden, L., & Meghir, C. (2010). When You Are Born Matters: The Impact of Date of Birth on Educational Outcomes in England 1. *Institute for Fiscal Studies, IFS Working Papers*.

Datar, A. (2006). The impact of kindergarten entrance age policies on the childcare needs of families. *Journal of Policy Analysis and Management*, 25(1), 129–153. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/pam.20159.

Dobkin, C. & Ferreira, F. (2010). Do school entry laws affect educational attainment and labor market outcomes? *Economics of Education Review*, 29(1), 40–54.

Elder, T. E. & Lubotsky, D. H. (2009). Kindergarten Entrance Age and Children's Achievement: Impacts of State Policies, Family Background, and Peers. *Journal of Human Resources*, 44(3), 641–683. Publisher: University of Wisconsin Press Section: Articles.

Grenet, J. (2011). Academic performance, Educational Trajectories and the Persistence of Date of Birth Effects. Evidence from France.

McCrary, J. & Royer, H. (2011). The Effect of Female Education on Fertility and Infant Health: Evidence from School Entry Policies Using Exact Date of Birth. *American Economic Review*, 101(1), 158–195.

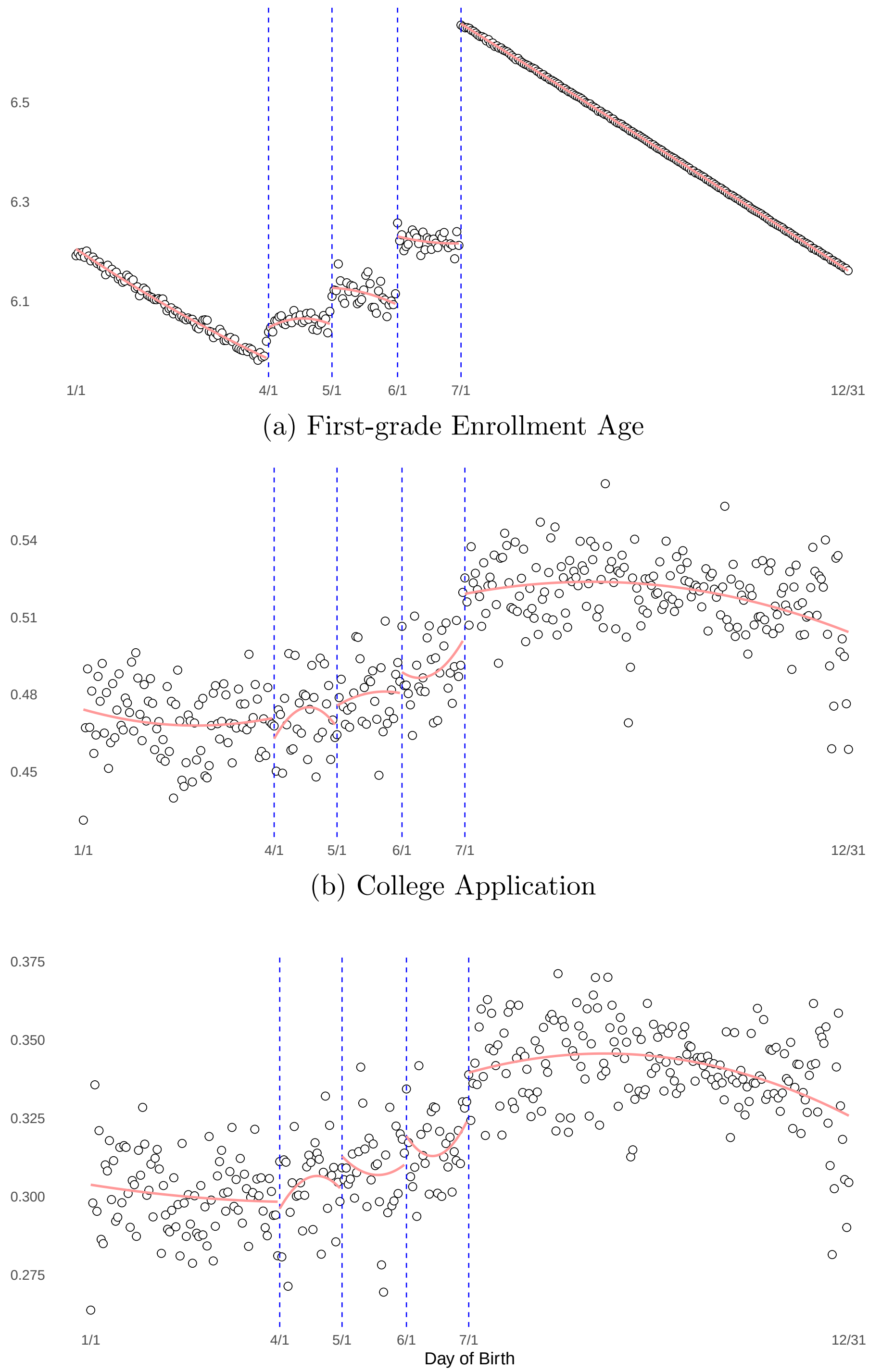
McEwan, P. & Shapiro, J. (2008). The Benefits of Delayed Primary School Enrollment: Discontinuity Estimates Using Exact Birth Dates. *Journal of Human Resources*, 43.

Peña, P. A. (2017). Creating winners and losers: Date of birth, relative age in school, and outcomes in childhood and adulthood. *Economics of Education Review*, 56, 152–176.

Shigeoka, H. (2015). School Entry Cutoff Date and the Timing of Births.

Van Der Klaauw, W. (2002). Estimating the Effect of Financial Aid Offers on College Enrollment: A Regression–Discontinuity Approach*. *International Economic Review*, 43(4), 1249–1287. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/1468-2354.t01-1-00055.

Figure 1: Day of Birth, First-grade Enrollment Age, and College Enrollment.



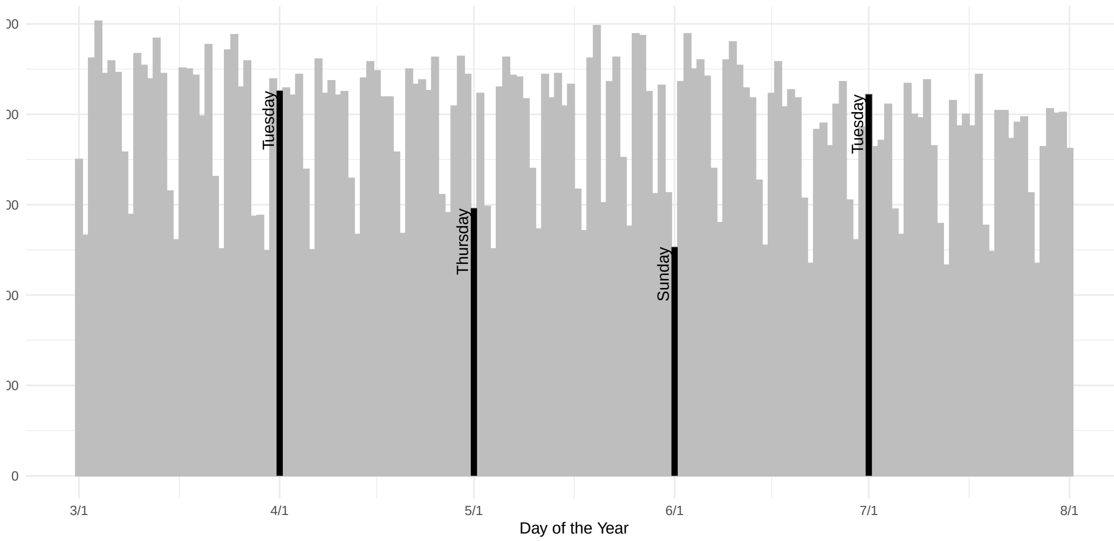
(a) First-grade Enrollment Age

(b) College Application

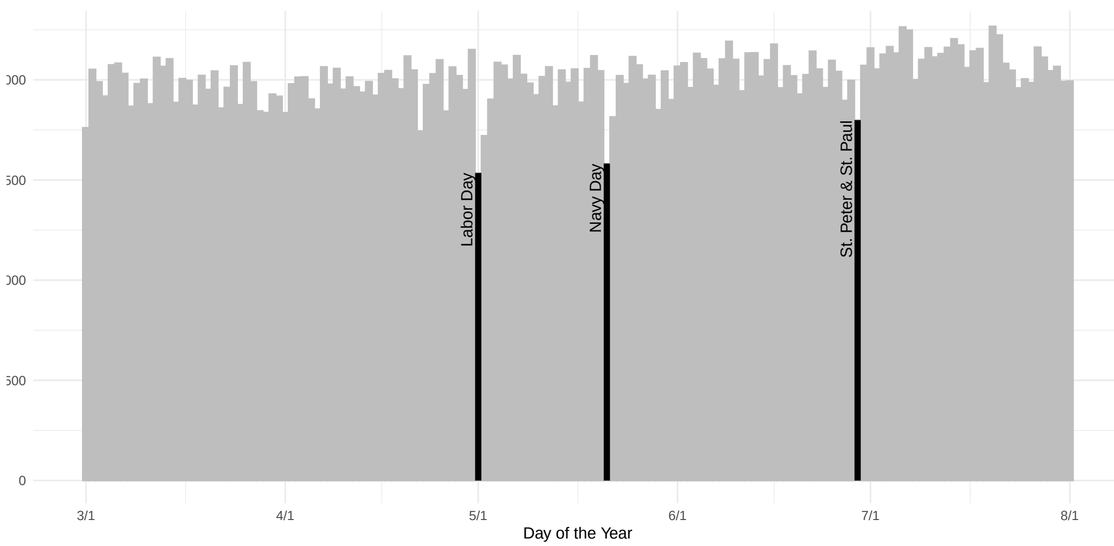
(c) College Enrollment

Note: Sample includes Chilean students born between the years 1994 and 2001, who attended first grade of school between the years 2002 and 2007, and who took the college admission test between the years 2014 and 2019. Dots are mean values of the y-axis variable within each day of birth.

Figure 2: Day of birth histograms



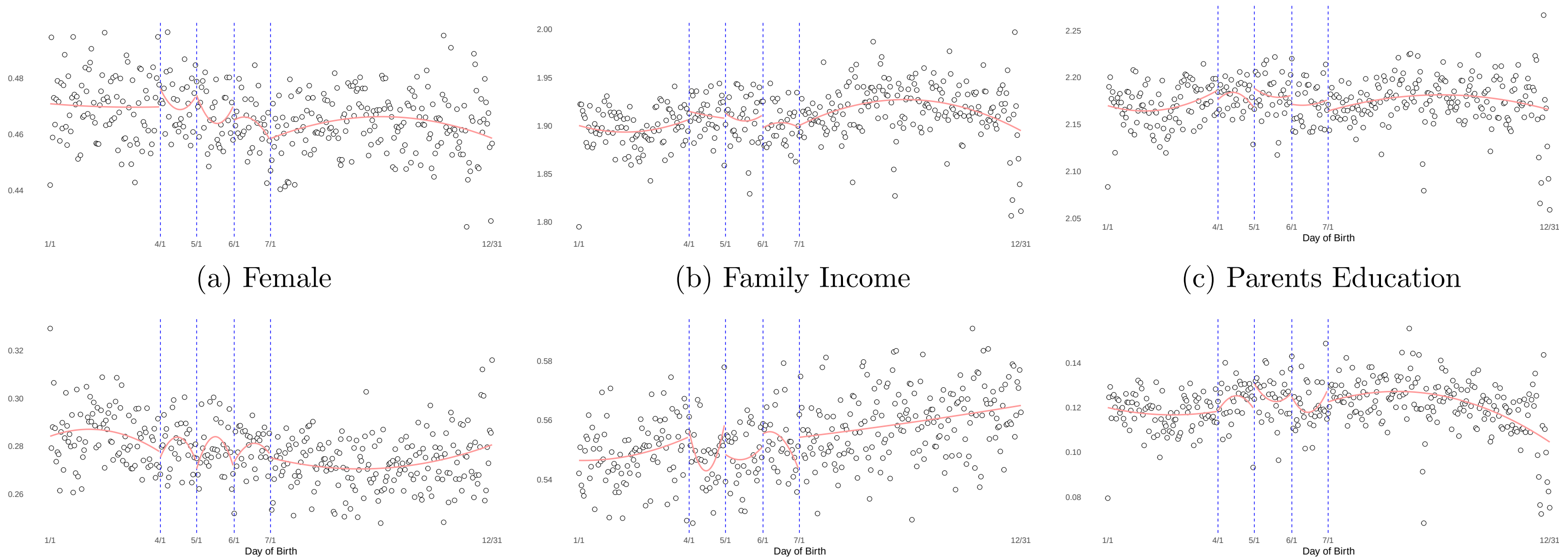
(a) Births 1997



(b) Births pooled across all years

Note: Figure includes all children born between 1995 and 2001. The four labeled bars in panel A represent the days of the week for the four cutoffs (April 1^{st}, May 1^{st}, June 1^{st}, and July 1^{st}) in 1997. The three labeled bars in panel B indicate holidays on May 1^{st} (Labor Day), May 21^{st} (Navy Day), and June 29^{th} (St. Peter and St. Paul).

Figure 3: Students characteristics across enrollment cutoffs.

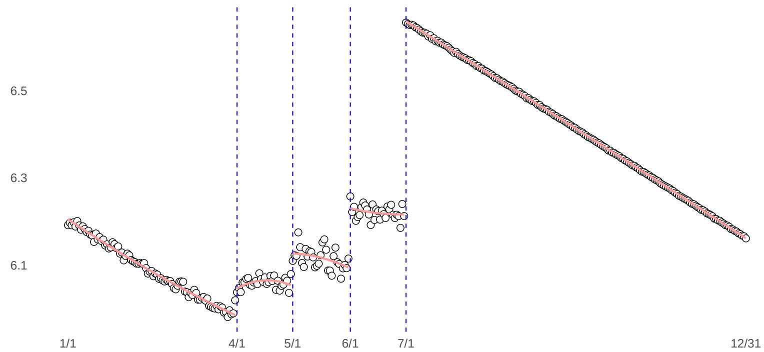


(a) Female (b) Family Income (c) Parents Education

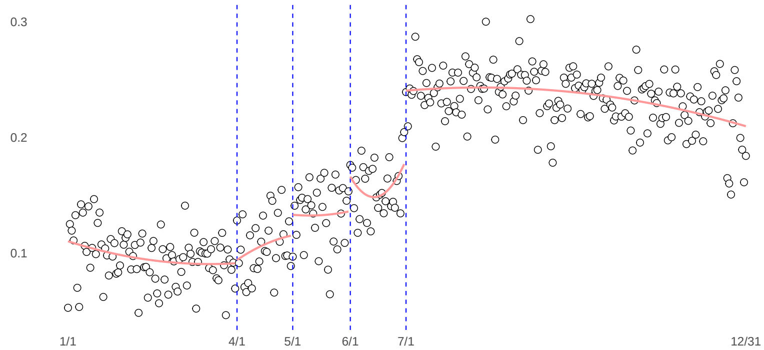
(d) 1^{st} grade public school (e) 1^{st} grade voucher school (f) 1^{st} grade private school

Note: Sample includes Chilean students born between the years 1994 and 2001, who attended first grade of school between the years 2002 and 2007, and who took the college admission test between the years 2014 and 2019. Dots are mean values of the y-axis variable within each day of birth.

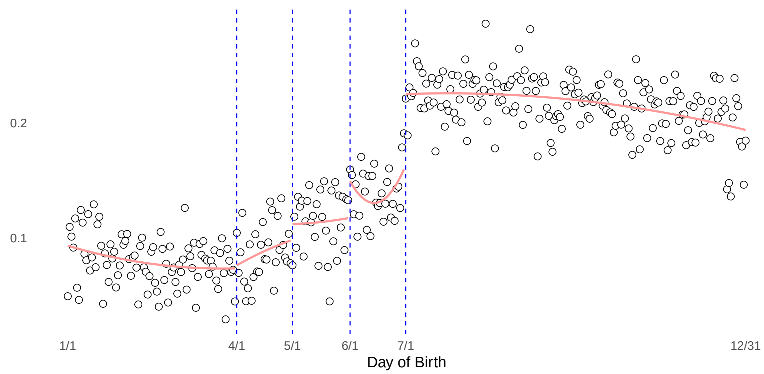
Figure 4: Day of Birth, First-grade Enrollment Age, and High school performance.



(a) First-grade Enrollment Age



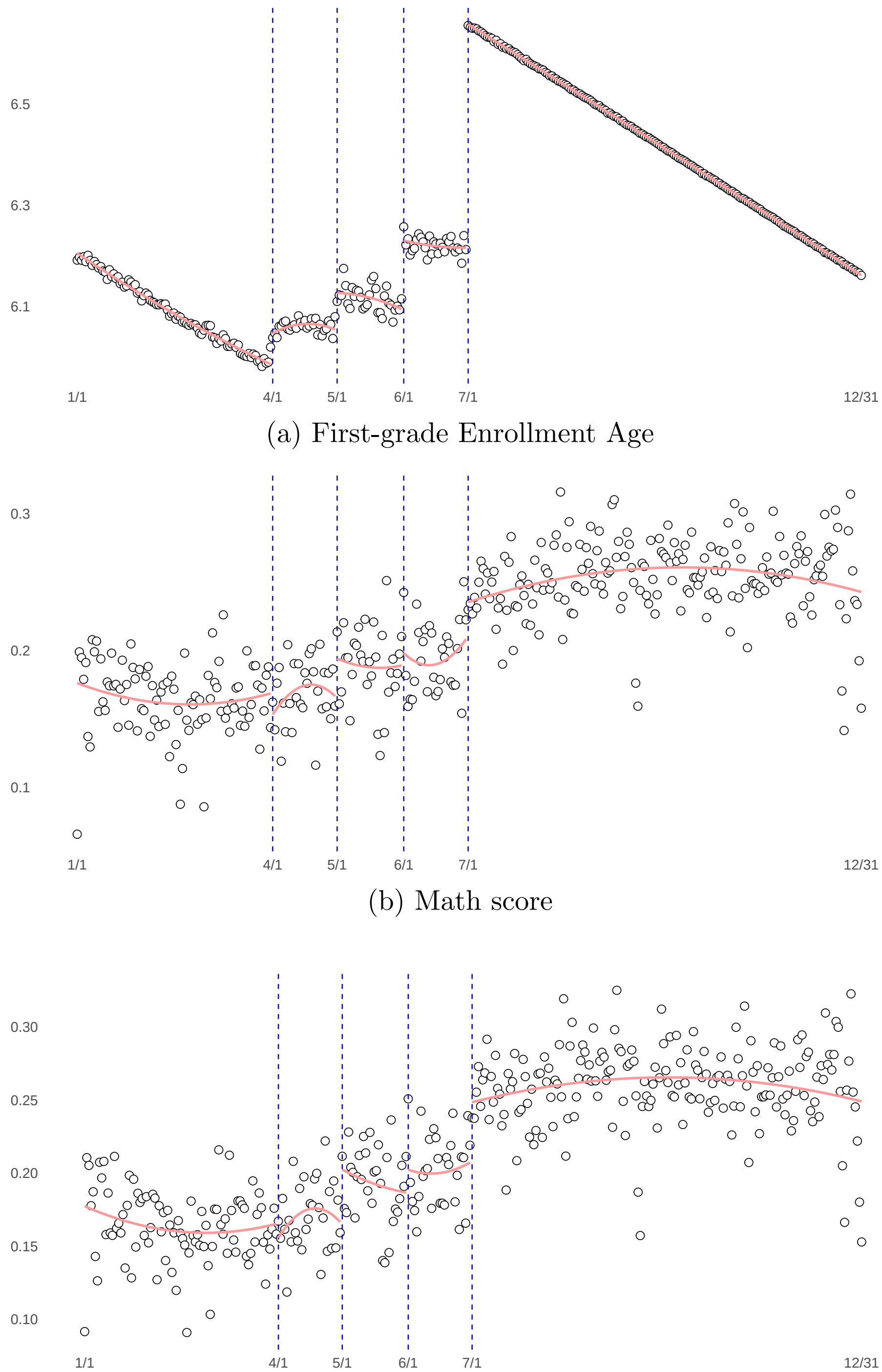
(b) GPA score



(c) GPA rank score

Note: Sample includes Chilean students born between the years 1994 and 2001, who attended first grade of school between the years 2002 and 2007, and who took the college admission test between the years 2014 and 2019. Dots are mean values of the y-axis variable within each day of birth.

Figure 5: Day of Birth, First-grade Enrollment Age, and College admission test scores.



(a) First-grade Enrollment Age

(b) Math score

(c) Verbal score

Note: Sample includes Chilean students born between the years 1994 and 2001, who attended first grade of school between the years 2002 and 2007, and who took the college admission test between the years 2014 and 2019. Dots are mean values of the y-axis variable within each day of birth.

Table 1: Descriptive statistics by birthday segment

	First-grade enrollment age			
	< 5.67	$[5.67 ; 6.67[$	$[6.67 ; 6.92[$	≥ 6.92
Panel A: Students Characteristics				
Male (1=Yes)	0.431**	0.458	0.497***	0.554***
	(0.495)	(0.498)	(0.5)	(0.497)
Middle-income	0.46	0.468	0.397***	0.16***
	(0.499)	(0.499)	(0.489)	(0.367)
High-income	0.227	0.212	0.354***	0.74***
	(0.419)	(0.409)	(0.478)	(0.439)
Parents secondary education	0.473**	0.505	0.406***	0.164***
	(0.499)	(0.5)	(0.491)	(0.37)
Parents college education	0.374***	0.327	0.456***	0.779***
	(0.484)	(0.469)	(0.498)	(0.415)
Panel B: First-grade enrollment				
Retained in first grade	0.08***	0.013	0.007***	0.009***
	(0.272)	(0.111)	(0.081)	(0.094)
Public school	0.341**	0.371	0.278***	0.087***
	(0.474)	(0.483)	(0.448)	(0.281)
Private school	0.196***	0.088	0.25***	0.742***
	(0.397)	(0.284)	(0.433)	(0.438)
Panel C: College-admission test				
Age at college-admission test	17.641***	18.05	18.508***	18.77***
	(0.353)	(0.322)	(0.102)	(0.112)
Math Score	0.258**	0.201	0.378***	0.859***
	(0.893)	(0.901)	(0.93)	(0.899)
Verbal Score	0.251*	0.21	0.369***	0.744***
	(0.853)	(0.892)	(0.908)	(0.878)
GPA Score	0.189	0.175	0.319***	0.665***
	(0.873)	(0.87)	(0.896)	(0.881)
Applies to college	0.511	0.496	0.57***	0.762***
	(0.5)	(0.5)	(0.495)	(0.426)
Enrol into college	0.345**	0.317	0.399***	0.647***
	(0.475)	(0.465)	(0.49)	(0.478)
Observations	1308	629822	54169	7948
Share of sample	0.002	0.909	0.078	0.011

Note: Sample includes students who attended first grade of school between years 2002 and 2007, and took the college-admission test between years 2013 and 2018. Each row reports variable means by four first-grade enrollment age categories. Stars indicate that a row mean is statistically different from the mean of students in second category.

Table 2: Effects of enrollment age on applying to college (first-stage and reduced-form regressions)

	Dependent variable: First grade enrollment age			Dependent variable: applying to college		
	(1)	(2)	(3)	(4)	(5)	(6)
D_{april}	0.063***	0.059***	0.057***	-0.009	-0.008	-0.006
	(0.004)	(0.004)	(0.004)	(0.007)	(0.007)	(0.007)
D_{may}	0.077***	0.073***	0.075***	0.008	0.006	0.004
	(0.008)	(0.007)	(0.007)	(0.009)	(0.009)	(0.009)
D_{june}	0.138***	0.128***	0.127***	0.008	0.008	0.012
	(0.008)	(0.008)	(0.008)	(0.009)	(0.009)	(0.008)
D_{july}	0.441***	0.417***	0.422***	0.017**	0.020***	0.020***
	(0.006)	(0.005)	(0.006)	(0.007)	(0.007)	(0.007)
Observations	759990	759990	697177	759990	759990	697177
F-statistic	323836.2	82043.9	60650.4	-	-	-
P-value	0	0	0	-	-	-
Piecewise quadratic of B	Yes	Yes	Yes	Yes	Yes	Yes
Birth-year dummies	No	Yes	Yes	No	Yes	Yes
Day-of-week dummies	No	Yes	Yes	No	Yes	Yes
Student control variables	No	No	Yes	No	No	Yes

Note: Robust standard errors, adjusted for clustering within 366 day-of-birth cells, appear in parentheses. The F-statistic corresponds to a test of the null hypothesis that the four instruments, D_1 to D_4, are jointly zero (p-values underneath). Student control variables include Female, dummy variables indicating discrete categories of family income and dummy variables indicating discrete categories of parents' schooling. All regressions include a constant. * p<0.1, ** p<0.05, *** p<0.01.

Table 3: Effects of enrollment age on enrolling into college (first-stage and reduced-form regressions)

	Dependent variable: First grade enrollment age			Dependent variable: enrolling into college		
	(1)	(2)	(3)	(4)	(5)	(6)
D_{april}	0.063***	0.059***	0.057***	-0.003	-0.002	-0.001
	(0.004)	(0.004)	(0.004)	(0.006)	(0.006)	(0.006)
D_{may}	0.077***	0.073***	0.075***	0.010	0.007	0.007
	(0.008)	(0.007)	(0.007)	(0.008)	(0.008)	(0.008)
D_{june}	0.138***	0.128***	0.127***	0.008	0.010	0.012
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
D_{july}	0.441***	0.417***	0.422***	0.015**	0.016**	0.016***
	(0.006)	(0.005)	(0.006)	(0.006)	(0.006)	(0.006)
Observations	759990	759990	697177	759990	759990	697177
F-statistic	323836.2	82043.9	60650.4	-	-	-
P-value	0	0	0	-	-	-
Piecewise quadratic of B	Yes	Yes	Yes	Yes	Yes	Yes
Birth-year dummies	No	Yes	Yes	No	Yes	Yes
Day-of-week dummies	No	Yes	Yes	No	Yes	Yes
Student control variables	No	No	Yes	No	No	Yes

Note: Robust standard errors, adjusted for clustering within 366 day-of-birth cells, appear in parentheses. The F-statistic corresponds to a test of the null hypothesis that the four instruments, D_1 to D_4, are jointly zero (p-values underneath). Student control variables include Female, dummy variables indicating discrete categories of family income and dummy variables indicating discrete categories of parents' schooling. All regressions include a constant. * p<0.1, ** p<0.05, *** p<0.01.

Table 4: Effects of enrollment age on applying to college (TSLS regressions)

	Dependent variable: applying to college		
	(1)	(2)	(3)
Panel A: Excluded instruments: $\mathbf{D_{april}, D_{may}, D_{june}, D_{july}}$			
First grade enrollment age	0.039***	0.046***	0.048***
	(0.014)	(0.015)	(0.014)
χ^2 (Hansen J test)	2.849	2.593	2.123
p-value	0.415	0.459	0.547
χ^2 (Durbin-Wu-Hausman)	48.843	47.783	3.933
p-value	0	0	0.047
Panel B: Excluded instruments: $\mathbf{D_{april}}$ (born on or after April 1)			
First grade enrollment age	-0.136	-0.135	-0.104
	(0.111)	(0.118)	(0.120)
Panel C: Excluded instruments: $\mathbf{D_{may}}$ (born on or after May 1)			
First grade enrollment age	0.108	0.085	0.059
	(0.113)	(0.120)	(0.116)
Panel D: Excluded instruments: $\mathbf{D_{june}}$ (born on or after June 1)			
First grade enrollment age	0.060	0.066	0.091
	(0.063)	(0.067)	(0.067)
Panel E: Excluded instruments: $\mathbf{D_{july}}$ (born on or after July 1)			
First grade enrollment age	0.039**	0.047***	0.047***
	(0.015)	(0.016)	(0.016)
Observations	759990	759990	697177
Piecewise quadratic of B	Yes	Yes	Yes
Birth-year dummies	No	Yes	Yes
Day-of-week dummies	No	Yes	Yes
Student control variables	No	No	Yes

Note: Robust standard errors, adjusted for clustering within 366 day-of-birth cells, appear in parentheses. Cells in Columns 1, 2, and 3 report the coefficient on First grade enrollment age from separate TSLS regressions that use different excluded instruments (see panel headings) and controls (see bottom of table). Student control variables include Female, dummy variables indicating discrete categories of family income and dummy variables indicating discrete categories of parents' schooling. All regressions include a constant. * p<0.1, ** p<0.05, *** p<0.01.

Table 5: Effects of enrollment age on enrolling into college (TSLS regressions)

	Dependent variable: enrolling into college		
	(1)	(2)	(3)
Panel A: Excluded instruments: $\mathbf{D_{april}, D_{may}, D_{june}, D_{july}}$			
First grade enrollment age	0.038***	0.042***	0.045***
	(0.013)	(0.014)	(0.013)
χ^2 (Hansen J test)	1.572	1.163	1.265
p-value	0.666	0.762	0.737
χ^2 (Durbin-Wu-Hausman)	61.918	60.693	9.538
p-value	0	0	0.002
Panel B: Excluded instruments: $\mathbf{D_{april}}$ (born on or after April 1)			
First grade enrollment age	-0.045	-0.042	-0.018
	(0.100)	(0.107)	(0.109)
Panel C: Excluded instruments: $\mathbf{D_{may}}$ (born on or after May 1)			
First grade enrollment age	0.134	0.101	0.091
	(0.106)	(0.113)	(0.109)
Panel D: Excluded instruments: $\mathbf{D_{june}}$ (born on or after June 1)			
First grade enrollment age	0.062	0.076	0.096
	(0.059)	(0.063)	(0.063)
Panel E: Excluded instruments: $\mathbf{D_{july}}$ (born on or after July 1)			
First grade enrollment age	0.035**	0.038**	0.039***
	(0.014)	(0.015)	(0.015)
Observations	759990	759990	697177
Piecewise quadratic of B	Yes	Yes	Yes
Birth-year dummies	No	Yes	Yes
Day-of-week dummies	No	Yes	Yes
Student control variables	No	No	Yes

Note: Robust standard errors, adjusted for clustering within 366 day-of-birth cells, appear in parentheses. Cells in Columns 1, 2, and 3 report the coefficient on First grade enrollment age from separate TSLS regressions that use different excluded instruments (see panel headings) and controls (see bottom of table). Student control variables include Female, dummy variables indicating discrete categories of family income and dummy variables indicating discrete categories of parents' schooling. All regressions include a constant. * p<0.1, ** p<0.05, *** p<0.01.

Table 6: Effects of enrollment age on applying to college (RD regressions)

	Dependent variable: applying to college		
	(1)	(2)	(3)
Panel A: IV Excluded instruments: $\mathbf{D_{april}, D_{may}, D_{june}, D_{july}}$			
First grade enrollment age	0.039***	0.046***	0.048***
	(0.014)	(0.015)	(0.014)
χ^2 (Hansen J test)	2.849	2.593	2.123
p-value	0.415	0.459	0.547
χ^2 (Durbin-Wu-Hausman)	48.843	47.783	3.933
p-value	0	0	0.047
Panel B: RD estimate around $\mathbf{D_{april}}$ (born on or after April 1)			
First grade enrollment age	-0.003	-0.004	-0.002
	(0.009)	(0.009)	(0.008)
Panel C: RD estimate around $\mathbf{D_{may}}$ (born on or after May 1)			
First grade enrollment age	0.003	-0.001	-0.001
	(0.008)	(0.007)	(0.006)
Panel D: RD estimate around $\mathbf{D_{june}}$ (born on or after June 1)			
First grade enrollment age	0.002	0.002	0.004
	(0.007)	(0.007)	(0.005)
Panel E: RD estimate around $\mathbf{D_{july}}$ (born on or after July 1)			
First grade enrollment age	0.022***	0.024***	0.024***
	(0.008)	(0.008)	(0.008)
Observations	759990	759990	697177
Piecewise quadratic of B	Yes	Yes	Yes
Birth-year dummies	No	Yes	Yes
Day-of-week dummies	No	Yes	Yes
Student control variables	No	No	Yes

Note: Robust standard errors, adjusted for clustering within 366 day-of-birth cells, appear in parentheses. Cells in Columns 1, 2, and 3 report the coefficient on First grade enrollment age from separate regressions that use different controls (see bottom of table). Panel A shows 2SLS estimates using all excluded instruments, Panel B-E show RD estimates for each of the cutoffs in our data following Cattaneo et al. (2019). Student control variables include Female, dummy variables indicating discrete categories of family income and dummy variables indicating discrete categories of parents' schooling. * p<0.1, ** p<0.05, *** p<0.01.

Table 7: Effects of enrollment age on enrolling into college (RD regressions)

	Dependent variable: enrolling into college		
	(1)	(2)	(3)
Panel A: IV Excluded instruments: $\mathbf{D_{april}, D_{may}, D_{june}, D_{july}}$			
First grade enrollment age	0.038***	0.042***	0.045***
	(0.013)	(0.014)	(0.013)
χ^2 (Hansen J test)	1.572	1.163	1.265
p-value	0.666	0.762	0.737
χ^2 (Durbin-Wu-Hausman)	61.918	60.693	9.538
p-value	0	0	0.002
Panel B: RD estimate around $\mathbf{D_{april}}$ (born on or after April 1)			
First grade enrollment age	0.009	0.006	0.010
	(0.008)	(0.008)	(0.008)
Panel C: RD estimate around $\mathbf{D_{may}}$ (born on or after May 1)			
First grade enrollment age	0.003	-0.002	-0.001
	(0.007)	(0.007)	(0.006)
Panel D: RD estimate around $\mathbf{D_{june}}$ (born on or after June 1)			
First grade enrollment age	0.005	0.005	0.007
	(0.007)	(0.007)	(0.005)
Panel E: RD estimate around $\mathbf{D_{july}}$ (born on or after July 1)			
First grade enrollment age	0.021***	0.021***	0.004
	(0.007)	(0.007)	(0.009)
Observations	759990	759990	697177
Piecewise quadratic of B	Yes	Yes	Yes
Birth-year dummies	No	Yes	Yes
Day-of-week dummies	No	Yes	Yes
Student control variables	No	No	Yes

Note: Robust standard errors, adjusted for clustering within 366 day-of-birth cells, appear in parentheses. Cells in Columns 1, 2, and 3 report the coefficient on First grade enrollment age from separate regressions that use different controls (see bottom of table). Panel A shows 2SLS estimates using all excluded instruments, Panel B-E show RD estimates for each of the cutoffs in our data following Cattaneo et al. (2019). Student control variables include Female, dummy variables indicating discrete categories of family income and dummy variables indicating discrete categories of parents' schooling. * p<0.1, ** p<0.05, *** p<0.01.

Table 8: Heterogeneous effects of enrollment age on the probability of applying to college (RD regressions)

	Female			Male		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: IV Excluded instruments: $\mathbf{D_{april}, D_{may}, D_{june}, D_{july}}$						
First grade enrollment age	0.041**	0.051***	0.050***	0.038*	0.041*	0.046**
	(0.019)	(0.020)	(0.019)	(0.021)	(0.022)	(0.022)
χ^2 (Hansen J test)	7.053	6.837	6.263	5.800	5.747	6.099
p-value	0.070	0.077	0.100	0.122	0.125	0.107
χ^2 (Durbin-Wu-Hausman)	23.649	22.249	1.031	33.657	34.443	3.106
p-value	0	0	0.310	0	0	0.078
Panel B: RD estimate around $\mathbf{D_{april}}$ (born on or after April 1)						
First grade enrollment age	-0.012	-0.014	-0.014	0.010	0.008	0.016
	(0.010)	(0.010)	(0.009)	(0.012)	(0.012)	(0.011)
Panel C: RD estimate around $\mathbf{D_{may}}$ (born on or after May 1)						
First grade enrollment age	0.005	0.001	0.001	0.002	-0.003	-0.001
	(0.010)	(0.010)	(0.009)	(0.010)	(0.010)	(0.009)
Panel D: RD estimate around $\mathbf{D_{june}}$ (born on or after June 1)						
First grade enrollment age	-0.013	-0.013	-0.019**	0.012	0.018*	0.012
	(0.010)	(0.009)	(0.008)	(0.010)	(0.010)	(0.008)
Panel E: RD estimate around $\mathbf{D_{july}}$ (born on or after July 1)						
First grade enrollment age	0.031***	0.035***	0.030***	-0.014	-0.011	-1.031***
	(0.010)	(0.010)	(0.009)	(0.014)	(0.014)	(0.010)
Observations	405768	405768	374946	354222	354222	322231
Piecewise quadratic of B	Yes	Yes	Yes	Yes	Yes	Yes
Birth-year dummies	No	Yes	Yes	No	Yes	Yes
Day-of-week dummies	No	Yes	Yes	No	Yes	Yes
Student control variables	No	No	Yes	No	No	Yes

Note: Robust standard errors, adjusted for clustering within 366 day-of-birth cells, appear in parentheses. Cells in Columns 1, 2, and 3 report the coefficient on First grade enrollment age from separate regressions that use different controls (see bottom of table). Panel A shows 2SLS estimates using all excluded instruments, Panel B-E show RD estimates for each of the cutoffs in our data following Cattaneo et al. (2019). Student control variables include Female, dummy variables indicating discrete categories of family income and dummy variables indicating discrete categories of parents' schooling. * p<0.1, ** p<0.05, *** p<0.01.

Table 9: Heterogeneous effects of enrollment age on the probability of enrolling into college (RD regressions)

	Female			Male		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: IV Excluded instruments: $\mathbf{D_{april}, D_{may}, D_{june}, D_{july}}$						
First grade enrollment age	0.036**	0.041**	0.039**	0.043**	0.045**	0.052**
	(0.017)	(0.018)	(0.018)	(0.020)	(0.021)	(0.021)
χ^2 (Hansen J test)	6.273	5.432	5.325	3.245	3.520	4.351
p-value	0.099	0.143	0.149	0.355	0.318	0.226
χ^2 (Durbin-Wu-Hausman)	36.031	34.339	6.430	33.346	33.977	3.364
p-value	0	0	0.011	0	0	0.067
Panel B: RD estimate around $\mathbf{D_{april}}$ (born on or after April 1)						
First grade enrollment age	-0.005	-0.009	-0.009	0.026**	0.026**	0.035***
	(0.010)	(0.010)	(0.008)	(0.012)	(0.012)	(0.012)
Panel C: RD estimate around $\mathbf{D_{may}}$ (born on or after May 1)						
First grade enrollment age	0.005	0.000	-0.001	-0.001	-0.008	-0.029***
	(0.009)	(0.009)	(0.007)	(0.011)	(0.011)	(0.010)
Panel D: RD estimate around $\mathbf{D_{june}}$ (born on or after June 1)						
First grade enrollment age	0.003	-0.765***	0.003	0.009	0.007	0.010
	(0.008)	(0.008)	(0.006)	(0.008)	(0.008)	(0.006)
Panel E: RD estimate around $\mathbf{D_{july}}$ (born on or after July 1)						
First grade enrollment age	0.016	-0.165***	0.016**	-0.007	-0.005	-0.162***
	(0.011)	(0.010)	(0.008)	(0.013)	(0.013)	(0.009)
Observations	405768	405768	374946	354222	354222	322231
Piecewise quadratic of B	Yes	Yes	Yes	Yes	Yes	Yes
Birth-year dummies	No	Yes	Yes	No	Yes	Yes
Day-of-week dummies	No	Yes	Yes	No	Yes	Yes
Student control variables	No	No	Yes	No	No	Yes

Note: Robust standard errors, adjusted for clustering within 366 day-of-birth cells, appear in parentheses. Cells in Columns 1, 2, and 3 report the coefficient on First grade enrollment age from separate regressions that use different controls (see bottom of table). Panel A shows 2SLS estimates using all excluded instruments, Panel B-E show RD estimates for each of the cutoffs in our data following Cattaneo et al. (2019). Student control variables include Female, dummy variables indicating discrete categories of family income and dummy variables indicating discrete categories of parents' schooling. * p<0.1, ** p<0.05, *** p<0.01.

Table 10: Effects of enrollment age on high school GPA (RD regressions)

	Dependent variable: high school GPA		
	(1)	(2)	(3)
Panel A: IV Excluded instruments: $\mathbf{D_{april}, D_{may}, D_{june}, D_{july}}$			
First grade enrollment age	0.155***	0.176***	0.190***
	(0.022)	(0.024)	(0.023)
χ^2 (Hansen J test)	0.467	0.269	0.435
p-value	0.926	0.966	0.933
χ^2 (Durbin-Wu-Hausman)	32.694	34.869	2.734
p-value	0	0	0.098
Panel B: RD estimate around $\mathbf{D_{april}}$ (born on or after April 1)			
First grade enrollment age	0.019	0.013	0.018
	(0.015)	(0.015)	(0.016)
Panel C: RD estimate around $\mathbf{D_{may}}$ (born on or after May 1)			
First grade enrollment age	0.011	0.003	0.019*
	(0.013)	(0.013)	(0.011)
Panel D: RD estimate around $\mathbf{D_{june}}$ (born on or after June 1)			
First grade enrollment age	0.017	-0.009	0.022**
	(0.012)	(0.012)	(0.010)
Panel E: RD estimate around $\mathbf{D_{july}}$ (born on or after July 1)			
First grade enrollment age	0.046***	0.051***	0.080***
	(0.017)	(0.016)	(0.011)
Observations	753318	753318	691291
Piecewise quadratic of B	Yes	Yes	Yes
Birth-year dummies	No	Yes	Yes
Day-of-week dummies	No	Yes	Yes
Student control variables	No	No	Yes

Note: Robust standard errors, adjusted for clustering within 366 day-of-birth cells, appear in parentheses. Cells in Columns 1, 2, and 3 report the coefficient on First grade enrollment age from separate regressions that use different controls (see bottom of table). Panel A shows 2SLS estimates using all excluded instruments, Panel B-E show RD estimates for each of the cutoffs in our data following Cattaneo et al. (2019). Student control variables include Female, dummy variables indicating discrete categories of family income and dummy variables indicating discrete categories of parents' schooling. * p<0.1, ** p<0.05, *** p<0.01.

Table 11: Effects of enrollment age on GPA ranking (RD regressions)

	Dependent variable: GPA ranking		
	(1)	(2)	(3)
Panel A: IV Excluded instruments: $\mathbf{D_{april}, D_{may}, D_{june}, D_{july}}$			
First grade enrollment age	0.146***	0.161***	0.165***
	(0.025)	(0.026)	(0.027)
χ^2 (Hansen J test)	0.974	1.188	1.670
p-value	0.807	0.756	0.644
χ^2 (Durbin-Wu-Hausman)	22.335	23.645	2.195
p-value	0	0	0.139
Panel B: RD estimate around $\mathbf{D_{april}}$ (born on or after April 1)			
First grade enrollment age	0.019	0.014	0.019
	(0.015)	(0.015)	(0.016)
Panel C: RD estimate around $\mathbf{D_{may}}$ (born on or after May 1)			
First grade enrollment age	0.008	0.000	-0.000
	(0.013)	(0.013)	(0.011)
Panel D: RD estimate around $\mathbf{D_{june}}$ (born on or after June 1)			
First grade enrollment age	0.025**	0.035***	0.019*
	(0.011)	(0.011)	(0.010)
Panel E: RD estimate around $\mathbf{D_{july}}$ (born on or after July 1)			
First grade enrollment age	0.043**	0.048***	0.055***
	(0.017)	(0.017)	(0.015)
Observations	759990	759990	697177
Piecewise quadratic of B	Yes	Yes	Yes
Birth-year dummies	No	Yes	Yes
Day-of-week dummies	No	Yes	Yes
Student control variables	No	No	Yes

Note: Robust standard errors, adjusted for clustering within 366 day-of-birth cells, appear in parentheses. Cells in Columns 1, 2, and 3 report the coefficient on First grade enrollment age from separate regressions that use different controls (see bottom of table). Panel A shows 2SLS estimates using all excluded instruments, Panel B-E show RD estimates for each of the cutoffs in our data following Cattaneo et al. (2019). Student control variables include Female, dummy variables indicating discrete categories of family income and dummy variables indicating discrete categories of parents' schooling. * p<0.1, ** p<0.05, *** p<0.01.

Table 12: Effects of enrollment age on math score (RD regressions)

	Dependent variable: math score		
	(1)	(2)	(3)
Panel A: IV Excluded instruments: $\mathbf{D_{april}, D_{may}, D_{june}, D_{july}}$			
First grade enrollment age	0.051**	0.042	0.053**
	(0.025)	(0.027)	(0.025)
χ^2 (Hansen J test)	3.760	3.916	5.101
p-value	0.289	0.271	0.165
χ^2 (Durbin-Wu-Hausman)	83.682	85.045	14.778
p-value	0	0	0
Panel B: RD estimate around $\mathbf{D_{april}}$ (born on or after April 1)			
First grade enrollment age	-0.007	-0.015	-0.085***
	(0.017)	(0.016)	(0.014)
Panel C: RD estimate around $\mathbf{D_{may}}$ (born on or after May 1)			
First grade enrollment age	0.018	0.013	-0.043***
	(0.015)	(0.014)	(0.010)
Panel D: RD estimate around $\mathbf{D_{june}}$ (born on or after June 1)			
First grade enrollment age	-0.009	0.003	0.224***
	(0.015)	(0.014)	(0.010)
Panel E: RD estimate around $\mathbf{D_{july}}$ (born on or after July 1)			
First grade enrollment age	0.033**	0.022	0.013
	(0.015)	(0.016)	(0.015)
Observations	751344	751344	689533
Piecewise quadratic of B	Yes	Yes	Yes
Birth-year dummies	No	Yes	Yes
Day-of-week dummies	No	Yes	Yes
Student control variables	No	No	Yes

Note: Robust standard errors, adjusted for clustering within 366 day-of-birth cells, appear in parentheses. Cells in Columns 1, 2, and 3 report the coefficient on First grade enrollment age from separate regressions that use different controls (see bottom of table). Panel A shows 2SLS estimates using all excluded instruments, Panel B-E show RD estimates for each of the cutoffs in our data following Cattaneo et al. (2019). Student control variables include Female, dummy variables indicating discrete categories of family income and dummy variables indicating discrete categories of parents' schooling. * p<0.1, ** p<0.05, *** p<0.01.

Table 13: Effects of enrollment age on verbal score (RD regressions)

	Dependent variable: verbal score		
	(1)	(2)	(3)
Panel A: IV Excluded instruments: $\mathbf{D_{april}, D_{may}, D_{june}, D_{july}}$			
First grade enrollment age	0.099***	0.100***	0.105***
	(0.025)	(0.026)	(0.025)
χ^2 (Hansen J test)	3.804	3.704	4.836
p-value	0.283	0.295	0.184
χ^2 (Durbin-Wu-Hausman)	51.619	52.710	1.805
p-value	0	0	0.179
Panel B: RD estimate around $\mathbf{D_{april}}$ (born on or after April 1)			
First grade enrollment age	0.006	-0.001	-0.209***
	(0.016)	(0.016)	(0.013)
Panel C: RD estimate around $\mathbf{D_{may}}$ (born on or after May 1)			
First grade enrollment age	0.029**	-0.226***	0.035***
	(0.014)	(0.014)	(0.010)
Panel D: RD estimate around $\mathbf{D_{june}}$ (born on or after June 1)			
First grade enrollment age	0.009	-0.031**	0.023**
	(0.014)	(0.014)	(0.009)
Panel E: RD estimate around $\mathbf{D_{july}}$ (born on or after July 1)			
First grade enrollment age	0.052***	0.049***	0.038***
	(0.015)	(0.015)	(0.014)
Observations	752059	752059	690171
Piecewise quadratic of B	Yes	Yes	Yes
Birth-year dummies	No	Yes	Yes
Day-of-week dummies	No	Yes	Yes
Student control variables	No	No	Yes

Note: Robust standard errors, adjusted for clustering within 366 day-of-birth cells, appear in parentheses. Cells in Columns 1, 2, and 3 report the coefficient on First grade enrollment age from separate regressions that use different controls (see bottom of table). Panel A shows 2SLS estimates using all excluded instruments, Panel B-E show RD estimates for each of the cutoffs in our data following Cattaneo et al. (2019). Student control variables include Female, dummy variables indicating discrete categories of family income and dummy variables indicating discrete categories of parents' schooling. * p<0.1, ** p<0.05, *** p<0.01.