UNIVERSIDAD DEL CEMA
Buenos Aires
Argentina

Serie
DOCUMENTOS DE TRABAJO

Área: Economía

SCHOOL STARTING AGE AND THE IMPACT ON SCHOOL ADMISSION

Julio Cáceres-Delpiano y Eugenio Giolito

Noviembre 2022
Nro. 840

https://ucema.edu.ar/publicaciones/doc_trabajo.php
UCEMA: Av. Córdoba 374, C1054AAP Buenos Aires, Argentina
ISSN 1668-4575 (impreso), ISSN 1668-4583 (en línea)
Editor: Jorge M. Streb; asistente editorial: Valeria Dowding ved@ucema.edu.ar

School Starting Age and the impact on School Admission

Julio Cáceres-Delpiano*

jcaceres@eco.uc3m.es

Universidad Carlos III de Madrid

Eugenio P. Giolito

egiolito@ucema.edu.ar

Universidad del CEMA, Buenos Aires and IZA.

November 14, 2022

Abstract

Using administrative data for Chile, we study the impact of School Starting Age (SSA) on the characteristics of the school of first enrollment. After addressing the usual concerns of endogeneity using minimum age requirements and an RD-design, we uncover gains associated with a delay of school entry at the start of the student's school life. SSA is associated with an enrollment in a school with an approximately 0.1 standard deviations higher average in standardized test scores, an increase of approximately 0.17 years in the average education of the peers' parents, and an increase of 4 percentage points in the probability of being enrolled in a private school. The heterogeneity analysis by parents' education reveals the largest gain in the probability of enrollment in a voucher school among less-educated families. We also show that the impact on school's standardized test scores occurs among girls. This heterogeneity by parents' education and student's gender differs from that reported in previous studies.

JEL: A21, I24, I25, and I28

Keywords: Latin America; Chile; Early Entry; Schools' characteristics

*Cáceres-Delpiano gratefully acknowledges financial support from the Spanish Ministry of Education (Grant ECO2009-11165 and ECO2019-00419-001), the Spanish Ministry of Economy and Competitiveness (MDM 2014-0431), and the Comunidad de Madrid MadEco-CM (S2015/HUM-3444). Errors are ours. The views of the authors do not necessarily represent the position of the Universidad del Cema.

1 Introduction

We study the impact of the School Starting Age (SSA) on the characteristics of the school of first enrollment using public administrative data for Chile. To address the potential correlation of SSA with unobserved factors, we use minimum age requirement rules that create a discontinuous jump in the probability of delaying school entry as the source of variation in the analysis.

Since Deming and Dynarski (2008) document a trend for the US in delaying school entry age,[1] several studies have explored the short and long-term effects of SSA on several dimensions with mixed results.[2] For children in primary/middle school, on the one hand, SSA has been negatively associated with grade retention (Elder and Lubotsky, 2008; Caceres-Delpiano and Giolito, 2019), mental health problems (Dee and Sievertsen, 2018), ADD/ADHD diagnoses (Elder and Lubotsky 2008), and the probability of receiving special education (Dhuey and Lipscomb, 2010). Moreover at these early stages, SSA has been positively linked to test scores (McEwan and Shapiro, 2006; Elder and Lubotsky 2008; Attar and Cohen-Zada, 2018) and the probability of following an academic-oriented track (Mühlenweg and Puhani, 2010).

The evidence on the impact of SSA on long-run outcomes is less conclusive, on the other hand. Angrist and Kruger (1991) show a negative correlation between SSA and schooling and earnings among young individuals in the US. Cascio and Lewis (2006) find that SSA has an insignificant effect on the Army Forces Qualifying Test. Black et al. (2011) show that, despite a negative effect on IQ and earnings for young adults in Norway, SSA reduces the likelihood of teen pregnancy and is associated with better mental health among boys. Moreover, Fredriksson and Öckert (2014) find that, in Sweden, even though SSA leaves prime-age earnings on average unchanged, a positive effect is observed among individuals with less-educated parents. In the crime literature, Cook and Kang (2016) show that, despite a positive effect on academic achievements, SSA increases the probability of dropping out and the likelihood of committing a felony offense by age 19. On the contrary, Landersø et al. (2016) find that SSA decreases the likelihood to commit a crime at a younger age in Denmark. In general, this positive effect of delaying school entry in early outcomes that wear off over time with an ambiguous long-term effect is also present in different institutional settings and countries (Elder and Lubotsky, 2009; Caceres-Delpiano and Giolito, 2019).

Behind the heterogeneous impacts reported in the literature that seem to wear off over time, SSA may affect educational achievement through different channels. First, holding back a student is associated with a higher degree of maturity to learn (school readiness). Second, delaying school entry means that a student will be evaluated at an older age than students who started earlier (age

[1] As mentioned by the authors, one-fourth of this change is explained by legal changes, while the rest is attributed to families, teachers, and schools.

[2] A related part of the literature has focused on the impact of age rather than SSA. Among these studies, we find Bedard and Dhuey (2006), who show that younger children obtain considerably lower scores than older ones at fourth and at eighth grades for a sample of OECD countries.

at test). It is difficult to set apart these two channels from each other by comparing children in the same grade (Black et al., 2011), even though the effect of school readiness should wear off over time. Third, the combination of minimum age entry rules and rules on mandatory education may affect dropout decisions (Angrist and Kruger, 1991), which might offset initial gains in delaying school entry. Fourth, a late start in school implies delaying entry into the labor market, and consequently the accumulation of labor experience (Black et al., 2011). Finally, recent evidence shows that SSA may alter other margins at the household level, such as family structure, mother's labor participation, or the achievements of the other siblings, with a consequent potential indirect effect on the student's educational performance (Landersø et al., 2020).

In this paper, we study the impact of SSA on the characteristics of the school of first enrollment, that is, we explore another channel by which age of entry can alter a student's achievements. In this paper, we try to put together two central issues in the literature: school choice and SSA. By exploring the impact of SSA on school characteristics, we provide evidence about the interaction of school entry rules and school choice. In a revision of the literature, Epple et al. (2017) point out that the heterogeneity in the design of several voucher programs is responsible for the mixed findings and the lack of a clear position in favor or against school choice among researchers and policymakers. Our paper shows the role of other policies/institutions, like the age of entry rule, as a determinant in the ability of families to profit from school choice. Moreover, by studying the impact of SSA on school characteristics, we shed light on school choice as an investment process. Specifically, by looking at the age of entry as an investment margin in the hands of the household, this paper, on the one hand, stresses student's circumstances as a key factor behind educational achievements, a well-known fact since the Coleman Report. On the other hand, it helps to understand some of the limits of school choice: when the cost of searching for a school differs between families, they might have a different willingness to delay school entry. Therefore, our paper speaks to literature linking educational policies, specifically school entry rules, with family investments in human capital.

Our findings reveal that delaying school entry induces a relevant shift in the opportunity to start in a school perceived as "better". First, students who delay entry enroll in schools whose average standardized test scores are 0.1 standard deviations higher (measured a year before enrollment). Second, students who delay school entry face peers whose parents have 0.17 more years of education. Third, a delay in enrollment increases the probability of enrolling in a private school (funded with or without a voucher) by 4 percentage points (8%). Our heterogeneity analysis by parents' education and gender reveals higher impacts for children of less-educated parents and girls over boys.

Complementary to other channels in the literature, our findings also suggest that age of entry is associated with a different starting point in a student's educational life. Elder and Lubotsky (2009) find an early impact of SSA on test scores, particularly on children from upper-income families, which they interpret as a reflection of skill accumulation previous to kindergarten. Our results by education of the parents suggest that lower-income families "use" SSA to invest in their children by enrolling them in better schools. Moreover, the larger effect for girls, despite their relative advantage in terms of school readiness, may suggest a channel embedded in the process of school choice. If the competitive pressure supposedly induced by a voucher scheme translates into a more competitive selective process, we know from recent literature that girls react worse on average to competition (Niederle, 2017; Sutter and Glätzle-Rützler, 2015). With those findings in mind, our results may suggest that girls' delayed entry helps them to access better (and likely more selective) schools.

We also provide evidence that delaying school entry increases the probability of being enrolled in a school with some degree of academic selection, which would explain the increase in the average tests scores of the starting school in standardized tests. We also find that families delaying school entry are more likely to express being constrained by the cost or the availability of the schools in their counties (municipalities), which we interpret as an increase of investment effort from the side of the families.

Finally, when we look at the reasons behind their school selection, we find that families who delay school entry, despite enrolling their children in better schools (in terms of their average tests scores), prioritize school proximity and the prestige and values associated with the school.

The paper is organized as follows. In Section 2, we briefly sketch Chile's educational system. Section 3 describes our empirical strategy. In Section 4 we present the data set used in the analysis, and define the selected outcomes in the analysis. In Section 5 we show evidence supporting the validity of our empirical specification, and Section 6 provides the main results. Finally Section 7 concludes.

2 Institutional background

Since a major educational reform took place in the early 1980s, the Chilean primary and secondary educational system has been characterized by its decentralization and by significant participation of the private sector. [3],[4] Despite mixed evidence on the impact of the voucher system on the quality of education (see, for example Hsieh and Urquiola (2006)), the Chilean educational system is comparable in terms of coverage to those in developed countries. The population of students is approximately 3.5 million, distributed throughout three types of schools: public or municipal (41% of total enrollment), subsidized private (51% of total enrollment), and unsubsidized private schools (7%

[3] The management of primary and secondary education was transferred to counties, payment scales and civil servant protection for teachers were abolished, and a voucher scheme was established as the funding mechanism for municipal and (originally) non- fee-charging private schools. For more details, see Gauri and Vawda (2003).

[4] For a review of these and other reforms since the early 1980s, see Contreras et al. (2005).

of total enrollment).[5] Municipal schools are managed by the *comunas* (counties), while the other two types of schools are controlled by the private sector. Though both municipal and subsidized private schools get state funding through a voucher scheme, the latter are usually called *voucher* schools. Two key characteristics of this system are: i) the amount of the subsidy for municipal and voucher schools depend on the demand (enrollment) for each school and ii) parents were not constrained to choose a school located in their county of residence (even though around 90% of the parents do so). The Chilean system is comparable with other large-scale voucher programs around the world like those currently in place in the Netherlands, Denmark or Sweden, or a number of small-scale programs operating in the USA (see Epple et al. (2017)).

While initially conceived as non-fee-charging schools, since 1994, public and subsidized private schools were allowed to charge tuition and fees (on top of the subsidy). Law 20.845 of 2015 determined that schools charging tuition should gradually decrease the co-payment to zero unless they become unsubsidized private. By that time, despite the fact that most public schools were free of charge, around 30% of students were attending voucher schools with co-payment. Moreover, the law explicitly prevented voucher schools from selecting students and established a Centralized Admission System for public and voucher schools, applied gradually by regions and in place at the national level in 2020.

3 Empirical specification

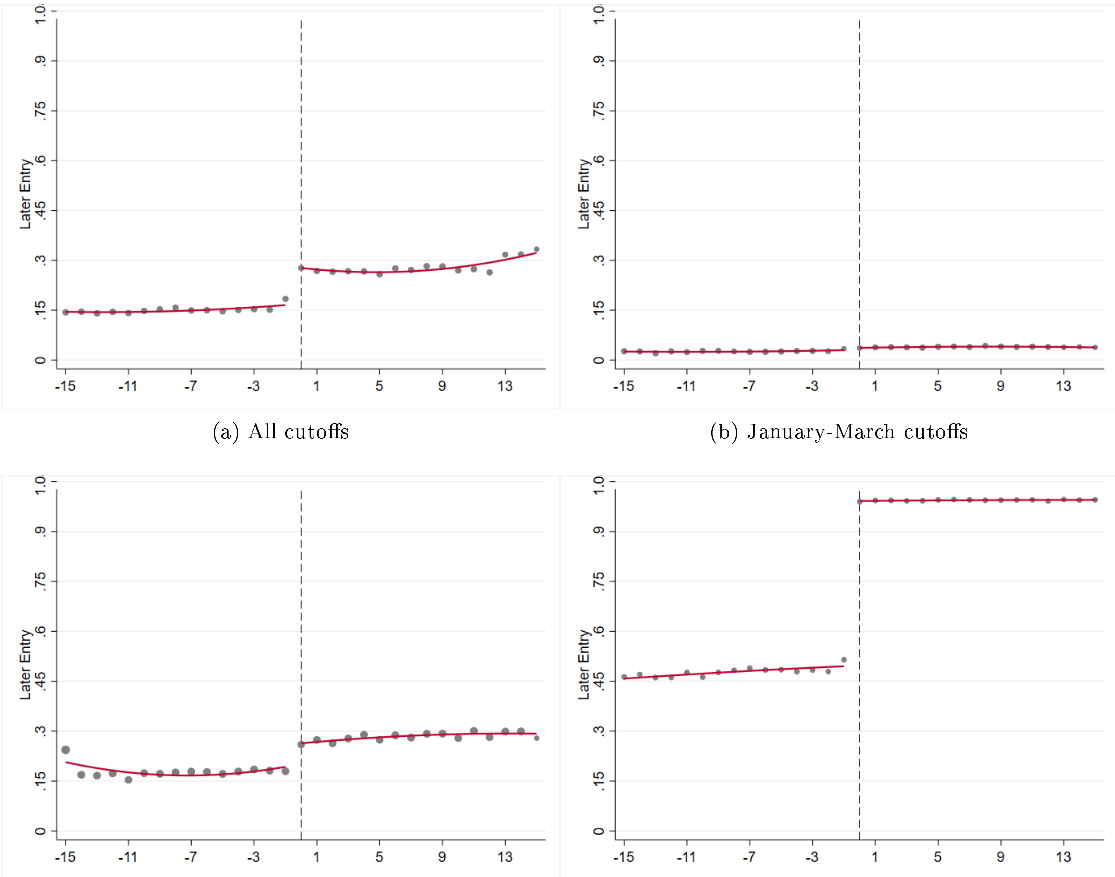
As a consequence of the non-random nature of SSA and researchers' limited information, the estimation of the impact of SSA is non-trivial. To circumvent this problem, and following extensive literature, we use the minimum age requirement rule as a quasi-experimental variation. These rules establish that children, to be enrolled in first grade at primary school in a given school, must have turned six before a given *date* during the academic year.[6] That is, children whose birthday takes place before this cutoff date are entitled to start school the year they turn six. Those whose birthday is after this specific date must wait until the next academic year to start school. Although Chile's official enrollment cutoff was set initially on April 1^{st}, from 1992 until recently, the Ministry of Education has provided schools with some flexibility for setting other dates but never later than July 1^{st}.[7] In practice, schools have distributed themselves over seven cutoffs between January 1^{st} and July 1^{st}.[8] McEwan and Shapiro (2006), to estimate the impact of SSA in Chile, use the four most used cutoffs in practice: April 1^{st}, May 1^{st}, June 1^{st} and July 1^{st}.

[5] There is a fourth type of school, "corporations", which are vocational schools administered by firms or enterprises with a fixed budget from the state. In 2012, they constituted less than 2% of the total enrollment. Throughout our analysis, we treat them as municipal schools.

[6] In Chile, the academic year goes from March to December.

[7] In 2016, the Ministry of Education implemented a Centralized Admission System for Public and Voucher schools with an unified cutoff date.

[8] See http://bcn.cl/1yw2h



(a) All cutoffs

(b) January-March cutoffs

(c) April-June cutoffs

(d) July cutoff

Figure 1: Fraction of students delaying entry around different cutoffs.

Figure 1 presents graphically the source of variation induced by minimum age requirements rules. Each panel of Figure 1 presents the fraction of students who start elementary education a year later than the academic year closest to their sixth birthday ("late entry") within a 15 days window around the different age cutoffs. Panel (a) presents this fraction stacking up all the cutoffs. While approximately 15 percent of the students at the left of any cutoff delay school entry, this fraction approximately doubles at the right of these cutoffs.

Behind this evident discontinuity when stacking up all the cutoffs in one, notice in Panels (b) to (d) that, among those students with birthdays early in the year, only a small fraction delay school entry. Practically none of the students with a birthday before April 1^{st} delay enrollment, but this fraction increases progressively to reach approximately 45 percent among students with a birthday in June. Two elements explain this pattern. First, students born early in the year are older at the start of the academic year, that is, they are probably less likely to postpone enrollment. Second, due to the Chilean institutional setting of multiple cutoffs across schools, children with a birthday closer to July 1^{st} would not only profit from a direct effect of delaying school entry but also from choosing among a larger pool of schools in case of delaying enrollment.[9] These two elements are probably also behind the heterogeneity in the discontinuity induced by minimum age rules. Notice first that those children born after June 30^{th} are "forced" to postpone enrollment until the next year, with approximately a 55 percentage point increase in the likelihood of deferred entry, compared to those born before July $1.^{st}$ Notice that the change in the probability of a delayed entry is less than 1 percentage point for the January-March cutoff (Panel (c)) and 10 percentage points around the April-June (Panel (d)) cutoff. From now on, we will concentrate our analysis on the July cutoff, which presents the largest discontinuity, to avoid any concerns of weak instruments.[10]

The discontinuity in the SSA, together with the assumption that parents cannot fully manipulate the date of birth, has been used in the literature as a quasi-experimental variation in the SSA at the core of a "fuzzy" Regression Discontinuity (RD) strategy.[11] Consider a scenario where students are indexed by i, and birthday over the calendar year by b. The specification we use to estimate the impact of SSA can be expressed as follows,

$$y_{ict}=\rho_t + \rho_c + X_i^{'}\psi + \alpha DelayedEntry_i + g(b_i) + \epsilon_{it} \quad (1)$$

with y_{ict} representing one of the outcomes for student i, born the day b of year t, and living in

[9] In a previous paper (Caceres-Delpiano and Giolito, 2018) we use the distribution of school vacancies across cutoffs and counties to study the effect of variations in the school choice set in educational outcomes.

[10] Nevertheless, we present the results using all the cutoffs in the appendix. The results are robust to considering children born around the other six discontinuities.

[11] In a fuzzy RD design, the probability of treatment (starting primary school at the age of six) changes in magnitude lower than one. On the other hand, in the case where the treatment is a deterministic function of the day of birth, the probability of treatment would change from one to zero at the cut-off day. For more details, see Lee and Lemieux (2010).

municipality c.

The variable $DelayedEntry_i$ is a dummy variable taking a value of one in the case in which a child started primary school one academic year later than that closest to her/his sixth birthday and zero, otherwise. Moreover, ρ_t, ρ_c and X_i represent the year of birth, municipality of residence and X_i a vector of individual covariates, respectively.[12],[13] Finally, $g(b_i)$ is a flexible polynomial specification in the day of birth for a student, taking the form:

$$g\left(b_i\right) = \sum_{k=1}^{K} \left[\beta_k \left(b_i - C\right)^k + \rho_k \left(b_i - C\right)^k \times 1\{b_i - C > 0\}\right],$$

where k is the degree of the polynomial, $1\{*\}$ is an indicator operator and C the July cutoff. That is, $1\{b_i - C > 0\}$ defines whether or not an individual has a birthday after the given cutoff. [14]

Hahn et al. (2001) show that the estimation of causal effects in this regression discontinuity framework is numerically equivalent to an instrumental variable (IV) approach within a small interval around the discontinuity. By focusing on observations around the discontinuity, we concentrate on those observations where we can consider the treatment (SSA) as good as randomly assigned. This randomization of the treatment ensures that all other factors (observed and unobserved) determining a given outcome must be balanced at each side of the discontinuity.

Following Calonico et al. (2014), we select two data-driven bandwidths for each of the outcomes. The first method tries to balance some form of bias-variance trade-off by minimizing the Mean Squared Error (MSE) of the local polynomial RD point estimator (MSE-optimal bandwidth).[15] While this bandwidth is optimal for point estimates, it is not so for inference, because it is not "small" enough to remove the leading bias term to conduct statistical inference. To address this difficulty, we follow the under-smoothing approach, that is, we use more observations for point estimation than for inference. The second bandwidth minimizes an approximation to the coverage error of the confidence intervals, that is, the discrepancy between the empirical coverage of the confidence interval and the theoretical level (CER). We report these selected bandwidths in Table 2. These bandwidths go from 5 to 12 days around the discontinuity, with the CER-optimal bandwidths being consistently smaller.

In equation (1), the impact of SSA is captured by the population parameter α. As stated above, the variable indicating whether or not a child delays school entry ($DelayedEntry_i$) is a non-random variable. To address this endogeneity, we use the discontinuity determining SSA and estimate the

[12] Covariates include fixed effects for parents' education and student's gender. We use as parents' education the highest observed education between the father and the mother reported in parents' survey of the National standardized test (SIMCE). We define as an additional category for the case when the educational level is missing simultaneously for mother and father. We also include six dummy variables indicating the day of the week on which a student was born, and a dummy to define whether or not a student was born on a national holiday.

[13] We cluster the error term at the student's day of birth.

[14] Using Akaike's information criterion (AIC), we get a degree for $g(.)$ that is either one or two depending on the outcome. We use a local linear specification in our preferred specification.

[15] Choosing a smaller bandwidth reduces the bias of the local polynomial approximation, but simultaneously increases the variance of the estimated coefficients because fewer observations will be used for estimation.

following first-stage regression:

$$DelayedEntry_i \quad = \quad \rho_t + \rho_c + X_i^{'}\psi + \delta \times 1\{b_i - C > 0\} + g(b_i) + v_i, \tag{2}$$

where δ, the parameter of interest, captures the discontinuity in the endogenous variable around the cutoff.

By adding to the specification a fully flexible polynomial specification in the day of birth, $g(b_i)$,[16] we deal with the possibility that students born at different dates differ in a systematic manner. [17] Nevertheless, even if the mother's characteristics were correlated simultaneously with the child's birthday and educational outcomes, it does not invalidate our approach. What is required is that the effect of these observed and unobserved factors do not change discontinuously in the mentioned cutoffs.[18],[19]

We must also stress that, in Chile, where schools can sort over the different thresholds, the parameter α captures the whole benefit of delaying school entry, that is, also the one associated with an increase in the number of available schools.

4 Data and variables

The primary data source in our analysis comes from public administrative records provided by the Ministry of Education of Chile for the period 2002-2014. We restrict the analysis to students born between the years 1996 and 2001 to ensure a sample of students who can be observed until they complete primary school, independently of when the effective enrollment took place.[20]

Our second data source is the SIMCE standardized tests records, obtained from the *Agencia de Calidad de la Educación*.[21] We use the individual records to calculate average scores at the school level for the year previous to each cohort entry. We also use the parents' survey from SIMCE tests to, on the one hand, find out the individual school choice process and, on the other hand, to aggregate their opinions regarding school characteristics, specifically on their degree of selectivity.[22]

[16]Given a small interval around the discontinuity and a parametrization of $g(.)$, we can see the estimated function as a non-parametric approximation of the true relationship between a given outcome and the variable *day of birth*. Therefore, we face fewer concerns that the estimated impacts are driven by an incorrect specification of $g(.)$.

[17]Buckles and Hungerman (2013) show, for the United States, that the season of birth is correlated with the mother's characteristics. Specifically, they show that children born in winter are more likely to have a less-educated mother, a teen mother, or an African-American mother.

[18]In a context of "intrinsic heterogeneity" (Heckman et al., 2006), the estimated parameters can be interpreted as weighted "Local Average Treatment Effects" (LATE) across *all* individuals (Lee and Lemieux, 2010). That is, this fuzzy RD design does not estimate the impact just for those individuals around the discontinuity but overall *compliers*. How close this weighted LATE is to the traditional LATE depends on how flat these weights are (Lee and Lemieux, 2010).

[19]Since the last cutoff (July 1^{st}) is associated with practically perfect compliance in the age of entry, by using the last discontinuity the estimation is closer to a sharp RD where the interpretation of the estimated parameter is a weighted "Average Treatment Effects".

[20]For students born before year 1995, for example, we observe them in first grade only in the case of delaying school entry or repeating first grade.

[21]The national standardized test (SIMCE: *Sistema de Medición de la Calidad de la Educación*), is usually given in 4th, 8th and 10th grades.

[22]In addition, we use the parents' survey to obtain parents' level of education.

With this data, we construct, first, six outcome variables to characterize the school where a student is enrolled in first grade of elementary school and the consequence of a potential mismatch due to an early entry.

The first three outcomes are dummy variables indicating if a student is enrolled in a *private* school (voucher or unsubsidized), or specifically in a *voucher* school, or in a school that charges some type of *tuition* (unsubsidized private schools or voucher with co-payment), respectively.

The following two variables correspond to the average education of the parents in the school and the school's average score in a standardized test (SIMCE), measured the year before effective enrollment.

"*Switch school*" dummy variable which takes a value of one if a child is observed in two (or more) different schools for two consecutive years. We define this variable only for students in elementary school since more than 50% of the students change school when they move from elementary to high school.[23] This last variable aims to capture a potential gain for students/families who, by delaying school entry, increase the likelihood to enroll in a school that is a better match, reducing their probability to switch school later on.

Until recent reforms, the process of selecting a school reflected, in many cases, an active role of both family and schools. To understand the mechanism of school choice we use the parents' survey of the SIMCE standardized test to construct seven variables.

First, using parents' answers about the selection process of the schools, we construct two variables at a school level. Each of these two variables correspond to the fraction of parents reporting if their child was subject to a selection process based on academic or family background dimensions, respectively.[24] As well for the average school SIMCE scores, these variables are imputed for each student based on the information measured the year before effective enrollment.

The remaining outcomes are dummy variables regarding the elements driving the family choice of the school. The first one, "*Family constrained in selection*" takes a value of one when families justify the selection on grounds of being an affordable school, or when it is the only school in the county/municipality or proximity, and zero otherwise. The remaining variables in specifically take a value of one when a family selects a school due to proximity, standardized test scores, its environment (values or prestige), or because of other relatives in the school, respectively, and zero otherwise. A potential limitation of these four last variables is the fact that they reflect the reasons behind the

[23]The reason for a forced school switching comes from the fact that many public schools with elementary education do not offer high school education.

[24]We consider the presence of academic selection when a student is required to take an exam or to be subject to a play session, or when the school requests information about previous academic records (for example, whether or not the child had preschool education). We consider that a school selects based on the family background when the selection process consists of a meeting with parents, or when the school requires wedding or baptism certificates, or when parents are asked about earnings. In the construction of these two variables, the questions allow parents to choose multiple selection methods.

school attended when taking the SIMCE examination. [25] Though this is not ideal, we can interpret the parameter as a reduced form estimate of the impact on the reason behind the selection of their current school (which in most cases is the school of their first enrollment).

Table 1 presents the descriptive statistics. The upper panel reports the statistics for the whole student population, while the bottom panel column shows the statistics for students with birthdays within an 8-day interval around July's cutoff. Noticeable is the similarity in means between the students in the working sample and the whole population. Specifically, each cohort of students is composed of approximately 250,000 students, half of which are boys. The average age of entry is 6.14 years of age, with around 21% of these students delaying school entry. This fraction is higher around July's cutoff (Panel B) due to the institutional setting that forces all children with a birthday after June 30th to wait until the next academic year to start elementary education. In terms of the selected outcomes, around 14% of students switch schools during their elementary education.

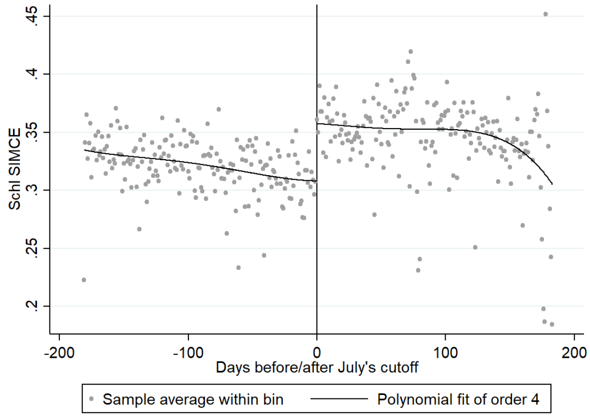
Regarding the type of school, around 8% and 50% of the students are enrolled in a private or voucher school, respectively. The average education of the parents of the classmates is approximately 11 years. In terms of the selection process, around 60% of students are enrolled in schools that selected students based on academic grounds in the past. Moreover, 40% of students are enrolled in schools reportedly selecting based on family backgrounds. Finally, the most common reasons reported by the families for choosing a given school are either the proximity to the residence or having other relatives in the school, with approximately 50% of the families claiming each one as the main reason. Less than 30% of the families/students report the schools' average test scores as the main reason for choosing a given school.

Figures 2 and 3 show the relationship between each of the outcomes and the student's birthday. [26]Notice that the jump at the discontinuity is evident for most of the variables considered. Specifically, we observe that students with a birthday just after the cutoff start on average in schools with higher tests scores and average parents' education. That is, at the descriptive level, delayed enrollment appears to be correlated with an increase in school's "quality." Moreover, students born after the cutoff are more likely to be enrolled in a paid school (unsubsidized or voucher school with co-payment) which is in line with an increase in the probability to pay for education.

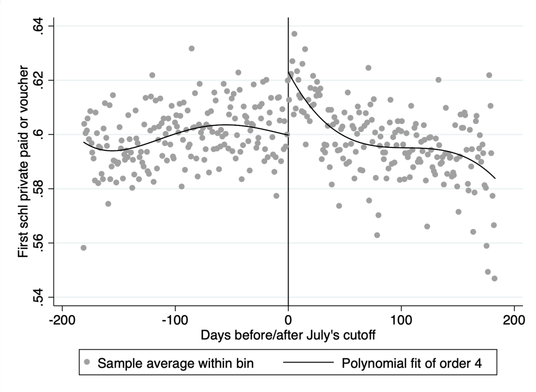
For the outcomes that aim to characterize the driving force behind the school choice, Figure 3 shows, first, that students at the right of the cutoff are more likely to start in an elementary school with some kind of academic selection than those born before the threshold. We also observe that families who have to delay school entry are more likely to report that they choose an elementary school based on average test scores, school environment, or due to having a relative in that school.

[25] Usually, the SIMCE examination is taken in the fourth and eighth grades of primary school.

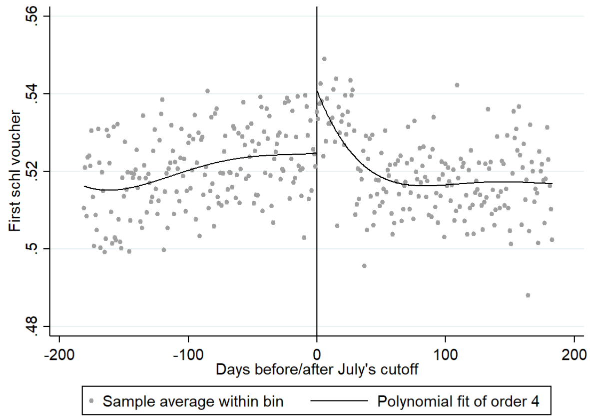
[26] For each outcome, we fit a flexible fourth-degree polynomial at every side of the four discontinuities. Specifically, we use the *rdplot* STATA command for the graphical analysis.



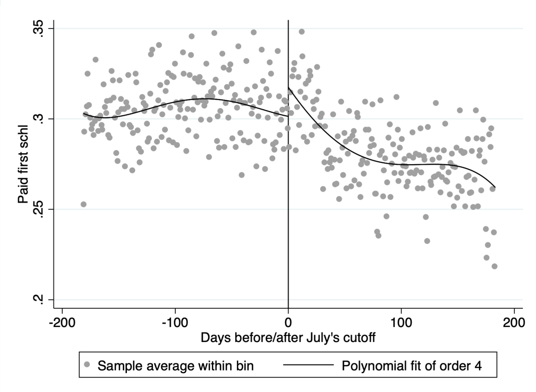
(a) School SIMCE.



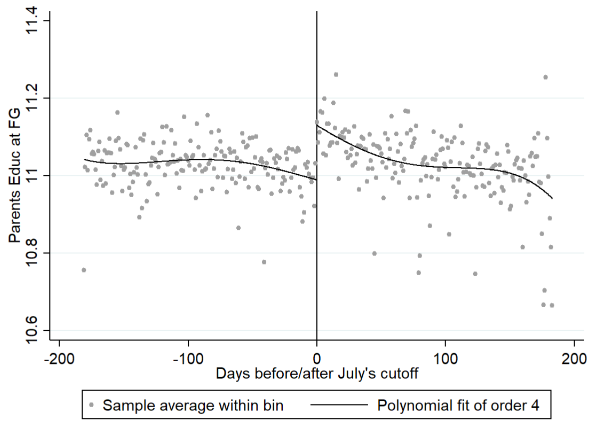
(b) Private or voucher school



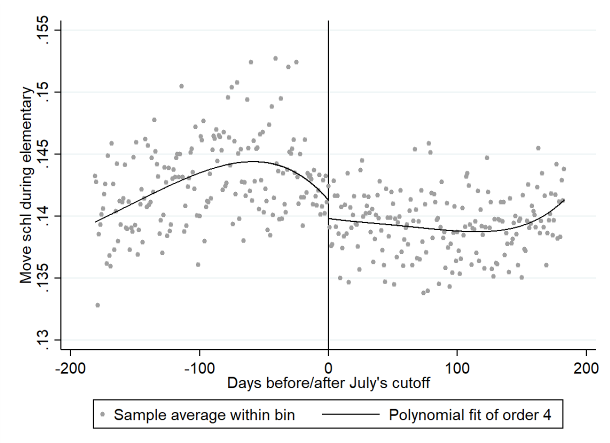
(c) Voucher school



(d) Tuition paid

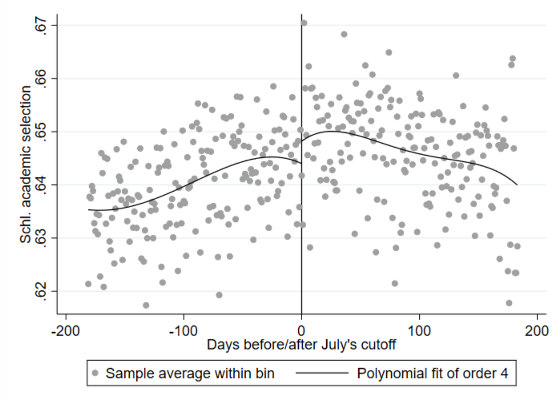


(e) Mean Parents' Schooling

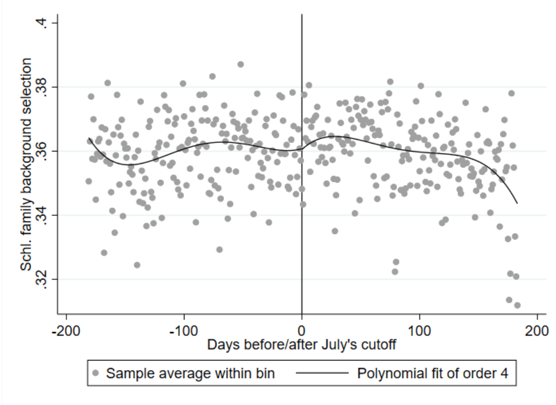


(f) Switch schools

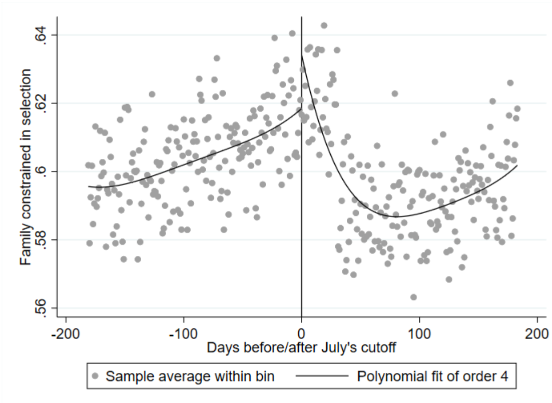
Figure 2: Evolution of selected outcomes according to the student's birthday.



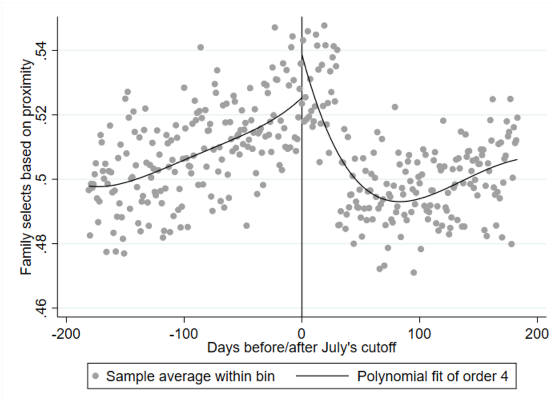
(a) School academic selection



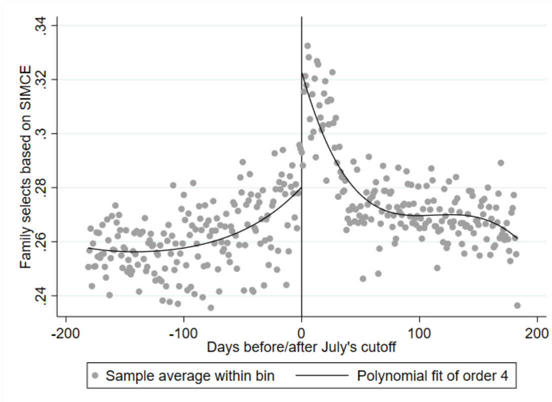
(b) Schl. selects on family background



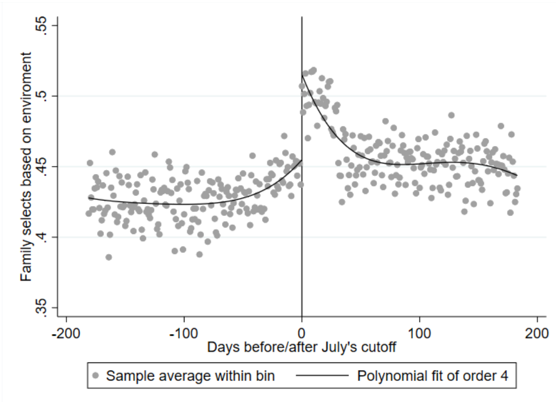
(c) Family constrained in selection



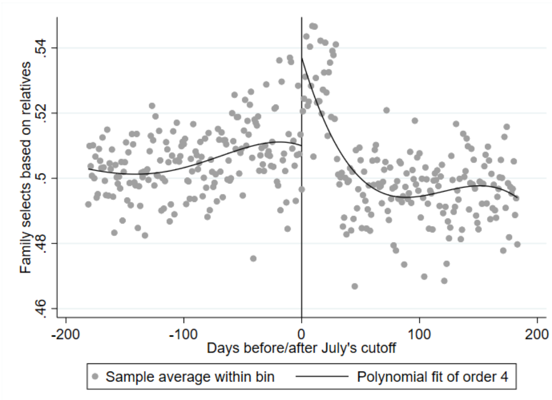
(d) Family selects based on proximity



(e) Family selects based on SIMCE



(f) Family selects based on environment



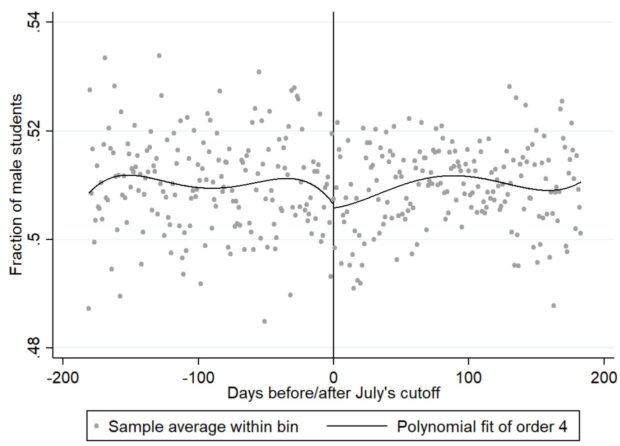
(g) Family selects based on relatives

Figure 3: Evolution of selected outcomes according to the student's birthday.

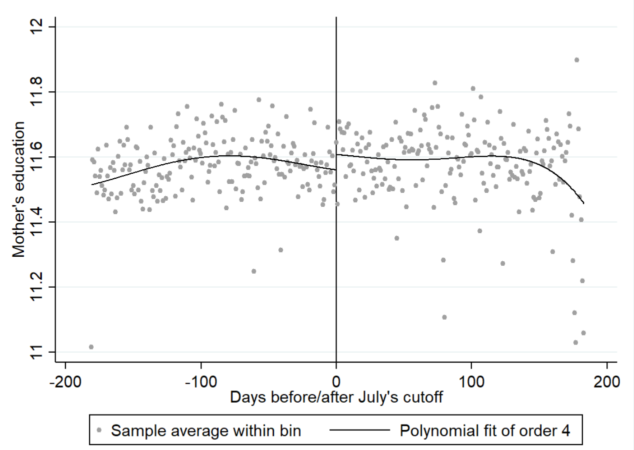
5 Validity of the source of variation

5.1 Continuity in predetermined variables

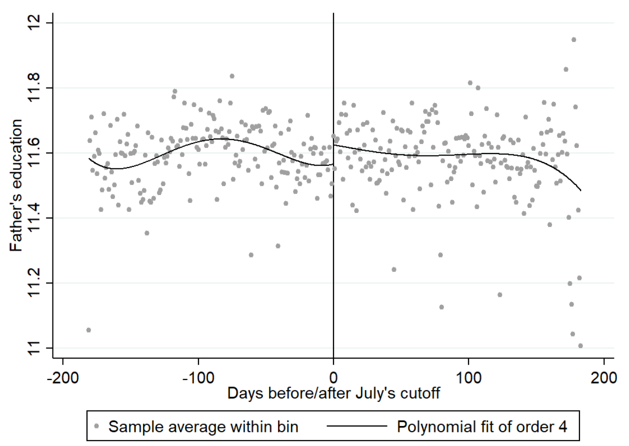
Our analysis builds on the fact of changes in school eligibility for early enrollment around a cutoff being “as good as a randomized assignment” for students whose birthday is close to a cutoff. Then, as in any random assignment, those pre-determined characteristics at the time of the randomization should be similar between the treated (students with a birthday just after one of the seven cutoffs) and the control group (students with a birthday just before one of these cutoffs). Evidence of a systematic difference in these pre-determined characteristics around these dates would compromise the underlying assumption that individuals cannot precisely manipulate the running variable (Lee and Lemieux, 2010).



(a) Fraction of males



(b) Mother’s education



(c) Father’s education

Figure 4: Balancing covariates

Figure 4 inspects graphically the existence of a potential discontinuity among four baseline characteristics available in the data set: gender (fraction of male), mother’s and father’s education, and the highest education of the student's parents. The graphical representation does not show any sizable discontinuity or outliers for these selected variables, compared to those presented in Figure 2. [27] We formally test for discontinuities in the baseline characteristics using different polynomial specifications and three bandwidths (5,10, and 15 days) and report the p-values in Table 3.[28] In Columns (2) and (4) of Table 3, we check for discontinuities around July's cutoff, and in columns (1) and (3), we consider all the cutoffs other than January. For only four out of 36 specifications, we reject the null at a significance level of 5%. Notice that discontinuities are detected only for a quadratic or cubic polynomial, probably due to an overfitting effect in the model.

5.2 Manipulation of the running variable

The randomization of treatment in the neighborhood of the discontinuities rests on the assumption that families cannot *precisely* select their children's date of birth. That is, the validity of an RD design can be compromised in cases in which individuals are able to manipulate the running variable (Lee and Lemieux, 2010). Since minimum-age entry rules are of public knowledge, it may be the case that benefits/costs associated with a delay in the age of entry might induce some families to choose the season of birth. Moreover, it is worth noting that Chile (along with Turkey and Mexico) is one of the countries with the highest rates of c-sections in the world. [29] These two facts suggest some power to select the running variable.

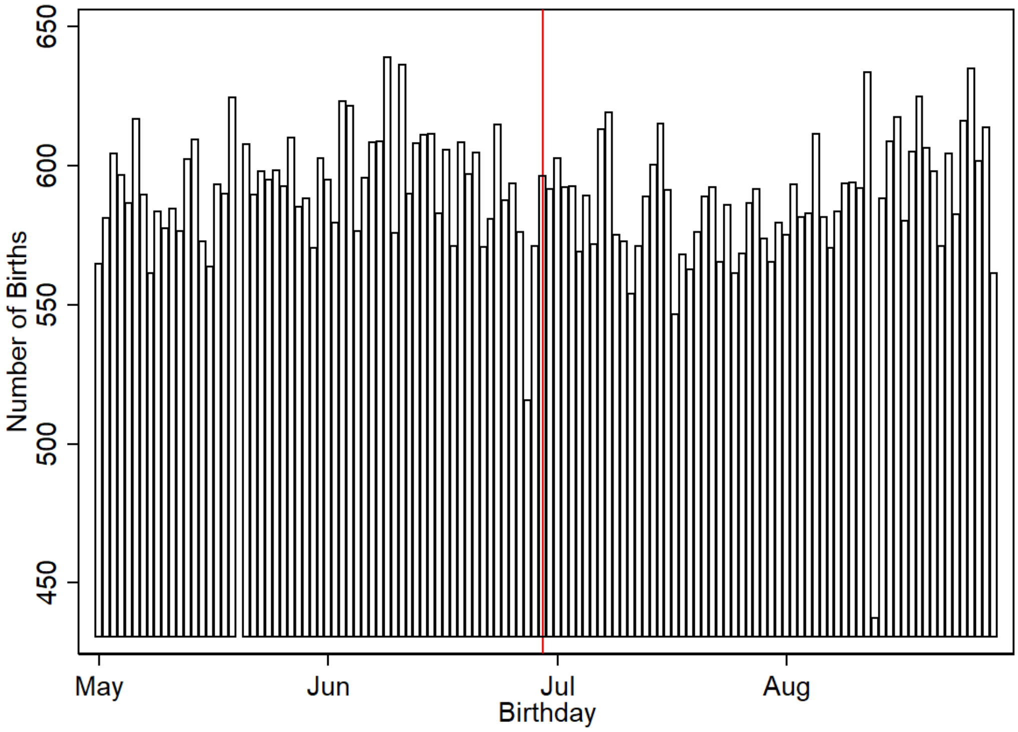
Nevertheless, the fact that families can sort themselves over the calendar year does not invalidate this quasi-experimental design. The critical identification assumption is that individuals lack the power to *precisely* sort themselves around these discontinuities. Under precise manipulation, we would find observations stacking up around the discontinuities, or in other words, we would observe a discontinuous distribution for the day of birth (the running variable).

Panel (a) of Figure 5 presents the raw histogram for the day of birth for all individuals born from December 15th to July 6th. Despite the high volatility, the figure hides a quite uniform distribution of births during the calendar year but with a high dispersion across weekdays and years of birth. Specifically, we observe an average of 650 births within a range of 500 to 800 births per day. However, when we control for the day of the week, holidays, and year of birth fixed effects, as shown in Panel (b)

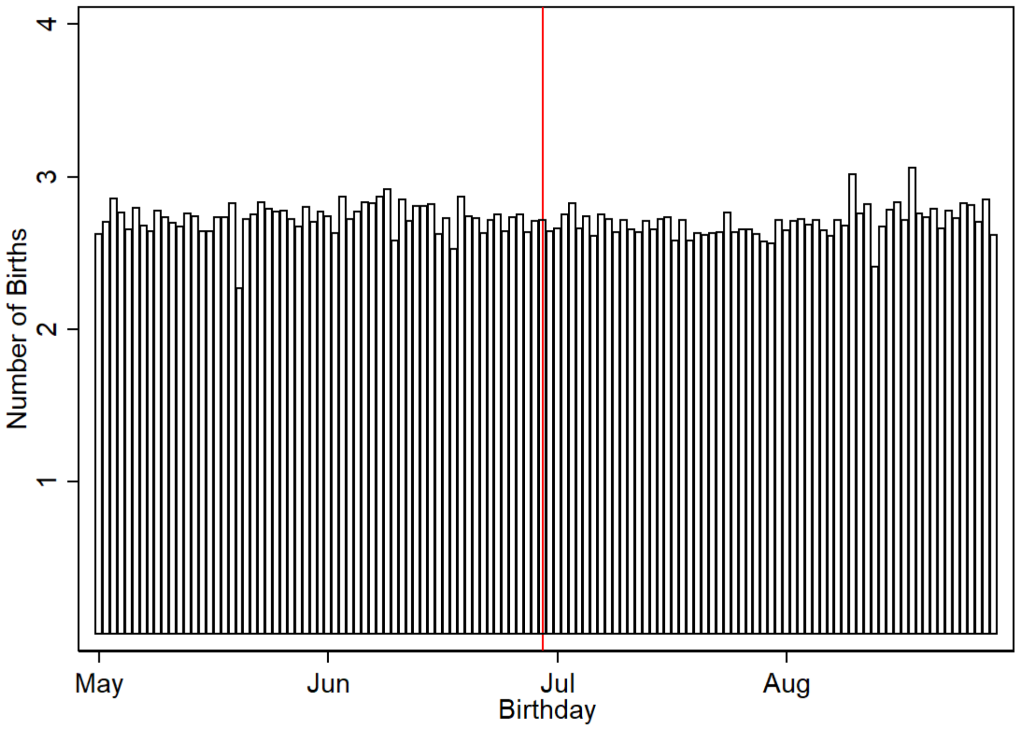
[27] Away from July's cutoff, Caceres-Delpiano and Giolito (2018) find discontinuities around January 1^{st}, May 1^{st}, and September 18^{th}, corresponding to three major holidays (Christmas-New Year's, Labor Day, and Independence Day, respectively). Therefore, these discontinuities can be explained by a drop in the number of births during holidays, associated with the high C-section rate in Chile. That is why, in our preferred specification, we control for whether or not a particular student was born on a legal holiday.

[28] Specifically, we use the regression model $covariate_i = \eta_{wh} + \phi_b + \gamma * 1\{b_i - C > 0\} + g(b_i) + v_i$, for each of the predetermined variables. Here $1\{b_i - C > 0\}$ is an indicator variable taking a value of one for students whose birthday (b_i) is over the cutoff (C), and zero otherwise. As in equation (1) , ϕ_b represents year of birth fixed effects, and η_{wh}, week day-holiday fixed effects. The null hypothesis for which the p-values are reported, $\gamma = 0$, corresponds to the scenario where there are not differences in the predetermined variables between children over and those below the cutoff.

[29] Buckles and Hungerman (2013) show that in the United States the season of birth is correlated with some mother's characteristics.



(a) Day of Birth. Raw Histogram.



(b) Day of Birth. Conditional Histogram

Figure 5: Raw and conditional histogram for the day of birth

of Figure 5, there is a uniform distribution of births across the calendar year and no discontinuities in the distribution of the running variable over the cutoffs. Therefore, in a given municipality and year of birth, we observe approximately 3 births per day.[30] Finally, following McCrary (2008), we formally test for a discontinuity in the distribution of the running variable by running a regression that has as the dependent variable the frequency of birthdays over the calendar year, as we did with the previous three pre-determined variables.[31] We report the p-values in Table 4. We check for sorting considering all cutoffs (column 1) or just July's (column 2). Only for one specification (quadratic polynomial) and bandwidth (five days), do we reject the null at a significance level of 5%. Therefore, parents might be able to select approximately the week or month of the birth, but they cannot precisely choose the day of birth.

6 Results

6.1 First stage

The evidence in Section 5 suggests that our empirical strategy meets the first requirement: other observed and unobserved factors seem uncorrelated with the potential source of variation at the age cutoff. Specifically, we find that the induced variation resulting from the different cutoffs is uncorrelated with other pre-determined variables. We also find no sign that families can precisely manipulate the day of birth. Nevertheless, we also need the minimum age requirements to produce a relevant variation in our potentially endogenous variable, SSA, measured by the dummy variable $DelayedEntry$. As shown in Figure 1 on page 6, the age cutoffs are associated with an evident jump in the fraction of students delaying entry.

We formally test for a relevant variation in the endogenous variable in the analysis by estimating equation (2). For clarity in the exposition, in Table 5 we report just the estimates of δ, that is, the discontinuity in the probability of delaying school entry between those children whose birthday occurs after the cutoff and those whose birthday is before that date, or, in other words, the impact of the excluded instruments. The first three columns report the estimates for a 5-day window around the cutoff, and the last three, for a 12-day window.[32] For each of the bandwidths, in the first two columns, we use a local linear specification for $g(b_i)$, without and with controls, and in the last column, we use a quadratic specification. Looking at panel A, the complete sample of students, the first element to be noticed is the qualitative robustness of the estimates when we include additional variables in the model. This robustness to the inclusion of other covariates is consistent with the evidence presented

[30] Notice in Figure 5b that, even though we do observe discontinuities around January 1st and May 1st, they correspond to Christmas- New Year's and the Labor Day holidays, respectively.

[31] See footnote 28.

[32] We report the results for these two-day windows since they are the upper and lower ranges of the optimal data-driven bandwidths found for the different outcomes and reported in Table 2.

in the previous section about the lack of correlation between observed predetermined variables and the treatments around the discontinuities. The second element to be noticed is the robustness of the estimates to the bandwidth used.

Regarding the estimated δ, notice, first, that being born after the cutoff increases the probability of delaying school entry by approximately 45 percentage points compared to students with a birthday just before July 1^{st}. Recall also from Figure 1 that a little less than half of the children born before July delay school entry to after the academic year closest to their sixth birthday. Therefore, our estimates for the last cutoff suggest perfect compliance with the minimum age rule.

Following the equivalence with an IV approach, the value of the F-statistic (reported at the bottom of the table) suggests disregarding any concern about weak instruments.

Panels B and C of Table 5 explore the heterogeneity in the source of variation by parents' education and student's gender. This analysis helps us to confirm a relevant variation for these two subsamples. In addition, it also shows the robustness of our estimates across families with different educational levels or student gender. The difference in point estimates suggests the highest compliance among families with less-educated parents. Independently of the specification, the increase in the probability of delaying school entry is approximately 10 percentage points higher among students with less-educated parents. The differences by student's gender are smaller, with girls being approximately 5 percentage points more likely to postpone starting due to minimum age requirements. Despite these differences in the degrees of compliance, for all the subsamples, we observe a relevant source of variation induced by the discontinuity determined by the entry rules.

6.2 School of entry characteristics

Complete sample

Table 10 shows the estimates of equation (1) on school characteristics. Specifically, Panel A of Table 10 shows the OLS estimates for the impact of delaying school entry, and Panel B shows the RD estimates using the two data-driven bandwidths described in Section 3: the point estimate optimal and the inference-optimal bandwidths.[33] For each of these bandwidths, we report a test for weak instruments. From these tests, we confirm that the variation induced by minimum age eligibility requirements is such that it allows us to rule out a concern about weak instruments.

The RD estimates reveal that delaying school entry is associated with a change in terms of the type of school of first enrollment. First, as shown in columns (1) and (2) of Table 10, we observe that a delay in enrollment increases the likelihood that a child would start primary school in a private school by approximately 3.7/4.4 percentage points, and specifically in a voucher school by 3.1/3.8 percentage points, an increase of about 6.5% in terms of the sample mean. This finding is

[33] OLS estimates were obtained by using a sample of students with a birthday 15 days around the cutoffs.

consistent with a positive sorting of students under a voucher system and suggests the well-known importance of family/students' background in educational decisions. That is, families' decision for an early enrollment introduces additional noise in the process of school sorting, which harms the entry to schools that are more likely to select students, as private schools (and specifically voucher schools) are.

Column (4) of Table 10 shows that students delaying entry have a gain of approximately 0.17 years in the average education of the parents in the school, smaller than the OLS estimate of 0.3 years. Column (5) of Table 10 shows that students delaying enrollment start in schools with an average higher achievement measured by their standardized test scores from previous years. Specifically, we observe an average statistical difference of 0.09 to 0.11 standard deviations between students depending on whether or not they delay school entry. This impact is smaller than the OLS estimate of approximately 0.13 standard deviations. This overestimation in the gain using OLS, together with the well-documented positive sorting of students/families across schools under a voucher system, suggests a positive selection of students/families into a delayed school entry. Consistently with other outcomes, the OLS estimate incorrectly reflects a reduction in the probability of switching schools of approximately 1 percentage point (10% in terms of sample mean) compared with the RD estimates. As shown in column (7), we find no impacts on the probability of switching schools during elementary school. This result for the probability of switching schools when using an RD-design is not only due to a less precise estimate but the point estimates are considerably lower than those from OLS, being quite robust to the bandwidths used.

Heterogeneity

Table 7 shows the heterogeneity of the previous results by parents' education level. Panel A presents the estimates for both data-driven bandwidths (MSE-optimal and the inference valid) for parents with 12 years of education or less, while the lower panel shows those families with more than 12 years of education. [34]

Notice in Panel B of Table 7 that, for more educated parents, we only find a significant impact in the average SIMCE test score (0.1 standard deviations) and average parents' education (0.2 years). However, these effects are not robust to the bandwidth used. Therefore, our results for the complete sample are mainly driven by children with parents with 12 years of education or less.

As we can see in columns (1) and (2) of Panel A of Table 7, delaying school entry is associated with an increase of approximately 6.5 (6) percentage points in the probability of attending a private (voucher) school or 13% in terms of the sample mean for children of less-educated parents. For this group of students, we observe that delaying school entry implies a 0.12 standard deviations gain in

[34] Parents with missing information about years of education are pooled with those having 12 or fewer years of education.

the average school SIMCE score. Moreover, Column (3) shows, an increase of 3 percentage points in the probability of being enrolled in a paid school (unsubsidized private or with a copay), for one of the bandwidths. This impact means an increase of approximately 16% in terms of the sample mean for this group. Column (4) indicates that children with less-educated parents starting early face a decrease in the average education of their peer's parents of about 0.2 years, or 2% in terms of the sample mean.

Elder and Lubotsky (2009) find that the early benefits of SSA in terms of test scores are higher among upper-income families. They use this evidence (plus the timing of the effect) to show that SSA is a function of the skill accumulation before kindergarten. In a context of a voucher system where school choice is itself an investment margin, and with schools also actively selecting students, our findings support an attempt from lower-income families to partially compensate for these differences in previous skill accumulation.

Table 8 reports the estimates by student's gender. Despite the reported disadvantage of boys in terms of school readiness (Janus and Duku, 2007), our estimates reveal a higher benefit for girls. The first three columns of Panel B show that a delayed entry increases the probability of girls starting in a private or voucher school by approximately 5 percentage points. We also observe an increase of around 0.20 years in mean parent education (Column (4)) and a positive impact of approximately 0.11-0.12 standard deviations in the average school SIMCE score (Columns(5)) when a female student delays school entry.

As shown in Panel A, the estimates for boys reveal that, despite the positive effect on the probability of being enrolled in a private school, the remaining estimates are either imprecisely estimated, or not robust to the bandwidth used. This clear gain for girls may be interpreted in light of recent literature revealing that girls react worse to competition (Niederle, 2017), and this disadvantage is set early in life (Sutter and Glätzle-Rützler, 2015). Specifically, our results suggest that girls are taking advantage of the delayed entry to access better (and perhaps more selective) schools. This effect speaks against an effect operating through a channel of school readiness. Girls usually mature before boys, so they should be the ones gaining less by waiting.

School characteristics and determinants of school choice from parents' surveys

In Table 11, we use parent surveys to provide some evidence about the mechanisms behind the impact of delaying school entry on the characteristics of the chosen school. In an institutional setting like the one in Chile with a decentralized system funded via a voucher scheme, schools are expected to select the best students. Moreover, families pushed to delay school entry might spend more time finding the best-fitting school for their children.

Columns (1) and (2) of Table 11 show the estimates for two variables constructed at the school level, based on parents' answers from previous years. The results in the first column reveal that students who delay school entry are enrolled in a school where a higher fraction of parents reported some kind of academic selection. Specifically, we find that delaying school entry implies a 7 to 8 percentage point increase in this fraction, (an 11% increase in terms of the sample mean). As column (2) shows, we also observe a positive effect on the fraction of parents reporting a selection based on family background, even though this estimate is not robust to the bandwidth selected.

Columns (3) to (6) of Table 11 refer to the reasons driving the choice of their children's current school by families. Despite the fact that these variables reflect the reason to select a school at the time of the SIMCE exam, these reduced-form estimates are likely to be driven by the selection of the first school, considering that we do not observe any effect on the probability of switching school. As shown in column (5), we do not see those families delaying school entry being more likely to report that the average test score is the main reason for choosing a school. However, columns (4) and (6) show us that families delaying school entry are more likely to report school proximity (2.5 percentage points or 5% in terms of the sample mean) or school environment (16 percentage points or 50% in respect to the sample mean) as the reason to select the school.

Finally, as shown in column (3), we find that those families who delay school entry are more likely to report being constrained (affordability or proximity) in their choice of school than those students not delaying entry. This result, first, might reflect a larger investment effort by the families, which is consistent with the positive effect on attending a private school or paying tuition (for some of the bandwidths).

The previous results suggest that gains in terms of school characteristics seem to be driven jointly by schools and families. On the one hand, schools actively choose students, so older children are more likely to be admitted to selective schools, which explains the gains in terms of average school test scores. Secondly, families are more likely to invest in the search process (with a higher probability to report being constrained in school choice), and prioritizing environment, which might be behind the choice of a private school or the average education of the parents in the selected school.

7 Conclusions

Our findings provide evidence about an alternative channel in the impact of SSA. Delaying school entry induces a relevant shift in the opportunity to start in a "better" school. First, students who delay entry are enrolled in a school with an approximately 0.1 standard deviations higher average standardized test score. Second, students who delay school entry face peers whose parents have 0.17 more years of education. Third, a delay in enrollment in elementary education increases the probability to start in a private school funded via voucher by 4 percentage points. (8%). The analysis by level of education of the parents reveals that this last result is driven by parents with lower levels of education and girls.

We also provide evidence that delaying school entry increases the probability of being enrolled in a school with some kind of academic selection, which would explain the increase in the average score of the starting school in standardized tests. We also find that families delaying school entry are more likely to express being constrained by the cost or the availability of the schools in their counties (municipalities), which we interpret as an increase in investment effort from the side of the families.

References

Angrist, J. D. and A. B. Kruger (1991). Does compulsory schooling attendance affect schooling and earnings? *Quarterly Journal of Economics 106*, 979–1014.

Attar, I. and D. Cohen-Zada (2018). The effect of school entrance age on educational outcomes: Evidence using multiple cutoff dates and exact date of birth. *Journal of Economic Behavior & Organization 153*, 38–57.

Bedard, K. and E. Dhuey (2006, 11). The Persistence of Early Childhood Maturity: International Evidence of Long-Run Age Effects*. *The Quarterly Journal of Economics 121*(4), 1437–1472.

Black, S., P. Devereaux, and P. Salvanes (2011). Too young to leave the nest? the effect of school starting age. *The Review of Economics and Statistics 93*(2), 455–467.

Buckles, K. and D. Hungerman (2013). Season of birth and later outcomes: Old questions, new answers. *Review of Economics and Statistics 95*(3), 711–724.

Caceres-Delpiano, J. and E. Giolito (2018). Minimum Age Requirements and the Impact of School Choice. *IZA Discussion paper series 11420*.

Caceres-Delpiano, J. and E. P. Giolito (2019). *The impact of age of at entry on academic progression*, Chapter Data-Driven Policy Impact Evaluation: How Microdata is Transforming Policy Design. Springer.

Calonico, S., M. Cattaneo, and R. Titiunik (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica 82*(6), 2295–2326.

Cascio, E. and E. Lewis (2006). Schooling and the armed forces qualifying test. evidence from school-entry laws. *Journal of Human Resources 41*(2), 294–318.

Contreras, D., O. Larranaga, L. Flores, F. Lobato, and V. Macias (2005). Politicas educacionales en chile: Vouchers, concentracion, incentivos y rendimiento. *In Uso e impacto de la informacion educativa en America Latina, ed. Santiago Cueto. Santiago: PREAL.*, 61–110.

Cook, P. J. and S. Kang (2016, January). Birthdays, schooling, and crime: Regression-discontinuity analysis of school performance, delinquency, dropout, and crime initiation. *American Economic Journal: Applied Economics 8*(1), 33–57.

Dee, T. and H. H. Sievertsen (2018). The gift of time? School starting age and mental health. *Health Economics 27*(5), 781–802.

Deming, D. and S. Dynarski (2008, Summer). The lengthening of childhood. *The Journal of Economic Perspectives 22*(3), 71–92.

Dhuey, E. and S. Lipscomb (2010). Disabled or young? relative age and special education diagnoses in schools. *Economics of Education Review 29*(5), 857–872.

Elder, T. and D. H. Lubotsky (2008). Kindergarten entrance age and children's achievement. *Journal of Human Resource 44*(3), 641–683.

Elder, T. E. and D. H. Lubotsky (2009). Kindergarten entrance age and children's achievement: Impacts of state policies, family background, and peers. *The Journal of Human Resources 44*(3), 641–683.

Epple, D., R. E. Romano, and M. Urquiola (2017). School vouchers: A survey of the economics literature. *Journal of Economic Literature 55*(2), 441–492.

Fredriksson, P. and B. Öckert (2014). Life-cycle effects of age at school start. *The Economic Journal 124*(579), 977–1004.

Gauri, V. and A. Vawda (2003). Vouchers for basic education in developing countries. a principal-agent perspective. *World Bank Policy Research Working Paper 3005*.

Hahn, J., P. Todd, and W. van der Klaauw (2001). Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica 69*(1), 201–209.

Heckman, J., S. Urzua, and E. Vytlacil (2006). Understanding instrumental variables in models with essential heterogeneity. *Review of Economics and Statistics 88*(3), 389–432.

Hsieh, C. and M. Urquiola (2006). The effects of generalized school choice on achievement and stratification: Evidence from chile's voucher program. *Journal of Public Economics 90*(8-9), 1477–1503.

Janus, M. and E. Duku (2007). The school entry gap: Socioeconomic, family, and health factors associated with children's school readiness to learn. *Early Education and Development 18*(3), 375–403.

Landersø, R., H. S. Nielsen, and M. Simonsen (2016, 08). School Starting Age and The Crime Age Profile. *The Economic Journal 127*(602), 1096–1118.

Landersø, R. K., H. S. Nielsen, and M. Simonsen (2020). Effects of School Starting Age on the Family. *Journal of Human Resources 55*(4), 1258–1286.

Lee, D. S. and T. Lemieux (2010). Regression discontinuity designs in economics. *Journal of Economic Literature 48*, 281–355.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics 142*(2).

McEwan, P. and J. Shapiro (2006). The benefit of delayed primary school enrollment. discontinuity estimates using exact birth dates. *Journal of Human Resources 43*(1), 1–29.

Mühlenweg, A. M. and P. A. Puhani (2010). The Evolution of the School-Entry Age Effect in a School Tracking System. *Journal of Human Resources 45*(2), 407–438.

Niederle, M. (2017). A gender agenda: A progress report on competitiveness. *The American Economic Review 107*(5), 115–119.

Sutter, M. and D. Glätzle-Rützler (2015). Gender differences in the willingness to compete emerge early in life and persist. *Management Science 61*(10), 2339–2354.

Table 1: Descriptive Statistics

Panel A: Whole Sample	Mean	SD
Private schl.	0.60	0.49
Voucher schl.	0.52	0.50
Paid schl.	0.30	0.46
Mean schooling parents	11.03	2.16
Average Schl SIMCE	251.87	27.31
Switch schl.	0.14	0.35
Schl. academic selection	0.64	0.24
Schl. family background selection	0.36	0.28
Family constrained in selection	0.60	0.49
Family selects based on proximity	0.51	0.50
Family selects based on SIMCE	0.27	0.44
Family selects based on enviroment	0.43	0.50
Family selects based on relatives	0.51	0.50
Later Entry	0.21	0.41
SSA	6.14	0.35
Male students	0.51	0.50
Father's education	11.59	3.85
Mother's education	11.57	3.59
Panel B: Students with birthday within 8 days around cutoff		
Private school	0.61	0.49
Voucher school	0.53	0.50
Tuition Paid school	0.30	0.46
Mean schooling parents	11.07	2.17
Average School SIMCE	252.12	27.67
Switch school.	0.14	0.35
School academic selection	0.65	0.24
School family background selection	0.36	0.28
Family constrained in selection	0.62	0.49
Family selects based on proximity	0.53	0.50
Family selects based on SIMCE	0.30	0.46
Family selects based on enviroment	0.48	0.50
Family selects based on relatives	0.52	0.50
Later Entry	0.75	0.43
SSA	6.43	0.42
Male students	0.51	0.50
Father's education	11.60	3.88
Mother's education	11.61	3.59

Panels A and B correspond to the complete sample of students and those with a birthday 8 days around the July's cutoff, respectively.

Table 2: Optimal bandwidth by selected outcomes and cutoffs

	Method	Optimal bandwidth
Private schl.	MSE	12.6
	CER	10.3
Voucher schl.	MSE	12.6
	CER	10.3
Paid schl.	MSE	13.1
	CER	10.6
Parents Educ at FG	MSE	10.7
	CER	8.7
Schl SIMCE	MSE	10.0
	CER	8.1
Switch schl.	MSE	11.8
	CER	9.6
Schl. academic selection	MSE	8.2
	CER	6.6
Schl. family background selection	MSE	6.6
	CER	5.3
Family constrained in selection	MSE	8.3
	CER	6.7
Family selects based on proximity	MSE	8.3
	CER	6.7
Family selects based on SIMCE	MSE	11.9
	CER	9.7
Family selects based on enviroment	MSE	7.6
	CER	6.2
Family selects based on relatives in schl or other factors	MSE	11.8
	CER	9.6

MSE: Mean Squared Error (MSE) optimal bandwidth.
CRE: optimal bandwidth that minimizes the asymptotic coverage error rate of the robust bias corrected confidence interval.
Both bandwidths were chosen following Calonico (2017), and implemented with the command rdbwselect in STATA, a local linear polynomial and triangular kernel.

Table 3: Differences in predetermined variables between children born before and after selected cutoffs. p-values reported.

Bandwidth				
	Male		Parents education	
	Panel A. $g(x_i^s)$ of degree 3			
15 days	0.217	0.601	0.170	0.579
10 days	0.248	0.562	0.087	0.577
5 days	0.000	0.070	0.000	0.185
	Panel B. $g(x_i^s)$ of degree 2			
15 days	0.972	0.941	0.598	0.785
10 days	0.494	0.694	0.348	0.630
5 days	0.001	0.297	0.000	0.390
	Panel C. $g(x_i^s)$ of degree 1			
15 days	0.820	0.387	0.097	0.328
10 days	0.884	0.522	0.162	0.410
5 days	0.757	0.898	0.407	0.908
Cutoffs	Feb-July	July	Feb-July	July

For each of the variables (w_i), reported on the top of the columns, we run the regression, $w_i = \alpha_s + \eta_{wh} + \phi_b + \gamma^s * 1\{b_i - C >\} + g(x_i^s) + v_{it}$ where $1\{b_i - C > 0\}$ is an indicator variable taking a value of one for sudents whose birthday (b_i) is over the cutoff (C), and zero otherwise. α_s is a specific constant for individuals around the s cutoff. η_{wh} and ϕ_b represent week-day/holiday, and year of birth fixed effects, respectively. The null hypothesis for which the p-values are reported is $H0 : \gamma^s = 0$, that is, there are not differences in the predetermined variables between children over and below any of the cutoffs. Selected cutoffs indicated at the bottom of the table.

Table 4: Differences in the number of births before and after selected cutoffs. p-values reported.

Bandwidth		
	Panel A. $g(x_i^s)$ of degree 3	
15 days	0.219	0.076
10 days	0.260	0.499
5 days	0.318	0.122
	Panel B. $g(x_i^s)$ of degree 2	
15 days	0.352	0.114
10 days	0.195	0.063
5 days	0.891	0.008
	Panel C. $g(x_i^s)$ of degree 1	
15 days	0.072	0.473
10 days	0.058	0.280
5 days	0.519	0.166
Cutoffs	Feb-July	July

The dependent variable is the number of births in a specific day of the calendar year, in the different counties, for the different cohorts in the analysis. We run the regression, $w_i = \alpha_s + \eta_{wh} + \phi_b + \gamma^s * 1\{b_i - C >\} + g(x_i^s) + v_{it}$ where $1\{b_i - C > 0\}$ is an indicator variable taking a value of one for students whose birthday (b_i) is over the cutoff (C), and zero otherwise. α_s is a specific constant for individuals around the s cutoff. η_{wh} and ϕ_b represent week-day/holiday, and year of birth fixed effects, respectively. The null hypothesis for which the p-values are reported is $H0 : \gamma^s = 0$, that is, there are not differences in the predetermined variables between children over and below any of the cutoffs. Selected cutoffs indicated at the bottom of the table.

Table 5: First stage estimates. Impact of age eligibility requirement on the probability of delaying school entry.

	Bandwidth					
	5 days			12 days		
	[1]	[2]	[3]	[4]	[5]	[6]
Panel A:	**Complete sample**					
	0.469***	0.470***	0.447***	0.471***	0.471***	0.471***
	(0.012)	(0.011)	(0.008)	(0.006)	(0.005)	(0.011)
Observations	40846	40791	40791	88906	88768	88768
F excluded instr.	1415.9	1906.8	2823.3	5268.4	7398.7	1750.4
Panel B:	**By Parents' educational level**					
	12 years of education or less					
	0.505***	0.509***	0.496***	0.507***	0.509***	0.514***
	(0.011)	(0.009)	(0.012)	(0.007)	(0.006)	(0.012)
Observations	22551	22496	22496	49077	48939	48939
F excluded instr.	2087.4	3064.3	1771.5	5793.3	8473.5	1861.9
	More than 12 years of education					
	0.423***	0.419***	0.383***	0.425***	0.421***	0.416***
	(0.015)	(0.017)	(0.008)	(0.008)	(0.008)	(0.014)
Observations	18295	18295	18295	39829	39829	39829
F excluded instr.	761.5	635.7	2376.2	2731.3	2704.9	907.7
Panel C:	**By Student's gender**					
	Boys					
	0.445***	0.444***	0.430***	0.453***	0.454***	0.446***
	(0.014)	(0.010)	(0.012)	(0.008)	(0.007)	(0.011)
Observations	20823	20790	20790	45159	45080	45080
F excluded instr.	1013	2041.8	1205.7	2864.4	4491.2	1684.8
	Girls					
	0.494***	0.495***	0.462***	0.490***	0.488***	0.499***
	(0.012)	(0.013)	(0.008)	(0.006)	(0.007)	(0.014)
Observations	20023	20001	20001	43747	43688	43688
F excluded instr.	1780.3	1375.5	3154.1	5816.9	5513.9	1298.8
Controls		X	X		X	X
Degree Pol.	1	1	2	1	1	2

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered at the municipality level.
The dependent variable, *LaterEntry*, is a dummy variable that takes a value one for children who start primary school later than the closest academic year to when they turn six. Specifications with additional controls include municipality-year of birth, parents' education, gender, weekday of birth, and born on a holiday dummies.

Table 6: Impact of SSA on school characteristics. Complete Sample.

	Private	Voucher	Tuition paid	Mean Schooling Parents	Avg. SIMCE Score	Switch School
	[1]	[2]	[3]	[4]	[5]	[6]
Panel A: OLS						
	0.031***	-0.017***	0.036***	0.327***	0.128***	-0.010***
	(0.002)	(0.003)	(0.002)	(0.019)	(0.008)	(0.001)
Observations	739882	739882	739882	622996	734480	5501162
Panel B: RD estimates for two data-driven bandwidths						
			MSE-optimal bandwidth			
	0.037***	0.031***	0.011	0.145**	0.089***	0.001
	(0.010)	(0.009)	(0.011)	(0.058)	(0.019)	(0.003)
Observations	88768	88768	88768	63632	67873	608466
Mean	.61	.53	.31	11.06	.33	.1
Weak identification (a)	7402.8	7402.8	7402.8	6076.8	4807.5	5190.52
			CER-optimal bandwidth			
	0.044***	0.038***	0.018	0.201***	0.107***	0.001
	(0.010)	(0.012)	(0.011)	(0.068)	(0.019)	(0.003)
Observations	68393	68393	75291	51834	53796	507733
Mean	.61	.53	.31	11.06	.33	.1
Weak identification (a)	5073.1	5073.1	5212.4	3973	3031	4017.27

*** $p<0.01$, ** $p<0.05$, * $p<0.1$. Standard errors clustered at the municipality level.
Standard errors clustered at the municipality level.
Additional controls: municipality-year of birth, parents' education, gender, weekday of birth, and born on a holiday dummies. For the outcome of switching school, also years since first enrollment fixed effects are included. For each of the outcomes, two data driven bandwidths are used. These bandwidths are reported in Table 2.
(a) Kleibergen-Paap rk Wald F statistic. Stock-Yogo weak ID test critical values for 10, 15 and 20 percent maximal IV size are 16.38, 8.96, and 6.66, respectively.

Table 7: Impact of SSA on school characteristics by parents' education.

	Private	Voucher	Tuition paid	Mean Schooling Parents	Avg. SIMCE Score	Switch School
	[1]	[2]	[3]	[4]	[5]	[6]
Panel A:			*12 years of education or less*			
MSE-optimal bandwidth	0.056***	0.054***	0.016	0.178**	0.113***	0.004
	(0.010)	(0.011)	(0.012)	(0.079)	(0.023)	(0.003)
Observations	48939	48939	48939	29836	37402	322434
Mean	.49	.46	.19	9.94	.03	.1
Weak identification (a)	8483.1	8483.1	8483.1	8557.1	7981.3	5265.87
CER-optimal bandwidth	0.071***	0.065***	0.029***	0.222***	0.116***	0.005*
	(0.010)	(0.012)	(0.011)	(0.082)	(0.025)	(0.003)
Observations	37678	37678	41483	24275	29603	268909
Mean	.49	.46	.19	9.94	.03	.1
Weak identification (a)	8910	8910	6887.6	6051.7	4614.4	5021.68
Panel B:			*More than 12 years of education*			
MSE-optimal bandwidth	0.013	-0.000	0.010	0.097	0.053	-0.001
	(0.016)	(0.016)	(0.019)	(0.064)	(0.033)	(0.007)
Observations	39829	39829	39829	33796	30471	286032
Mean	.76	.61	.45	12.06	.71	.1
Weak identification (a)	2709.2	2709.2	2709.2	2078.1	1629.3	2935.09
CER-optimal bandwidth	0.010	0.001	0.007	0.175**	0.098***	-0.003
	(0.019)	(0.021)	(0.020)	(0.073)	(0.025)	(0.007)
Observations	30715	30715	33808	27559	24193	238824
Mean	.76	.61	.45	12.06	.71	.1
Weak identification (a)	1618.1	1618.1	2073.5	1245.1	1147.6	1809.77

*** $p<0.01$, ** $p<0.05$, * $p<0.1$. Standard errors clustered at the municipality level.
Standard errors clustered at the municipality level.
Additional controls: municipality-year of birth, parents' education, gender, weekday of birth, and born on a holiday dummies. For the outcome of switching school, also years since first enrollment fixed effects are included. For each of the outcomes in a given sample, two data driven bandwidths are used. These bandwidths are reported in Table 2.
(a) Kleibergen-Paap rk Wald F statistic. Stock-Yogo weak ID test critical values for 10, 15 and 20 percent maximal IV size are 16.38, 8.96, and 6.66, respectively.

Table 8: Impact of SSA on school characteristics by student's gender.

	Private	Voucher	Tuition paid	Mean Schooling Parents	Avg. SIMCE Score	Switch School
	[1]	[2]	[3]	[4]	[5]	[6]
Panel A:				*Boys*		
MSE-optimal bandwidth	0.031***	0.022**	0.011	0.078	0.059	0.002
	(0.011)	(0.009)	(0.014)	(0.071)	(0.037)	(0.006)
Observations	45080	45080	45080	31951	34432	309144
Mean	.6	.53	.3	11.03	.31	.1
Weak identification (a)	4496.2	4496.2	4496.2	2218.7	4065.9	2857.74
CER-optimal bandwidth	0.028**	0.020	0.009	0.141	0.113***	0.007
	(0.014)	(0.013)	(0.016)	(0.090)	(0.036)	(0.006)
Observations	34728	34728	38240	25950	27281	257435
Mean	.6	.53	.3	11.03	.31	.1
Weak identification (a)	4045.2	4045.2	2938.6	1867.6	2813.8	2916.58
Panel B:				*Girls*		
MSE-optimal bandwidth	0.041***	0.037***	0.009	0.197***	0.119***	0.000
	(0.014)	(0.014)	(0.013)	(0.066)	(0.016)	(0.005)
Observations	43688	43688	43688	31681	33441	299322
Mean	.61	.53	.31	11.09	.36	.1
Weak identification (a)	5520.1	5520.1	5520.1	7620.1	3176	4733.98
CER-optimal bandwidth	0.059***	0.054***	0.024**	0.248***	0.112***	-0.004
	(0.014)	(0.013)	(0.012)	(0.069)	(0.017)	(0.005)
Observations	33665	33665	37051	25884	26515	250298
Mean	.61	.53	.31	11.09	.36	.1
Weak identification (a)	3501.9	3501.9	4391.5	5235.3	2053.6	3237.95

*** $p<0.01$, ** $p<0.05$, * $p<0.1$. Standard errors clustered at the municipality level.
Standard errors clustered at the municipality level.
Additional controls: municipality-year of birth, parents' education, gender, weekday of birth, and born on a holiday dummies. For the outcome of switching school, also years since first enrollment fixed effects are included. For each of the outcomes in a given sample, two data driven bandwidths are used. These bandwidths are reported in Table 2.
(a) Kleibergen-Paap rk Wald F statistic. Stock-Yogo weak ID test critical values for 10, 15 and 20 percent maximal IV size are 16.38, 8.96, and 6.66, respectively.

Table 9: Impact of SSA on school selection characteristics.

	School selects based on			Family selects based on			
	Academic	Family background	Family constrained selection	Proximity	SIMCE	Environment	Relatives/ other
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Panel A: OLS							
	0.031***	0.048***	-0.008***	-0.013***	0.023***	0.131***	0.002
	(0.002)	(0.003)	(0.003)	(0.003)	(0.002)	(0.005)	(0.002)
Observations	2304789	2304789	604795	604795	604795	604795	604795
Panel B:				**RD estimates for two data driven bandwidth**			
				MSE-optimal bandwidth			
	0.075***	0.029**	0.036**	0.027*	0.016	0.163***	-0.020
	(0.013)	(0.015)	(0.015)	(0.016)	(0.015)	(0.015)	(0.020)
Observations	23520	20765	41535	41535	63204	41535	63204
Mean	.65	.36	.6	.5	.3	.5	.5
Weak identification (a)	124.7	48.4	1622.62	1622.62	2918.57	1622.62	2918.57
				CER-optimal bandwidth			
	0.073***	-0.001	0.035**	0.024*	0.020	0.164***	-0.014
	(0.013)	(0.025)	(0.015)	(0.014)	(0.016)	(0.016)	(0.018)
Observations	20765	14843	36465	36465	52377	31394	52377
Mean	.65	.36	.6	.5	.3	.5	.5
Weak identification (a)	48.4	61.5	809.66	809.66	2077.49	548.45	2077.49

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered at the municipality level.
Standard errors clustered at the municipality level.
Additional controls: municipality-year of birth, parents' education, gender, weekday of birth, and born on a holiday dummies. For the outcome of switching school, also years since first enrollment fixed effects are included. For each of the outcomes, two data driven bandwidths are used. These bandwidths are reported in Table 2.
(a) Kleibergen-Paap rk Wald F statistic. Stock-Yogo weak ID test critical values for 10, 15 and 20 percent maximal IV size are 16.38, 8.96, and 6.66, respectively.

Appendix

Table 10: Impact of SSA on school characteristics. Complete Sample. All cutoffs.

	Private	Voucher	Tuition paid	Mean Schooling Parents	Avg. SIMCE Score	Switch School
	[1]	[2]	[3]	[4]	[5]	[6]
	RD estimates for two data driven bandwidths					
	MSE-optimal bandwidth					
	0.040***	0.031***	0.011	0.188***	0.084***	0.002
	(0.011)	(0.010)	(0.010)	(0.051)	(0.020)	(0.003)
Observations	629684	629684	629684	446429	479635	4314637
Mean	.6	.52	.3	11.03	.32	.1
Weak identification (a)	3777.6	3777.6	3777.6	3486.2	2237.6	2649.57
	CER-optimal bandwidth					
	0.042***	0.034***	0.016*	0.240***	0.113***	0.001
	(0.012)	(0.012)	(0.009)	(0.052)	(0.022)	(0.003)
Observations	483187	483187	530442	365217	382946	3591275
Mean	.6	.52	.3	11.03	.32	.1
Weak identification (a)	2450	2450	2551.6	2523.6	1444.3	2067.4

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered at the municipality level.
Standard errors clustered at the municipality level.
Additional controls: municipality-year of birth, parents' education, gender, weekday of birth, and born on a holiday dummies. For the outcome of switching school, also years since first enrollment fixed effects are included. For each of the outcomes, two data driven bandwidths are used. These bandwidths are reported in Table 2.
(a) Kleibergen-Paap rk Wald F statistic. Stock-Yogo weak ID test critical values for 10, 15 and 20 percent maximal IV size are 16.38, 8.96, and 6.66, respectively.

Table 11: Impact of SSA on school selection characteristics. All cutoffs.

	School selects based on		Family	Family selects based on			
	Academic	Family background	constrained selection	promity	SIMCE	environment	relatives or other
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
RD estimates for two data driven bandwidth							
MSE-optimal bandwidth							
	0.057***	0.033*	0.049***	0.035**	-0.006	0.135***	-0.017
	(0.014)	(0.017)	(0.012)	(0.016)	(0.013)	(0.021)	(0.018)
Observations	139375	122183	315568	315568	474352	315568	474352
Mean	.64	.36	.6	.5	.3	.4	.5
Weak identification (a)	111	90	2494.14	2494.14	2546.15	2494.14	2546.15
CER-optimal bandwidth							
	0.052***	0.012	0.055***	0.041**	0.003	0.144***	-0.004
	(0.016)	(0.022)	(0.015)	(0.017)	(0.014)	(0.021)	(0.018)
Observations	122183	85600	276544	276544	394915	235522	394915
Mean	.64	.36	.6	.5	.3	.4	.5
Weak identification (a)	90	74.3	2030.49	2030.49	2531.69	1625.02	2531.69

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered at the municipality level.
Standard errors clustered at the municipality level.
Additional controls: municipality-year of birth, parents' education, gender, weekday of birth, and born on a holiday dummies. For the outcome of switching school, also years since first enrollment fixed effects are included. For each of the outcomes, two data driven bandwidths are used. These bandwidths are reported in Table 2.
(a) Kleibergen-Paap rk Wald F statistic. Stock-Yogo weak ID test critical values for 10, 15 and 20 percent maximal IV size are 16.38, 8.96, and 6.66, respectively.