Online Appendix not for publication

A Data-driven bandwidths

In line with Calonico et al. (2014), we opt for two data-derived bandwidths for each outcome. The initial approach aims to strike a balance between bias and variance, achieving this by minimizing the Mean Squared Error (MSE) of the local polynomial RD point estimator, often referred to as the MSE-optimal bandwidth.[41] Although the first bandwidth optimizes point estimates, it falls short in terms of inference since it's not sufficiently "narrow" to eliminate the predominant bias term for statistical reasoning. To address this issue, we adopt the under-smoothing approach, which involves utilizing a greater number of observations for point estimation than for making inferences.

For the second bandwidth, our focus shifts towards minimizing an approximation of the Coverage Error Rate (CER) of the confidence intervals. This quantifies the difference between the practical coverage of the confidence interval and its theoretical level. The selected bandwidths are presented in Table A.1. These bandwidths span a range of 5 to 12 days around the discontinuity, with the CER-optimal bandwidths consistently being narrower.

Table A.1: Optimal bandwidth by selected outcomes and cutoffs

	Method	Optimal bandwidth
Voucher School	MSE	12.6
	CER	10.3
Average schooling of classmates' parents	MSE	10.7
	CER	8.7
School SIMCE	MSE	10.0
	CER	8.1
Classmates age at entry	MSE	12.1
	CER	9.9
Teacher's Age	MSE	12.5
	CER	10.2
Teaching Hours / Total Hours Ratio	MSE	10
	CER	8.2
Share of full-time teachers at school	MSE	10.1
	CER	8.2
Share of teachers with 4-year college degree	MSE	8.4
	CER	6.8
School academic selection	MSE	8.2
	CER	6.6
School selects on family background	MSE	6.6
	CER	5.3

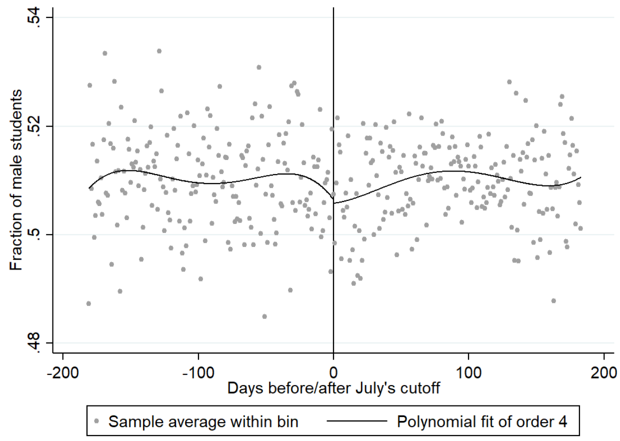
MSE: Mean Squared Error (MSE) optimal bandwidth.
CRE: optimal bandwidth that minimizes the asymptotic coverage error rate of the robust bias corrected confidence interval.
Both bandwidths were chosen following Calonico et al. (2017), and implemented with the command rdbwselect in STATA, a local linear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

[41] 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.

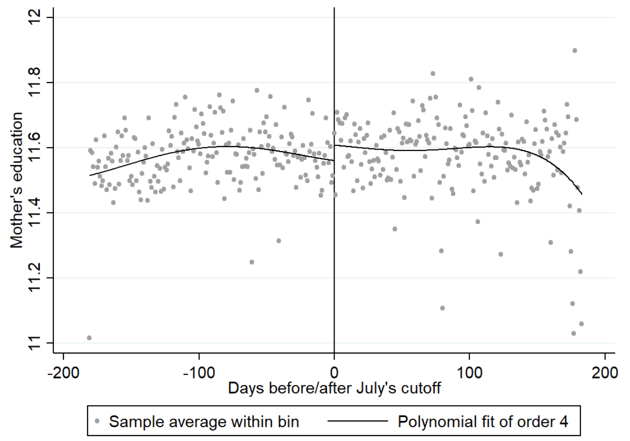
B Validity of the source of variation

B.1 Continuity in predetermined variables

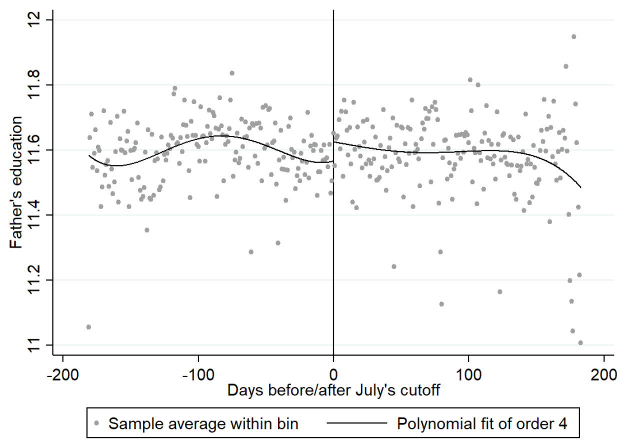
Our analysis relies on the concept that changes in school enrollment eligibility near a certain cutoff can be considered comparable to a randomized assignment. This assumption applies particularly to students whose birthdays fall near the cutoff date. Just as in standard random assignments, the attributes students have before this "randomization" should be comparable between the treated group (students born shortly after a cutoff) and the control group (students born just before the cutoff). If we identify consistent differences in these attributes around these dates, it would challenge 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 B.1: Balancing covariates

Figure B.1 visually examines the possible presence of a discontinuity in four fundamental characteristics found within the dataset: gender (percentage of males), educational levels of the mother and father, and the highest education attained by the student's parents. When contrasted with those in Figure 3, this graphical representation doesn't reveal any significant discontinuities or outliers for the specified variables. We carry out formal tests for potential discontinuities in these baseline characteristics, using various polynomial models and three bandwidths (5, 10, and 15 days). The resulting p-values are documented in Table B.2.[42] In Columns (2) and (4) of Table B.2, we check for discontinuities around July's cutoff, and in columns (1) and (3), we consider the February-July cutoffs. 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.

[42]We utilize a regression model for each of the predetermined variables, as follows: $covariate_i = \eta_{wh} + \phi_b + \gamma * 1\{b_i - C > 0\} + g(b_i) + v_i$. In this equation, $1b_i - C > 0$ is an indicator variable that takes a value of one for students with birthdays (b_i) after the cutoff (C), and zero otherwise. Similar to equation (1), ϕ_b represents fixed effects for the year of birth, and η_{wh} accounts for fixed effects related to week days and holidays. The p-values reported correspond to the null hypothesis $\gamma = 0$, which implies that there are no differences in the predetermined variables between children born above and below the cutoff.

Table B.2: Differences in predetermined variables between children born before and after selected cutoffs. p-values reported.

Bandwidth	Male		Parents education	
	Panel A. $g(b_i)$ 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. $g(b_i)$ 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. $g(b_i)$ 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, $w_i = \alpha_s + \eta_{wh} + \phi_b + \gamma^s * 1\{b_i - C >\} + g(b_i) + v_{it}$ 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.

Municipality variables

While it is not mandatory for students in Chile to reside in the same municipality as the school they attend, it is theoretically possible for parents to choose a neighborhood based on the school's cutoff date. However, considering that our observations begin at the commencement of schooling, our intention is to show that such a phenomenon does not appear to be prevalent in Chile. Figure B.2 illustrates the trajectory of three variables over the course of students' school life in the Santiago Metropolitan area, which comprises 52 municipalities. These variables include the fraction of students changing schools, the fraction of students changing their municipality of residence, and the fraction of students moving to a school located in a different municipality.

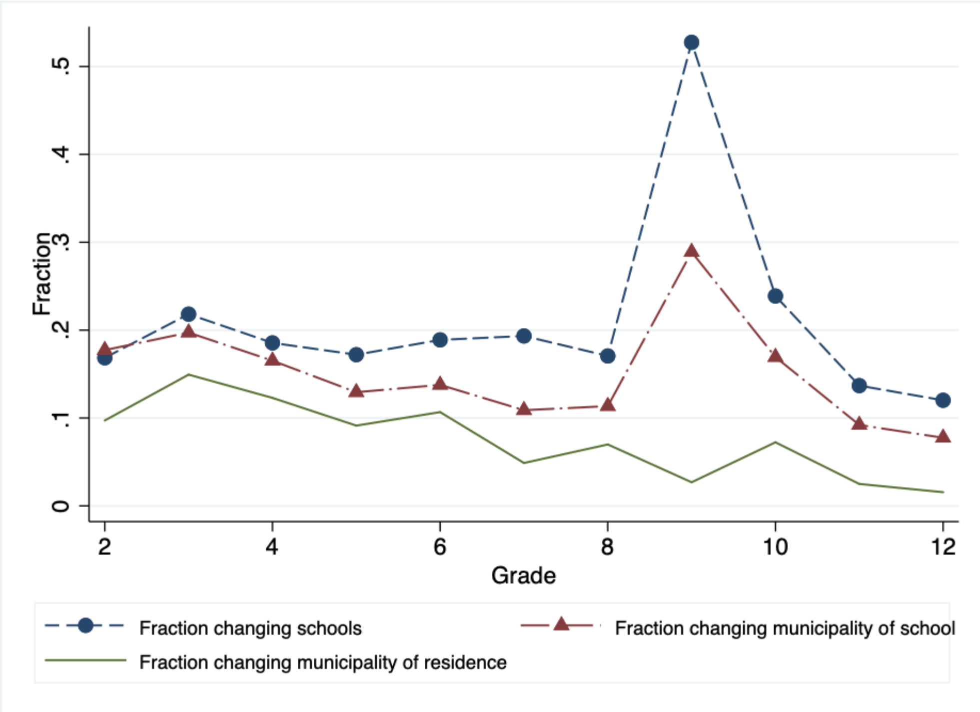


Figure B.2: Students changing schools vs changing municipality by grade. Santiago Metropolitan Area

To begin with, the average proportion of students changing their municipalities of residence annually is roughly 7.5 percent, and this percentage diminishes as their schooling progresses. In contrast, the mean proportion of students switching schools over the course of their school life is 21 percent per year, whereas the proportion of students changing the municipality of their school averages 15 percent per year (twice the proportion of those changing their municipality of residence). It is noteworthy that a marked spike occurs in grade 9 (the first year of secondary school), where over 50 percent of students change schools and about 30 percent of students change the municipality of their school. However, despite this considerable shift in the proportion of students changing municipalities for their schools, there is no corresponding surge in the proportion of students relocating to another municipality. This observation suggests that the likelihood of municipality changes during the school years is more heavily influenced by demographic factors, such as parents' age and their prospects of homeownership, rather than families actively choosing schools.

To strengthen our argument, we try to formally demonstrate that individuals don't sort themselves into different municipalities around the cutoff. If sorting were happening, we would expect to see differences between municipalities on each side of the cutoff. To investigate this, we select specific characteristics of municipalities and check for any abrupt changes in these characteristics, similar to how we looked at individual factors.

We focus on average family income, poverty rate,[43] and the average SIMCE (Math) score for municipalities in the Santiago Metropolitan Area. In Table B.3, we present the p-values for different mathematical specifications (ranging from linear to cubic) and three time intervals (5, 10, and 15 days). Our results don't provide enough evidence to reject the idea that there are no significant differences between municipalities at the 5% level of significance, and only one result is marginally significant at the 10% level.

	Avg. HH Income			Poverty rate			Avg. SIMCE Math		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$g(b_i)$ of degree 3	0.639	0.134	0.920	0.884	0.734	0.568	0.970	0.331	0.877
$g(b_i)$ of degree 2	0.851	0.843	0.599	0.983	0.892	0.999	0.637	0.427	0.949
$g(b_i)$ of degree 1	0.287	0.946	0.546	0.361	0.763	0.519	0.083	0.438	0.973
Bandwidth (days)	15	10	5	15	10	5	15	10	5
Observations	38,518	27,404	14,919	38,518	27,404	14,919	38,518	27,404	14,919

Note: For each of the variables (w_i), reported in the first row, we run the regression, $w_i = \eta_{wh} + \phi_b + \gamma * 1\{b_i > 0\} + g(b_i) + v_{it}$. $1\{b_i = BD_i - C^s > 0\}$ is an indicator variable taken a value of one for sudents whose birthday (BD_i) is over the cutoff (C), and zero otherwise. 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 $H0 : \gamma = 0$, that is, there are not differences in the variables at the municipality level between children over and below the July cutoff.

Table B.3: Differences in variables at the municipality level: Santiago Metropolitan Area

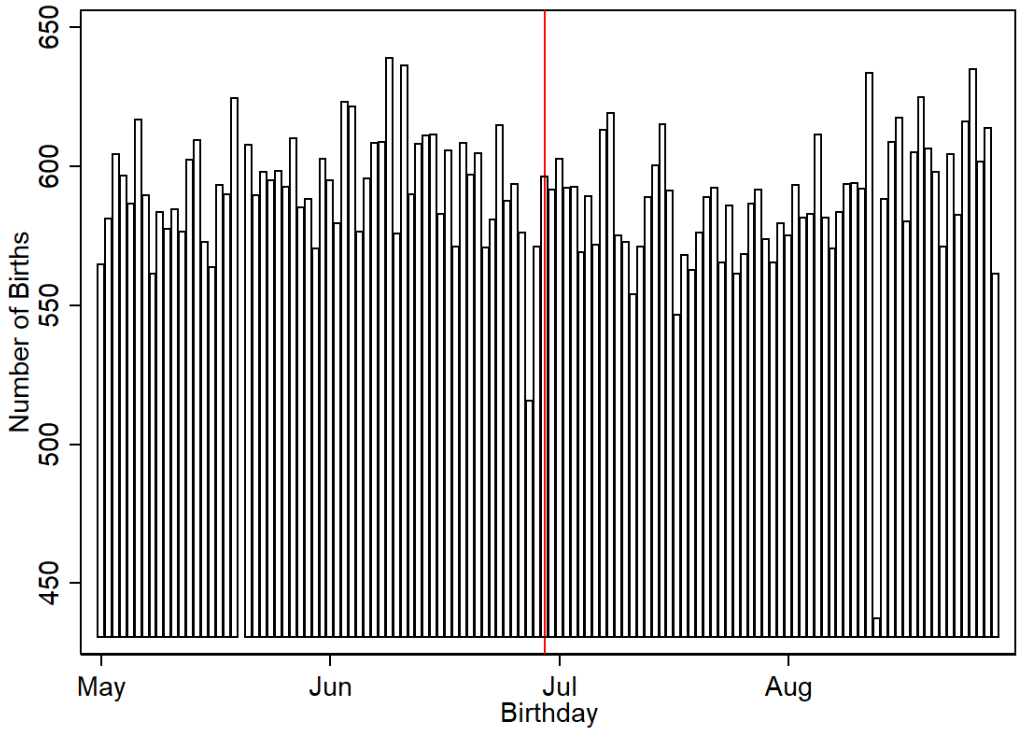
B.2 Manipulation of the running variable

The success of randomizing the treatment around the points of discontinuity hinges on the assumption that families cannot precisely control their children's birthdates. In simpler terms, the credibility

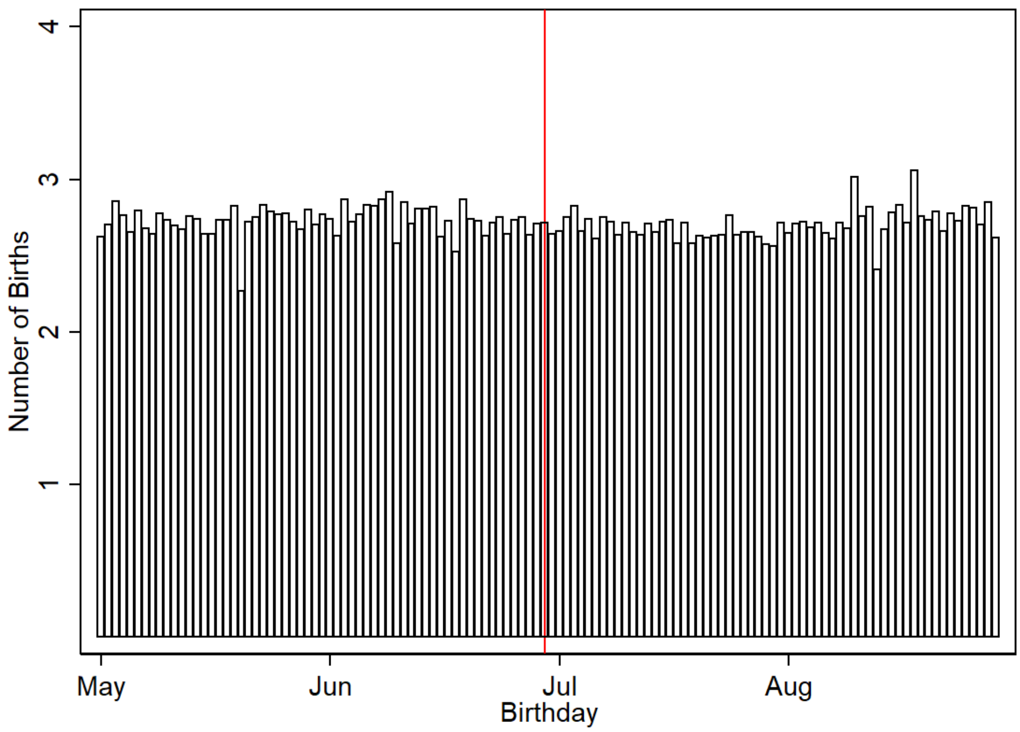
[43]The information for these two variables is taken from *Encuesta CASEN 2006. Ministerio de Desarrollo Social y Familia.* Available from https://datasocial.ministeriodesarrollosocial.gob.cl/portalDataSocial/descarga

of a Regression Discontinuity (RD) design could be compromised if individuals have the ability to manipulate the running variable (Lee and Lemieux, 2010). Given that the rules regarding the minimum age for school entry are publicly known, it's plausible that the potential advantages or disadvantages associated with delaying a child's entry into school might influence some families to strategically choose the birth season. Additionally, it's noteworthy that Chile, along with Turkey and Mexico, is among the countries with the highest rates of caesarean sections (c-sections) globally.[44] These two facts suggest some power to select the running variable.

[44] 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 B.3: Raw and conditional histogram for the day of birth

Nevertheless, the possibility of families arranging their choices throughout the calendar year does not invalidate the quasi-experimental design. The crucial assumption for identification is that individuals lack the capacity to precisely organize themselves around these points of discontinuity. If precise manipulation were at play, we would observe data clustering around these points, leading to a discontinuous distribution for the day of birth (the variable we are studying).

Panel (a) of Figure B.3 displays the original histogram for the day of birth, encompassing individuals born between December 15th and July 6th.[45] Despite considerable fluctuations, the figure conceals a fairly even distribution of births throughout the calendar year, with significant dispersion concerning weekdays and birth years. More specifically, we note an average of 650 births daily, varying between 500 and 800 births. However, when we factor in fixed effects for the day of the week, holidays, and birth years, as seen in Panel (b) of Figure B.3, we observe a uniform spread of births over the calendar year and no apparent disruptions in the distribution of the variable of interest across the discontinuity points. As a result, within a specific municipality and birth year, we observe an approximate average of 3 births per day.[46]

In line with the methodology introduced by McCrary (2008), we execute a formal assessment to ascertain any discontinuity in the distribution of the running variable. To achieve this, we carry out a regression where the frequency of birthdays across the calendar year serves as the dependent variable, mirroring the approach taken with the previous three predetermined variables.[47] These results are documented in Table B.4, with examinations conducted across all the cutoffs (column 1) and specifically for July's cutoff (column 2). Only in the case of one particular specification (quadratic polynomial) and a specific bandwidth (five days), do we reject the null at a significance level of 5%. This implies that parents might be able to select a certain week or month for their child's birth, but precision down to the exact birth date seems unlikely.

Finally, we perform a test for manipulation of the running variable that is consistent when the running variable is discrete, as proposed by Frandsen (2017). The obtained p-value is 0.761, indicating that we cannot reject the null hypothesis of no manipulation.[48]

[45] Data on births was obtained from *Departamento de Estadisticas e Información de Salud*, Ministerio de Salud de Chile, https://deis.minsal.cl/#datosabiertos

[46] Notice in Figure B.3b 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.

[47] See footnote 42.

[48] We employ the Stata command "rddisttestk" with parameter $k = 0$ to conduct the test. This parameter sets the maximum degree of nonlinearity in the probability mass function that is considered compatible with no manipulation. By choosing $k = 0$, we are stringent in detecting even small deviations from linearity, leading the test to reject with high probability.

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

Bandwidth		
	Panel A. $g(b_i)$ of degree 3	
15 days	0.219	0.076
10 days	0.260	0.499
5 days	0.318	0.122
	Panel B. $g(b_i)$ of degree 2	
15 days	0.352	0.114
10 days	0.195	0.063
5 days	0.891	0.008
	Panel C. $g(b_i)$ 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, $w_i = \alpha_s + \eta_{wh} + \phi_b + \gamma^s * 1\{b_i - C >\} + g(b_i) + v_{it}$ 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.

B.3 Monotonicity

The monotonicity assumption is essential because it allows us to interpret the estimated parameter as a Local Average Treatment Effect (LATE). In the context of our paper, the monotonicity assumption implies that a change in the instrument's value (in our case, being born after the July cutoff) should increase the school starting age (SSA) for all students or, at the very least, not decrease it.

The need for the monotonicity assumption arises due to the presence of heterogeneity, combined with students strategically choosing their SSA based on expected outcomes ("sorting on gains"). To gauge the extent of this heterogeneity, we include covariates that can reasonably serve as indicators for income, such as parental education, and specifically within the Chilean context, the municipality of residence. If violations of monotonicity were a substantial concern in our analysis, we would expect the estimates to show sensitivity when covariates like these are included (Dhuey et al., 2019). Nevertheless, our estimates are generally insensitive to the incorporation of these covariates (see Appendix C.2). This observation suggests that potential breaches of monotonicity are unlikely to significantly affect our results.

However, we acknowledge that we cannot rule out the presence of heterogeneity and sorting on gains. Therefore, we will now examine the plausibility of the monotonicity assumption in relation to our paper, assuming independence between the instrument and the potential treatment assignment.

	Born before cutoff	Born after cutoff
Early entry	0.000 (0.00)	0.006 (0.08)
On Time entry	0.500 (0.50)	0.994 (0.08)
Late entry	0.500 (0.50)	0.000 (0.00)

Table B.5: Fraction of students starting school on time, early or late by month of birth

Following the terminology in Fiorini and Stevens (2021), the term "On time" is used to describe children who start school as soon as they become eligible. "Late Entry" refers to the situation where children enter school one year later than their eligibility, while "Early Entry" pertains to children who begin school one year before becoming eligible. Consequently, if all children in our sample started "On time," we would expect to observe an increase in SSA from approximately 5.7 years for June-born children to 6.7 years for July-born children, thereby upholding the assumption of monotonicity.

As highlighted by Barua and Lang (2016), monotonicity is violated when there are defiers—

children who start early only when they are not allowed to do so. Table B.5 outlines the distribution of students in the Early, Late, and On-time groups on both sides of the cutoff in our working sample. Notably, we find no instances of June-born children starting school before their first eligible year (which would be at age 4.7), nor do we observe July-born children starting school a year later than their first eligible year (they should have entered at age which would be at age 7.7). This leaves us with four counterfactual scenarios:

(i) June-born students starting "On time" who would also start "On time" if they were born in July (resulting in an increase in SSA from approximately 5.7 to 6.7 at the cutoff).

(ii) June-born children starting "Late" but would start "On time" if they were born in July (resulting in a slight decrease in SSA).

(iii) June-born children starting "Late" but would start "Early" if they were born in July (leading to a decrease in SSA from 6.7 to 5.7).

(iv) June-born children starting "On time" but would start "Early" if they were born in July (also leading to a decrease in SSA from 6.7 to 5.7).

Monotonicity is violated if, in addition to the complier group (i), we have students whose SSA diminishes as the instrument changes. As shown in Table B.5, approximately half of the June-born students are redshirted (i.e., they delay their school entry). In contrast, almost all July-born children start school on time (99.4%). With this information in mind, we can rule out scenarios (iii) and (iv), both of which entail a year's decrease in SSA. Consequently, our primary concern regarding monotonicity pertains to scenario (ii), which we cannot dismiss. As shown in Table B.5, both the complier group (i) and the defier group (ii) are potentially of similar size: $p_{complier} = p_{defier} = 0.5$.

While we cannot conclusively rule out the possibility of monotonicity violation, the utilization of precise birth dates within the framework of a fuzzy RD design serves to alleviate potential concerns associated with this assumption. For instance, in a 1-day discontinuity sample, the complier group (i) would observe an increase in SSA at the cutoff by 364/365 years, while the defier group (ii) would encounter a decrease in SSA by 1/365 years. Within the context of essential heterogeneity, the annual benefits derived from SSA would vary between these two groups ($\bar{\beta}_{complier} \neq \bar{\beta}_{defier}$). Following the approach outlined by Fiorini and Stevens (2021):

$$\beta^{RD}\left(July\ 1^{st}, June\ 30^{th}\right) = \frac{\frac{364}{365}\bar{\beta}_{complier}p_{complier} - \frac{1}{365}\bar{\beta}_{defier}p_{defier}}{\frac{364}{365}p_{complier} - \frac{1}{365}p_{defier}} \approx \bar{\beta}_{complier}$$

This approximation arises from the fact, provided both groups are of equal size, that the complier group counts 364 times more than the hypothetical defier group. Even when using a larger discontinuity sample, as we do in our paper, the inclusion of flexible function on the day of birth enables us to capture the effect at the discontinuity. In summary, while we cannot definitively rule out a violation of monotonicity, we have reasonable confidence that the parameter we are estimating in our fuzzy RD design closely approximates the Local Average Treatment Effect (LATE) of interest.

C Robustness checks

C.1 CER optimal bandwidth

Table C.6: Impact of SSA on school characteristics. CER-optimal bandwidth

	Avg. SIMCE	Above Percentile 25	Above Percentile 50	Above Percentile 75	Voucher	Academic Selection
	[1]	[2]	[3]	[4]	[5]	[6]
Panel A: Complete Sample						
	0.107***	0.033***	0.053***	0.053***	0.038***	0.073***
	(0.019)	(0.011)	(0.006)	(0.013)	(0.012)	(0.013)
Observations	53796	54209	54209	54209	68393	20765
Mean	.31	.74	.49	.24	.52	.65
Weak Instr. (a)	3030.7	3215.8	3215.8	3215.8	5072.70	48.40
Panel B: By parents' educational level						
12 years of education or less						
	0.116***	0.039**	0.050***	0.038***	0.065***	0.086***
	(0.025)	(0.019)	(0.009)	(0.015)	(0.012)	(0.017)
Observations	29603	29824	29824	29824	37678	11024
Mean	.01	.65	.35	.13	.46	.58
Weak Instr. (a)	4613.5	5176.6	5176.6	5176.6	8908.6	55.90
More than 12 years of education						
	0.098***	0.027***	0.059***	0.077***	0.001	0.049
	(0.025)	(0.009)	(0.013)	(0.019)	(0.021)	(0.031)
Observations	24193	24385	24385	24385	30715	9741
Mean	.67	.86	.66	.39	.61	.73
Weak Instr. (a)	1147.4	1124.5	1124.5	1124.5	1617.8	37.8
Panel C: By students' gender						
Boys						
	0.113***	0.018	0.071***	0.060**	0.020	0.078***
	(0.036)	(0.013)	(0.013)	(0.028)	(0.013)	(0.019)
Observations	27281	27522	27522	27522	34728	10576
Mean	.29	.74	.48	.24	.52	.64
Weak Instr. (a)	2813.2	2855	2855	2855	4044.5	46.8
Girls						
	0.112***	0.051***	0.040***	0.048***	0.054***	0.073***
	(0.017)	(0.017)	(0.010)	(0.014)	(0.013)	(0.015)
Observations	26515	26687	26687	26687	33665	10189
Mean	.32	.75	.5	.25	.53	.65
Weak Instr. (a)	2053.2	2310.1	2310.1	2310.1	3501.3	47.7

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered by day of birth.

Additional controls: municipality, year of birth, parents' education, gender, weekday of birth, and born on a holiday dummies. For each of the outcomes, the CER-optimal bandwidth is reported in Table A.1. (a) Effective F statistic (Montiel-Pflueger robust weak instrument test). Critical values for 5, 10 and 20 percent maximal IV bias are 37.41, 23.11, and 15.06, respectively.

Table C.7: Impact of SSA on school Inputs. CER-optimal bandwidth.

	Avg. Parents' Education	Avg. Classmates' SSA	Teachers			
			Age	Full-time	Teaching / Total hours	4-year College
	[1]	[2]	[3]	[4]	[5]	[6]
Panel A: Complete Sample						
	0.264***	0.063***	-0.531***	0.015***	-0.010**	0.018**
	(0.036)	(0.006)	(0.159)	(0.005)	(0.005)	(0.008)
Observations	61251	68084	63841	50582	50582	36659
Mean	11.24	6.27	45.9	.32	.82	.60
Weak Instr. (a)	3948.2	5018.6	1790	1279.2	1279.2	1063.7
Panel B: By parents' educational level						
12 years of education or less						
	0.312***	0.061***	-0.716**	0.008	-0.013**	0.018**
	(0.040)	(0.006)	(0.310)	(0.009)	(0.005)	(0.008)
Observations	33683	37406	35076	27764	27764	21137
Mean	10.31	6.26	46.8	.32	.83	.60
Weak Instr. (a)	7577.70	9132.70	2478.2	1841.9	1841.9	1146.7
More than 12 years of education						
	0.203***	0.068***	-0.308	0.029***	-0.009	0.014
	(0.063)	(0.007)	(0.215)	(0.007)	(0.012)	(0.012)
Observations	27568	30678	28765	22818	22818	15522
Mean	12.41	6.29	44.77	.32	.81	.70
Weak Instr. (a)	1245.8	1608.8	758.80	590.9	590.9	701.5
Panel C: By students' gender						
Boys						
	0.221***	0.063***	-0.541***	0.017*	-0.014*	0.021
	(0.064)	(0.006)	(0.207)	(0.009)	(0.008)	(0.013)
Observations	31053	34564	32464	25725	25725	18656
Mean	11.22	6.27	45.87	.32	.82	.60
Weak Instr. (a)	3251.8	3734.2	1287.5	886.5	886.5	1247.5
Girls						
	0.303***	0.065***	-0.509**	0.012	-0.005	0.012
	(0.044)	(0.006)	(0.224)	(0.008)	(0.006)	(0.012)
Observations	30198	33520	31377	24857	24857	18003
Mean	11.27	6.27	45.92	.31	.82	.60
Weak Instr. (a)	2942.3	3761.3	1898.4	1582.7	1582.7	700.7

*** $p<0.01$, ** $p<0.05$, * $p<0.1$. Standard errors clustered by day of birth.

Additional controls: municipality, year of birth, parents' education, gender, weekday of birth, and born on a holiday dummies. For each of the outcomes, the CER-optimal bandwidth is reported in Table A.1.
(a) Effective F statistic (Montiel-Pflueger robust weak instrument test). Critical values for 5, 10 and 20 percent maximal IV bias are 37.41, 23.11, and 15.06, respectively.

C.2 Robustness analysis for the first stage and reduced form

C.2.1 First Stage. Impact of minimum age requirements on the probability to start school older.

Table C.8 show the sensitivity of the first stage estimates to three different bandwidths (5, 10 and 15 days) for quadratic and cubic polynomials. Notice that the results are robust to changes in both dimensions.

Table C.8: First stage estimates. Impact of age eligibility requirement on the probability of entry at an older age. Sensitivity to the degree of the polynomial specification.

	Bandwidth					
	5 days		10 days		15 days	
	[1]	[2]	[3]	[4]	[5]	[6]
Panel A:			**Complete sample**			
	0.447***	0.409***	0.467***	0.453***	0.471***	0.462***
	(0.008)	(0.009)	(0.011)	(0.012)	(0.010)	(0.013)
Observations	40791	40791	75291	75291	105831	105831
F excluded instr.	2823.3	2047.1	1865.1	1446.1	2114.4	1301.9
Panel B:			**By Parents' educational level**			
			12 years of education or less			
	0.496***	0.437***	0.504***	0.503***	0.513***	0.503***
	(0.012)	(0.012)	(0.009)	(0.014)	(0.010)	(0.013)
Observations	22496	22496	41483	41483	58301	58301
F excluded instr.	1771.5	1297.7	2886.5	1295.8	2414.1	1449
			More than 12 years of education			
	0.383***	0.367***	0.418***	0.386***	0.417***	0.407***
	(0.008)	(0.014)	(0.017)	(0.013)	(0.013)	(0.016)
Observations	18295	18295	33808	33808	47530	47530
F excluded instr.	2376.2	647.90	622.30	927.6	1054.4	686.2
Panel C:			**By Student's gender**			
			Boys			
	0.430***	0.388***	0.439***	0.439***	0.449***	0.433***
	(0.012)	(0.024)	(0.010)	(0.016)	(0.011)	(0.011)
Observations	20790	20790	38240	38240	53696	53696
F excluded instr.	1205.7	262.6	1837.4	757.5	1754.7	1436.6
			Girls			
	0.462***	0.432***	0.497***	0.465***	0.494***	0.494***
	(0.008)	(0.007)	(0.014)	(0.010)	(0.012)	(0.017)
Observations	20001	20001	37051	37051	52135	52135
F excluded instr.	3154.1	3368.2	1243.8	2318.2	1795.7	847.90
Degree Pol.	2	3	2	3	2	3

*** $p<0.01$, ** $p<0.05$, * $p<0.1$. Standard errors clustered by day of birth.
The dependent variable, *Older*, is a dummy variable that takes a value one for children who start primary school the year they turn seven. Specifications with additional controls include municipality, year of birth, parents' education, gender, weekday of birth, and born on a holiday dummies.

C.2.2 Reduced form. Impact of minimum age requirements on selected outcomes

Tables C.9 and C.10 show the reduced form estimates from the regression:

$$y_{ict} = \rho_t + \rho_c + X_i^{'}\psi + \delta \times 1\{b - C > 0\} + g(b_i) + v_i,$$

where y_{ict} represents one of the outcomes for student i, born the day b of year t, and living in municipality c. Here δ, the parameter of interest, captures the discontinuity around the cutoff, $1\{*\}$ is an indicator operator, and C the July cutoff.

Table C.9: Reduced form anaysisis: School characteristics

	Avg. SIMCE	Above Percentile 25	Above Percentile 50	Above Percentile 75	Voucher	Academic Selection
	[1]	[2]	[3]	[4]	[5]	[6]
			5 days bandwidths			
with controls						
	0.042***	0.007	0.023***	0.013**	0.027***	0.016**
	(0.012)	(0.005)	(0.003)	(0.005)	(0.007)	(0.006)
Observations	33590	33855	33855	33855	33855	14836
Mean	.33	.75	.5	.26	.53	.65
without controls						
	0.064***	0.017***	0.032***	0.022**	0.020***	-0.004
	(0.013)	(0.005)	(0.005)	(0.007)	(0.005)	(0.011)
Observations	33644	33908	33908	33908	33908	14863
Mean	.33	.75	.5	.26	.53	.65
			10 days bandwidths			
with controls						
	0.042***	0.011**	0.022***	0.023***	0.018***	0.020***
	(0.009)	(0.005)	(0.004)	(0.004)	(0.006)	(0.005)
Observations	67870	68390	68390	68390	68390	30000
Mean	.33	.75	.5	.26	.53	.65
without controls						
	0.053***	0.016**	0.024***	0.027***	0.013*	0.006
	(0.013)	(0.006)	(0.007)	(0.006)	(0.007)	(0.009)
Observations	67986	68507	68507	68507	68507	30048
Mean	.33	.75	.5	.26	.53	.65
			15 days bandwidths			
with controls						
	0.035***	0.012**	0.018***	0.017***	0.016***	0.018***
	(0.008)	(0.005)	(0.003)	(0.004)	(0.004)	(0.004)
Observations	101595	102338	102338	102338	102338	44283
Mean	.33	.75	.5	.26	.53	.65
without controls						
	0.052***	0.018***	0.023***	0.023***	0.012**	0.005
	(0.011)	(0.005)	(0.005)	(0.005)	(0.005)	(0.007)
Observations	101757	102503	102503	102503	102503	44353
Mean	.33	.75	.5	.26	.53	.65

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered by day of birth.

The specification without controls comprises a linear specification based on the day of birth, which can differ on each side of the threshold, and year of birth fixed effects. The specification with controls further incorporates variables such as municipality, parents' education, gender, weekday of birth, and a dummy variable for being born on a holiday.

Table C.10: Reduced form anaysisis: School Inputs

	Avg. Parents' Education	Avg. Classmates' SSA	Teachers: Age	Teachers: Full-time	Teachers: Teaching / Total hours	Teachers: 4-year College
	[1]	[2]	[3]	[4]	[5]	[6]
			5 days bandwidths			
with controls						
	0.103***	0.026***	-0.128*	0.003	0.001	0.009*
	(0.025)	(0.006)	(0.068)	(0.002)	(0.001)	(0.004)
Observations	33852	33714	31548	31549	31549	26113
Mean	11.31	6.29	45.84	.32	.82	.64
without controls						
	0.139**	0.026***	-0.085	-0.000	-0.000	0.008**
	(0.046)	(0.004)	(0.092)	(0.002)	(0.001)	(0.003)
Observations	33905	33769	31601	31602	31602	26159
Mean	11.31	6.29	45.84	.32	.82	.64
			10 days bandwidths			
with controls						
	0.115***	0.030***	-0.222***	0.005**	-0.006**	0.009**
	(0.018)	(0.003)	(0.071)	(0.002)	(0.002)	(0.003)
Observations	68381	68081	63838	63840	63840	52797
Mean	11.31	6.29	45.84	.32	.82	.64
without controls						
	0.133***	0.028***	-0.137*	0.004*	-0.007***	0.008***
	(0.037)	(0.003)	(0.075)	(0.002)	(0.002)	(0.003)
Observations	68498	68197	63954	63956	63956	52890
Mean	11.31	6.29	45.84	.32	.82	.64
			15 days bandwidths			
with controls						
	0.098***	0.030***	-0.259***	0.004	-0.004**	0.012***
	(0.017)	(0.003)	(0.065)	(0.002)	(0.002)	(0.003)
Observations	102328	101901	95457	95461	95461	79042
Mean	11.31	6.29	45.84	.32	.82	.64
without controls						
	0.141***	0.029***	-0.221***	0.004**	-0.005***	0.011***
	(0.032)	(0.002)	(0.076)	(0.002)	(0.002)	(0.003)
Observations	102493	102064	95618	95622	95622	79171
Mean	11.31	6.29	45.84	.32	.82	.64

*** $p<0.01$, ** $p<0.05$, * $p<0.1$. Standard errors clustered by day of birth.

The specification without controls comprises a linear specification based on the day of birth, which can differ on each side of the threshold, and year of birth fixed effects. The specification with controls further incorporates variables such as municipality, parents' education, gender, weekday of birth, and a dummy variable for being born on a holiday.

C.3 Results using January-July cutoffs

Table C.11: Impact of SSA on school characteristics. February-July cutoffs.

	Avg. SIMCE	Above Percentile 25	Above Percentile 50	Above Percentile 75	Voucher	Academic Selection
	[1]	[2]	[3]	[4]	[5]	[6]
RD estimates for two data driven bandwidths						
MSE-optimal bandwidth						
	0.055	-0.018	0.047**	0.040**	0.039**	0.033
	(0.037)	(0.021)	(0.019)	(0.019)	(0.017)	(0.020)
Observations	409484	412476	412476	412476	535654	119068
Mean	.32	.75	.5	.25	.52	.64
Weak Instr. (a)	32.8	32.8	32.8	32.8	43.1	28.5
CER-optimal bandwidth						
	0.096**	-0.003	0.055***	0.059***	0.037*	0.023
	(0.040)	(0.022)	(0.020)	(0.021)	(0.020)	(0.023)
Observations	326785	329183	329183	329183	412476	104030
Mean	.32	.75	.5	.25	.52	.64
Weak Instr. (a)	26.4	26.3	26.3	26.3	32.8	25.3

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered by day of birth.

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 A.1.
(a) Effective F statistic (Montiel-Pflueger robust weak instrument test). Critical values for 5, 10 and 20 percent maximal IV bias are 37.41, 23.11, and 15.06, respectively.

Table C.12: Impact of SSA on school Inputs. February-July cutoffs cutoffs.

	Avg. Parents' Education	Avg. Classmates' SSA	Teachers			
			Age	Full-time	Teaching / Total hours	4-year College
	[1]	[2]	[3]	[4]	[5]	[6]
RD estimates for two data driven bandwidth						
MSE-optimal bandwidth						
	0.067	0.050***	-0.592**	0.022**	-0.013	0.042***
	(0.075)	(0.006)	(0.250)	(0.010)	(0.008)	(0.015)
Observations	452224	492475	458208	352884	352884	235349
Mean	11.28	6.26	45.93	.32	.82	.60
Weak Instr. (a)	35.8	39.3	48.7	36.8	36.8	26.7
CER-optimal bandwidth						
	0.106	0.050***	-0.450*	0.033***	-0.013	0.037**
	(0.083)	(0.006)	(0.267)	(0.013)	(0.009)	(0.016)
Observations	370405	410921	352868	281475	281475	206103
Mean	11.28	6.26	45.93	.32	.82	.60
Weak Instr. (a)	29.4	32.9	36.8	29.2	29.2	23.1

*** $p<0.01$, ** $p<0.05$, * $p<0.1$. Standard errors clustered by day of birth.

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 A.1.
(a) Effective F statistic (Montiel-Pflueger robust weak instrument test). Critical values for 5, 10 and 20 percent maximal IV bias are 37.41, 23.11, and 15.06, respectively.

C.4 Placebo analysis. Reduced Form estimates for a threshold at the 1^{st} of September or 1^{st} of October.

Table C.13: Impact of SSA on school characteristics. Placebo analysis.

	Avg. SIMCE	Above Percentile 25	Above Percentile 50	Above Percentile 75	Voucher	Academic Selection
	[1]	[2]	[3]	[4]	[5]	[6]
Panel A: September and October 1st cutoffs. 10 days bandwidth; without controls.						
	0.000	-0.005	-0.004	0.001	-0.002	-0.007*
	(0.014)	(0.005)	(0.006)	(0.006)	(0.005)	(0.004)
Observations	169504	171155	171155	171155	171155	67840
Mean	.36	.76	.51	.26	.52	.65
Panel B: September and October 1st cutoffs. 10 days bandwidth; with controls.						
	0.002	-0.005	-0.004	0.002	-0.003	-0.003
	(0.008)	(0.004)	(0.004)	(0.003)	(0.004)	(0.002)
Observations	169286	170935	170935	170935	170935	67748
Mean	.36	.76	.51	.26	.52	.65

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered by day of birth.

The specification without controls comprises a linear specification based on the day of birth, which can differ on each side of the threshold, and year of birth fixed effects. The specification with controls further incorporates variables such as municipality, parents' education, gender, weekday of birth, and a dummy variable for being born on a holiday.

Table C.14: Impact of SSA on school Inputs. Placebo analysis.

	Avg. Parents' Education	Avg. Classmates' SSA	Teachers			
			Age	Full-time	Teaching / Total hours	4-year College
Panel A: September and October 1st cutoffs. 10 days bandwidth; without controls.						
	0.001	-0.002	0.089	-0.001	-0.000	-0.008**
	(0.036)	(0.003)	(0.076)	(0.003)	(0.002)	(0.003)
Observations	171135	170413	143634	143637	143637	118364
Mean	11.34	6.28	45.78	.32	.82	.64
Panel B: September and October 1st cutoffs. 10 days bandwidth; with controls.						
	0.004	0.000	0.065	-0.001	-0.001	-0.006*
	(0.017)	(0.001)	(0.054)	(0.002)	(0.001)	(0.003)
Observations	170915	170194	143419	143422	143422	118196
Mean	11.34	6.28	45.78	.32	.82	.64

*** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered by day of birth.

The specification without controls comprises a linear specification based on the day of birth, which can differ on each side of the threshold, and year of birth fixed effects. The specification with controls further incorporates variables such as municipality, parents' education, gender, weekday of birth, and a dummy variable for being born on a holiday.