Time Varying Effects of Elite Schools: Evidence from Mexico City *

Salvador Navarro † Marco Pariguana ‡

June 5, 2026

Abstract

We examine whether the academic effects of marginal admission to elite science high schools vary by admission year, reflecting changes in school quality over time. Using administrative data from Mexico City's centralized high school admission system between 2005 and 2009, we estimate year-specific regression discontinuity designs. We find that the effect on end-of-high-school math test scores declines steadily over the period: positive and significant in 2005, but statistically insignificant by 2009; while the effect on dropout is negative, sizeable, and relatively stable across cohorts. This decline is not explained by changes in peer quality or other observable school inputs, which do not change systematically over time. Instead, we document a reduction in the value-added of elite schools, suggesting that changes in the productivity of school inputs drive the trend. Our findings highlight the time-specific nature of treatment effects and the limits of external validity, even in internally valid designs.

Keywords: School choice, Upper-secondary education, Education policy.

JEL codes: I21, I24, I28, J24.

*We thank seminar participants at El Colegio de Mexico, the Society of Labor Economics Conference, the University of Western Ontario, and the University of Wisconsin for valuable feedback and suggestions. All remaining errors are our own. Navarro acknowledges support from the Social Sciences and Humanities Research Council of Canada.

†Corresponding author. Department of Economics, University of Western Ontario. E-mail: snavarr@uwo.ca.

‡WZB Berlin Social Science Center

1 Introduction

In centralized education systems, high-performing schools, hereafter elite schools, attract a considerable share of applicants. These students often expect that elite schools' advantages, such as better peers, superior infrastructure, and more qualified teachers, will translate into improved academic outcomes and, ultimately, better outcomes overall. Yet whether these expected benefits actually materialize, and under what conditions, remains unclear.

Recent research increasingly recognizes that the effects of elite schools may be heterogeneous. These studies have produced a range of estimates, reflecting differences in institutional context, student composition, and policy environment (e.g., Pop-Eleches and Urquiola (2013); Abdulkadiroğlu et al. (2014); Dobbie and Fryer Jr (2014); De Groote and Declercq (2021); Beuermann and Jackson (2022)). The literature recognizes that changes in school characteristics or in how those characteristics affect student outcomes can limit the generalizability of even internally valid causal estimates across periods or policy settings. For example, a policymaker considering whether to expand elite school capacity may find that past estimates reflect conditions or mechanisms that no longer apply.

This paper systematically examines whether, and why, the estimated effects of elite schools on academic outcomes vary across time. We study a fixed set of elite schools in Mexico City between 2005 and 2009, leveraging rich administrative and school census data from the city's centralized high school admission system. By combining a sharp regression discontinuity design (RDD) with year-specific estimates, we trace changes in school effects over time and explore potential mechanisms underlying these trends. Our findings raise questions about the external validity of causal estimates in education research when time heterogeneity is not taken into account.

Estimating the impact of elite schools is complicated by the fact that admission is nonrandom: students who attend elite schools may differ systematically from those who do not. A growing literature addresses this selection problem using RDDs that exploit centralized admissions systems, where oversubscribed schools generate score-based cutoffs that effectively randomize access among students at the margin. Comparing outcomes of marginally admitted and rejected students yields credible causal estimates, though typically only near the cutoff and for a specific cohort.

Yet many existing studies pool multiple cohorts to improve precision, often treating school effects as stable over time. However, if the relative quality of elite and non-elite schools shifts from year to year, such pooling may obscure meaningful variation and limit the relevance of findings for current policy decisions. Our analysis shows that this concern is not merely theoretical: even within a single market and institutional framework, the effects of marginal admission to elite schools can change substantially over a short horizon.

In particular, we examine the case of Mexico City's centralized high school admission system from 2005 to 2009. Each year, approximately 250,000 students participate in a centralized matching process that assigns them to public high schools based on their preferences and a standardized admission exam score. The system's structure and stability across years provides a rare opportunity to estimate year-specific school effects within a consistent institutional framework. However, during this period, Mexico introduced the *Reforma Integral de la Educación Media Superior* (RIEMS), a national upper-secondary reform that created a common curricular framework and a national baccalaureate system with the stated goal of aligning academic requirements across subsystems. In Section 2, we document the timing and stated objectives of this reform and discuss how it may have contributed to the evolution of elite schools' relative effectiveness. We cannot observe classroom-level curricula directly, so we cautiously interpret the reform as a plausible mechanism rather than as a separately identified causal effect.

Our dataset combines three key administrative sources. First, we use application and placement records to identify students' admission outcomes. Second, we track academic performance using standardized exit exam scores taken in the final quarter of high school. Third, we use yearly school census data to measure a wide range of school inputs, including peer composition, teacher qualifications, and infrastructure indicators. This rich, panel-like dataset allows us to estimate the impact of elite school admission on test scores and to investigate mechanisms driving any changes over time.

Three features of this setting enhance our empirical strategy. First, the centralized assignment process generates score-based cutoffs for oversubscribed elite schools, enabling a regression discontinuity design. Second, the large number of applicants allows us to estimate separate effects for each cohort with sufficient precision. Third, the availability of yearly input data enables us to explore how school-level changes correlate with the evolution of treatment effects. Finally, the repeated-cohort structure and stable assignment rules allow us to compare the same set of elite schools to similar next-best alternatives year after year, so that we can study time variation in treatment effects within a fixed market.

Our identification strategy builds on the use of Regression Discontinuity Designs (RDDs) in centralized school admission systems. Oversubscribed elite schools set implicit cutoffs based on students' composite scores, leading to clear discontinuities in the likelihood of admission. Students just above and just below each cutoff are plausibly similar in all respects except for school placement, allowing us to isolate the causal effect of elite school attendance for marginal applicants. We follow Kirkeboen et al. (2016), restricting attention to students whose local ranking places an elite school just above a non-elite alternative. For these students, our RDD compares outcomes for those who are marginally admitted to the elite school to outcomes for similar students who instead attend their next-best non-elite option.

Rather than pooling cohorts to obtain a single average effect, we estimate the RDD separately for each admission year from 2005 to 2009. This approach allows us to track whether and how the effects of elite school admission evolve over time. Given that the assignment mechanism and institutional setting remain stable during this period, any changes in estimated effects are likely attributable to shifts in school quality or in how school characteristics translate into academic outcomes.

We focus on end-of-high school math test scores as our primary outcome, complemented by secondary analysis of dropout rates. For the 2005 cohort, marginal admission to an elite science high school raises math scores by around 0.2 standard deviations for students at the cutoff, an effect similar in magnitude to previous estimates for Mexico City. Our analysis reveals that the academic benefits of elite school admission vary substantially across cohorts. The effect of marginal admission to an elite high school on students' math test scores is positive and statistically significant in 2005 but declines steadily, becoming statistically indistinguishable from zero by 2009. At the same time, marginal admission consistently increases the probability of not taking the exit exam, our measure of dropout, by roughly 6 to 10 percentage points across cohorts. In recent years of our sample, elite schools therefore expose marginal students to a higher risk of dropout without delivering measurable gains in test scores.

To understand these trends, we examine whether changes in peer quality or observable school inputs explain the time pattern. We find that elite schools continue to offer consistently better observable inputs, but the magnitude of these differences does not change substantially over time. Instead, most of the decline in the impact of elite schools appears to stem from a reduction in how effectively those inputs translate into academic gains, potentially reflecting shifts in unobserved school characteristics. One plausible explanation is the major curricular alignment reform during our study period, which may have reduced the comparative advantage of elite schools by standardizing academic content across institutions.

Our work contributes to two main strands of the literature. First, it contributes to the extensive literature that studies the academic effects of elite/selective schools.[1] Our analysis extends this research by providing year-specific estimates across five consecutive cohorts, revealing substantial within-market variation over time. This evidence complements prior work, which often attributes heterogeneity in elite school effects to differences across markets or institutional settings. Our results suggest that similar variation can arise within a single market as conditions evolve. The temporal variation in our estimated effect sizes mirrors the cross-market and cross-period variation documented in Beuermann and Jackson (2022). Within Mexico City, we also go beyond average effects by documenting how time trends look for more and less selective elite schools, showing that the decline in impacts is not driven by a small subset of institutions.

In the context of Mexico City, for example, Dustan et al. (2017) report large, positive, and statistically significant effects on math test scores for students marginally admitted to elite science schools based on two cohorts. Our analysis complements this work, by showing that the earlier positive effects are not stable: the impact of elite school admission declines and becomes statistically insignificant over time. Taken together, these findings imply that even within the same market, estimates from one or two cohorts may provide a misleading

[1] See Clark (2010); Jackson (2010); Pop-Eleches and Urquiola (2013); Abdulkadiroğlu et al. (2014); Dobbie and Fryer Jr (2014); Lucas and Mbiti (2014); Abdulkadiroğlu et al. (2017); Dustan et al. (2017); Beuermann and Jackson (2022); Angrist et al. (2023).

guide to the effects facing future students.

Second, we contribute to the literature on education production functions. As Todd and Wolpin (2003) highlight, policy effects such as those obtained using RDDs do not necessarily estimate production function parameters. Understanding whether and why elite school admission improves academic outcomes requires examining not just whether there is an effect, but also how school inputs interact to produce it. For example, some prior studies interpret the effect of marginal admission as primarily driven by improved peer quality (e.g., Jackson, 2010; Abdulkadiroğlu et al., 2014; Dobbie and Fryer Jr, 2014). However, in our setting, we observe a constant peer quality gain across cohorts, while the effect on test scores declines. This disconnect suggests that peer effects alone are not the primary driver of elite school impacts, or at least that their productivity may be changing over time. Our value-added decomposition shows that the part of school effects explained by observed inputs declines for elite schools relative to their main non-elite alternatives over time, even though the levels of those inputs remain relatively stable. This pattern is consistent with a change in how similar inputs map into achievement.

More broadly, following the framework of Altonji and Mansfield (2018), school effects depend on individual-level inputs, group-level inputs (such as peers and teachers), and the productivity of those inputs. As either the level or productivity of group-level inputs may vary across time and context, it follows that causal estimates from one setting or period may not generalize to another. Our findings underscore this point: even when school inputs appear stable on the surface, the way they translate into academic outcomes may shift due to changes in curricula, institutional incentives, or unobservable factors.

The remainder of the paper is organized as follows. Section 2 provides institutional background on Mexico City's high school admission system. Section 3 describes the data sources, sample construction, and outcome measures. Section 4 outlines the empirical strategy and presents evidence supporting the validity of the regression discontinuity design. Section 5 presents the main results. Sections 6 and 7 explore the mechanisms driving the time-varying effects. Section 8 concludes.

2 Institutional background

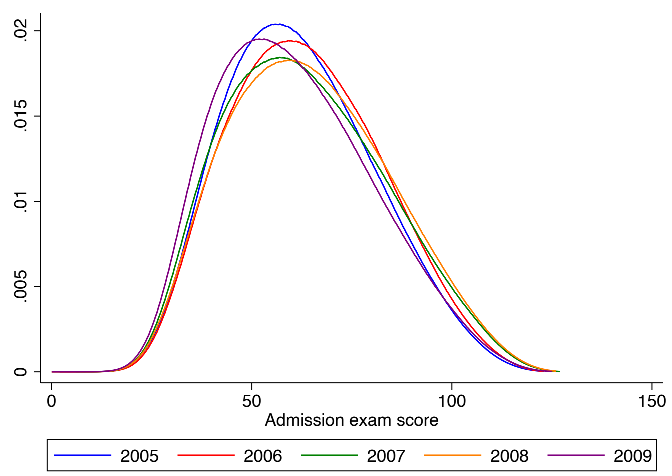
Elementary school in Mexico City is six years in length, middle school is three years, and high school is three years. The centralized high school admission process in Mexico City matches students with a middle school certificate to public high schools. Every year, the market has around 250,000 applicants applying for seats in around 600 high schools (see Appendix A). Importantly for our purposes, the admission process remained unchanged during our study period (2005-2009).

The timeline of the admission process is as follows. At the end of January, applicants receive a booklet that describes all the available public high schools in the market. Between late February and early March, students submit a rank-ordered list (ROL) of up to twenty high schools. At the end of June, students take a standardized admission exam. Exam scores are released at the end of July, and students are matched to high schools based on their exam scores, ROLs, and schools' available seats. The matching process follows the serial dictatorship algorithm. All applicants are ordered by their admission exam scores. Starting from the highest-scoring applicant and moving downwards, each student is assigned to her highest-ranked school with remaining available seats.

The admission exam evaluates students on middle school content knowledge as well as verbal and mathematical reasoning skills. Scores range from 31 to 128, with applicants scoring below 31 excluded from the admissions process. Between 2005 and 2009, the exam's subjects, length, duration, and overall structure remained unchanged. Figure 1 shows the distribution of exam scores for the application cohorts we study. Although the specific exam questions differed across years, the overall score distribution remained stable and comparable across cohorts.

Every high school in Mexico City belongs to one of nine subsystems, which are administrative units that oversee groups of schools. Two of these subsystems are commonly referred to as elite: IPN and UNAM. The IPN subsystem oversees 16 high schools, while the UNAM subsystem oversees 14. As an additional eligibility criterion, admission to any school affiliated with IPN or UNAM requires students to have a middle school GPA above 7/10. In practice, this constraint is not binding, as the GPA requirement is satisfied by more than

Figure 1: Admission exam score by application cohort



NOTE: This figure shows the empirical distribution of exam scores for each application cohort between 2005-2009.

90% of students each year.[2] An important feature of these elite high schools is that they are affiliated with the country's two most prestigious public universities, also named UNAM and IPN. Although only UNAM high schools offer preferential admission to tertiary education at UNAM, IPN high schools arguably provide the best preparation for admission to tertiary education at IPN.

From 2007 to 2014, students in the final year of high school—across both public and private institutions—took a standardized exam assessing their proficiency in mathematics and Spanish. This test was primarily used by the government to evaluate high school performance. Students enrolled in UNAM-affiliated high schools did not participate in this test. For the remainder of the paper, we refer to this assessment as the exit exam. Importantly, the exit exam was designed and administered annually by the same governmental institution, which made a documented effort to ensure comparability across years (CENEVAL, 2012, 2013, 2015).

During the period covered by our study, high school education in Mexico was affected by the *Reforma Integral de la Educación Media Superior* (RIEMS).[3] RIEMS, launched

[2]The minimum GPA required to obtain a middle school certificate is 6/10.

[3]Official legal documentation on the reform can be accessed here: https://www.gob.mx/sep/documentos/acuerdos-secretariales-que-determinan-la-reforma-integral-de-la-educacion-media-superior-riems

in 2007, created a *Sistema Nacional de Bachillerato* (national high school system) and a *Marco Curricular Común* (common core) with the stated goal of articulating the different upper-secondary subsystems and establishing common competency-based curricular standards across them. Implementation of the new curricular framework in schools began in the late 2000s. While we do not observe school-level curricula directly or have access to the exact timing of implementation by the schools in our sample, these stated objectives provide useful context for interpreting the convergence in school quality between elite and non-elite schools documented in later sections.

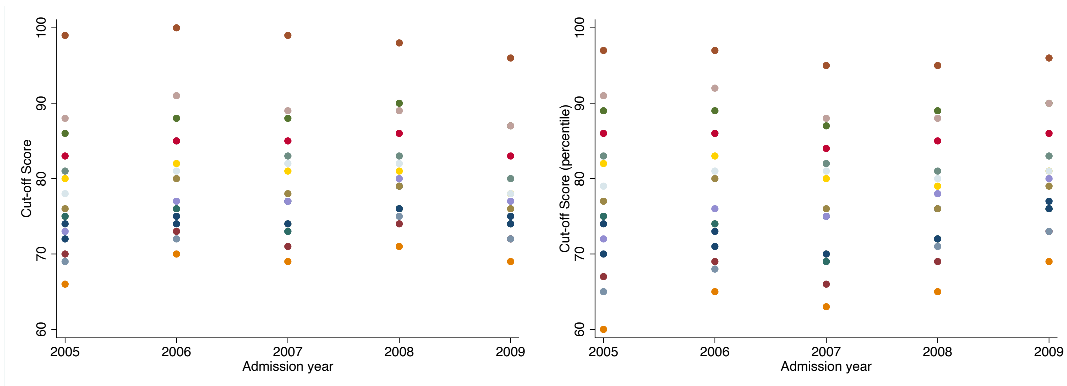
3 Data

We have data on all participants in the centralized education market in Mexico City during 2005-2009. For each student we observe her ROL, admission exam score, GPA, and socio-demographics such as gender and parental education. We combine the admission records with the high school exit exam records during 2008-2014 to measure educational outcomes. For each cohort of applicants we match students across datasets using their national IDs. We only work with students who participate for the first time in the admission process such that we only observe their outcomes at most once. We match students with their exam records between three to five years after application. Expected high school duration is three years.

Students admitted to UNAM-affiliated schools do not participate in the exit exam, so we do not have outcomes for these students. In the analysis we compare students admitted to IPN schools with students admitted to non-IPN and non-UNAM schools. Throughout the analysis period, the IPN subsystem has 16 affiliated schools that provide science oriented education. For the rest of the paper we will refer to these 16 high schools as elite high schools and all other high schools (excluding those affiliated to UNAM) as non-elite high schools.

Figure 2 shows the evolution of the sixteen elite schools' admission cut-offs during our study period. We define an admission cut-off as the lowest admission exam score of a student admitted to an over-subscribed school. All elite schools are over-subscribed each admission year. Panel (a) shows the admission cut-offs in levels. Panel (b) shows the admission cut-offs as percentiles of the test score distribution during a given admission year. We highlight

Figure 2: Elite schools cut-offs



(a) Cut-offs levels (b) Cut-offs percentiles

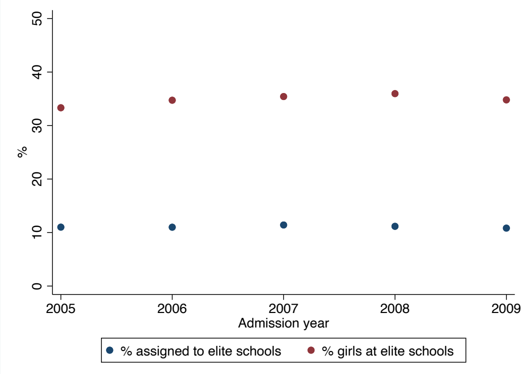
Note: This figure shows the admission cut-offs for each elite school over the period 2005-2009. Panel (a) shows the cut-offs in levels, where the score takes integer values from 31 to 128. Panel (b) shows the same cut-offs as percentiles of the distribution of test scores for each admission year.

two things from these figures. First, elite schools have relatively high admission cut-offs as these schools are heavily over-subscribed. Second, the admission cut-offs in levels and in percentiles are stable over time, implying that the students' ability at the cut-offs has not changed much over our study period.

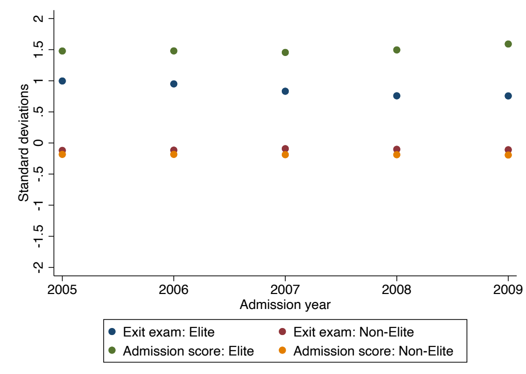
Our outcomes of interest are graduation/dropout and end-of-high school test scores. We consider a dropout a student assigned to a high school in the admission process who does not take the exit exam from three to five years later. For the students who do not drop out, we consider their performance on the mathematics and Spanish tests as a measure of skills at the end of high school.

We show some descriptive statistics in Figure 3. In panel (a), we show that in every year, only around 10% of applicants gain admission to elite schools. Among the admitted students, around 30% are girls. The lower share of girls among students in elite schools is partially related to gendered preferences for non-STEM schools (all of the elite schools in our sample are STEM-focused) and partially to the fact that boys outperform girls on the one-shot admission exam. In panel (b), we show that students admitted to elite schools have higher exit exam scores than students at non-elite schools throughout our study period. They also have higher admission exam scores (approximately 1.5 standard deviations above

Figure 3: Elite schools descriptives



(a) Applicants



(b) Exam scores

NOTE: Panel (a) shows the percentage of students assigned to elite schools each year and the share of girls among them. Panel (b) shows the admission and exit exam scores of students assigned to elite and non-elite schools. We standardize each score within the distribution of scores for each admission cohort.

the mean). The admission system creates stratification across schools by initial ability, and this has not changed much over time. However, this does not mean that elite schools have a causal effect on exit exam scores in any particular year. We use the empirical strategy outlined in the next section to separate the effect of elite schools on academic outcomes from differences that simply reflect the selection of higher-achieving students.

Regarding school-level information, we obtain school characteristics from the annual school census (Formato 911). The school census tracks information on all schools in Mexico at the campus level. The nearly 600 schools that are part of the centralized market are distributed across around 300 campuses, each belonging to one of the nine subsystems. The school census data include information on teachers, students, and classrooms for each admission year.

4 Empirical strategy

4.1 Design

We want to estimate the effect of elite schools on academic outcomes for each admission year. However, since students are not randomly allocated to schools, they may self-select

Table 1: Stylized example of two applicants at the margin (RD Sample)

ROL	Institutions	Cut-off
1st best	Non-Elite	82
2nd best	Elite	78
3rd best	Non-Elite	76
4th best	Non-Elite	53
Application score=79		
Local Institution Ranking		
Preferred	Elite	Yes
Next-best	Non-Elite	No
Application score=77		
Local Institution Ranking		
Preferred	Elite	No
Next-best	Non-Elite	Yes

NOTE: This table provides an example where two applicants are on the margin of receiving an offer for an elite school and a non-elite school.

into elite schools for observable or unobservable reasons. To deal with the selection problem, we compare the outcomes of students who prefer elite schools to non-elite schools and are marginally admitted or rejected from elite schools (i.e., same ability). Under this design, we look to obtain internally valid estimates of the effect of admission to elite schools for students at the elite school's admission cut-offs. We define a school admission cut-off as the score of the last admitted student to an oversubscribed school. All elite schools are oversubscribed.

We follow Kirkeboen et al. (2016) strategy to estimate the effects of university majors and institutions in a centralized education market. We consider the case where there are only two institutions/subsystems, elite and non-elite. We focus on students whose local institution ranking places an elite school just above a non-elite school (i.e., elite $\succ$ non-elite). Among these students, some are marginally admitted to the elite school, while others with very similar scores are instead assigned to their next-best non-elite alternative. Table 1 shows the types of applicants we include in our sample. This example considers two applicants with the same local institution ranking. The applicant with a score of 79 gains admission to her preferred institution, while the applicant with a score of 77 gains admission to her next-best institution.

Notice that students rank schools in cut-off descending order in our stylized example. The serial dictatorship algorithm guarantees that a student will never be assigned to a school that is not ranked in cut-off descending order. Therefore, we modify the observed ROLs to exclude all schools not ranked in cut-off descending order. ROLs subject to this modification would lead to the same equilibrium allocation as the unmodified ones.

We work with a sample of students with ROLs that list an elite school as their first best and a non-elite school as their next-best in their local institution ranking. We use a Regression Discontinuity Design (RDD) and focus on students close to the elite school admission cut-offs. Since there are 16 elite schools, we have 16 admission cut-offs. The intuition behind the identification strategy is that marginally admitted and rejected students from elite high schools have similar observable and unobservable characteristics.

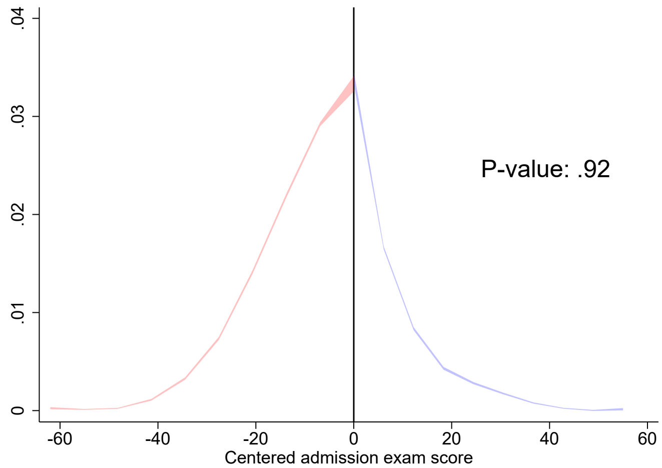
Our empirical model is:

$$Y_{ik} = \mu_k + \alpha_1 admit_i + \alpha_2(S_i - \underline{\text{s}}_k) + \alpha_3(S_i - \underline{\text{s}}_k) \times admit_i + \epsilon_{ik}, \tag{1}$$

where index i is for individual, index k is for elite school admission cut-off. We have $K = 16$ elite schools and stack individuals with different elite cut-offs in the estimation sample. Our specification includes cut-off fixed effects μ_k as recommended by Fort et al. (2022) for the case of multi cut-off RDDs. We include a dummy variable for elite school admission $admit_i$. We denote the admission exam score S_i and $\underline{\text{s}}_k$ is the elite school admission cut-off relevant to student i. In this specification $admit_i = 1$ when $S_i - \underline{\text{s}}_k \geq 0$.

The coefficient of interest is α_1. It measures the intent-to-treat (ITT) effect of gaining marginal admission to an elite high school instead of a non-elite one. For the estimation, we use the optimal bandwidth obtained by following Calonico et al. (2014), which minimizes the mean square error. Within the optimal bandwidth, we estimate the parameters in Equation 1 using a local linear regression with a triangular kernel and cluster the standard errors at the admitted high school level. Since the admission process remains the same from 2005-2009,

Figure 4: Density



NOTE: This figure shows the density of the centered running variable. The lines and shaded areas indicate the 95% confidence intervals for robust inference. The vertical line indicates the admission threshold.

we follow the same design and empirical model for each admission cohort.

4.2 Validity

Before proceeding with the results, we provide evidence of the validity of the design. Following Imbens and Lemieux (2008), certain conditions need to be met to support the validity of an RDD. Figure 4 shows the density of the centered admission score for the pooled sample for 2005-2009. There is no evidence that the density is discontinuous around the centered cut-offs, which indicates that manipulation of the running variable is unlikely. A formal statistical test does not reject the continuity of the density (p-value=0.92). As we estimate RDDs for each admission cohort, we also show evidence of a lack of manipulation for each cohort. We include these results in Appendix B.

As further support for the validity of the design, Table 2 presents results from estimating Equation 1 over predetermined covariates expected to be continuous around the centered admission cut-offs. For the pooled sample of years 2005-2019, being marginally admitted to an elite high school has no statistically significant effect on gender, students' GPA in middle school, fathers' education, or number of siblings. We report the estimates on predetermined covariates for each admission cohort in Appendix C.

Table 2: Covariates

	Girl	GPA	Father	Siblings
RD_Estimate	0.012	0.002	0.011	0.009
	(0.011)	(0.012)	(0.008)	(0.025)
Optimal BW	8.996	14.542	13.378	7.696
Mean	0.434	8.247	0.317	1.935
N	40,254	58,423	51,427	36,187

NOTE: This table shows RD estimates following the methods in Calonico et al. (2014) and the associated software package *rdrobust*. Each column indicates a different outcome variable. For each outcome, we also report the associated optimal bandwidth, the average outcome for the marginally rejected students, and the number of observations. Robust standard errors in parentheses.

Overall, we find no evidence of manipulation in the pooled sample or for each separate admission cohort, which we take as evidence in support of the validity of our design.

5 Results

We first show the results for the pooled sample of cohorts 2005-2009. Table 3 shows that marginal admission to an elite school has a positive and statistically significant effect on mathematics test scores, a non-statistically significant effect on Spanish test scores, and a negative and statistically significant effect on the probability of taking the exit exam (i.e., graduating). The point estimate for math, 0.053, is measured in standard deviations, so marginal admission increases end-of-high-school math scores by about 0.05 standard deviations when pooling all cohorts. The effect on Spanish is close to zero and not significant. The estimate of -0.072 for exit-exam take-up implies that marginal admission reduces the probability of taking the exit exam by about 7 percentage points on average over 2005-2009.

While pooling multiple years for estimation is a common approach in the literature to increase precision, it is important to consider that this may obscure heterogeneity in the effects across years. An important advantage of our setting is that our sample sizes are large

Table 3: Pooled sample 2005-2009

	Math	Spanish	Test Taker
RD_Estimate	0.053***	-0.027	-0.072***
	(0.019)	(0.021)	(0.010)
Optimal BW	12.611	10.355	12.659
Mean	0.143	0.137	0.601
N	31,070	27,532	53,088

NOTE: This table shows RD estimates following the methods in Calonico et al. (2014) and the associated software package *rdrobust*. Each column indicates a different outcome variable. For each outcome, we also report the associated optimal bandwidth, the average outcome for the marginally rejected students, and the number of observations. Robust standard errors in parentheses.

enough to allow us to implement the same empirical design on a yearly basis.

To study whether the effect of being marginally admitted to an elite high school on test scores changes over time, we estimate Equation 1 separately for each admission cohort between 2005 and 2009. Table 4 presents estimated parameters for mathematics performance as the outcome variable. A clear pattern emerges: the effect is positive and significant in 2005 but steadily decreases and becomes not significant for later cohorts. We reject the equality of coefficients at conventional statistical significance levels and show that the coefficients have a statistically significant negative linear time trend. Quantitavely, the estimated effect in 2005 is 0.199 standard deviations, and in 2006 it is 0.178 standard deviations, whereas the estimates for 2007-2009 are small in magnitude and not statistically different from zero. Consistent with our findings for the initial years of our sample, Dustan et al. (2017) also find positive and significant effects on mathematics performance for the 2005 and 2006 admission cohorts.[4] Our extended analysis over five years reveals that this effect is not time invariant, as it diminishes and loses statistical significance in later cohorts.

Table 5 shows that marginal admission to an elite science school does not have a statis-

[4]Their sample selection criteria are different from ours. Appendix I shows a replication of their results using their sample selection criteria.

Table 4: Math

	2005	2006	2007	2008	2009
RD_Estimate	0.199***	0.178***	0.033	-0.023	-0.041
	(0.042)	(0.038)	(0.040)	(0.038)	(0.040)
Optimal BW	11.576	13.308	13.006	11.565	12.615
Mean	0.111	0.102	0.144	0.179	0.161
N	4,766	5,663	6,482	6,800	7,258

H0: 2005=2006=2007=2008=2009, p-value: 0.000

Linear trend: coef -0.070, p-value 0.000

NOTE: This table shows RD estimates following the methods in Calonico et al. (2014) and the associated software package *rdrobust*. Each column indicates a different admission cohort. For each cohort, we also report the associated optimal bandwidth, the average outcome for the marginally rejected students, and the number of observations. Robust standard errors in parentheses.

tically significant effect on Spanish scores for any admission cohort during the study period 2005-2009. We believe this is because the elite schools we study here are highly focused on providing science education. Dustan et al. (2017) also find no effects on Spanish scores for admission cohorts 2005 and 2006. We do not reject the equality of coefficients across time and find no statistically significant linear time trend in them. Overall, the results show the time specificity of the estimated effects on mathematics performance.

We next show how marginal admission to an elite school affects exit-exam participation (our measure of dropout) for each cohort. Table 6 reports year-specific effects on the probability of taking the exit exam. For all cohorts between 2005 and 2009, the estimates are negative and statistically significant, ranging from -0.058 to -0.101. This means that marginal admission increases the probability of not taking the exit exam by roughly 6 to 10 percentage points in every cohort. We cannot reject the equality of these coefficients over time, and the estimated linear time trend is close to zero and not statistically significant. Thus, while the effect on math scores declines and becomes indistinguishable from zero by 2009, the increase in dropout risk for marginally admitted students remains sizeable and stable across cohorts.

Table 5: Effects on Spanish by year

	2005	2006	2007	2008	2009
RD_Estimate	-0.023	0.038	-0.026	-0.075	-0.038
	(0.050)	(0.043)	(0.034)	(0.052)	(0.039)
Optimal BW	12.628	12.134	11.188	11.858	11.623
Mean	0.101	0.080	0.145	0.165	0.111
N	5,064	5,417	5,827	6,800	6,813

H0: 2005=2006=2007=2008=2009, p-value: 0.769

Linear trend: coef -0.016, p-value 0.255

NOTE: This table shows RD estimates following the methods in Calonico et al. (2014) and the associated software package *rdrobust*. Each column indicates a different admission cohort. For each cohort, we also report the associated optimal bandwidth, the average outcome for the marginally rejected students, and the number of observations. Robust standard errors in parentheses.

Table 6: Effects on graduation by year

	2005	2006	2007	2008	2009
RD_Estimate	-0.082***	-0.058**	-0.064**	-0.062***	-0.101***
	(0.025)	(0.025)	(0.027)	(0.020)	(0.021)
Optimal BW	10.441	9.431	11.570	8.351	10.539
Mean	0.624	0.615	0.585	0.585	0.614
N	7,429	7,573	10,159	9,446	10,830

H0: 2005=2006=2007=2008=2009, p-value: 0.656

Linear trend: coef -0.004, p-value 0.562

NOTE: This table shows RD estimates following the methods in Calonico et al. (2014) and the associated software package *rdrobust*. Each column indicates a different admission cohort. For each cohort, we also report the associated optimal bandwidth, the average outcome for the marginally rejected students, and the number of observations. Robust standard errors in parentheses.

Table 7: Math elite below median selectivity

	2005	2006	2007	2008	2009
RD_Estimate	0.201***	0.208***	-0.060	0.029	-0.014
	(0.054)	(0.073)	(0.062)	(0.084)	(0.042)
Optimal BW	10.319	6.535	6.963	9.062	10.476
Mean	0.064	0.042	0.117	0.049	0.099
N	2,878	1,728	2,213	1,337	3,663

H0: 2005=2006=2007=2008=2009, p-value: 0.002

Linear trend: coef -0.058, p-value 0.000

NOTE: This table shows RD estimates following the methods in Calonico et al. (2014) and the associated software package *rdrobust*. Each column indicates a different admission cohort. For each cohort, we also report the associated optimal bandwidth, the average outcome for the marginally rejected students, and the number of observations. Robust standard errors in parentheses.

5.1 Heterogeneity by Selectivity

As a robustness check, we examine whether the time pattern of effects differs across elite schools that vary in selectivity. To do so, we split the sixteen elite schools into two groups based on their selectivity and re-estimate the year-specific effects for each group. We classify as "more selective" those elite schools whose average admission cut-off over 2005–2009 is above the median across elite schools, and as "less selective" those below the median.

Table 7 reports the RD estimates for math scores by cohort for the less selective elite schools. The estimated effects are 0.201 and 0.208 standard deviations in 2005 and 2006, respectively, and are statistically significant at conventional levels. For later cohorts the estimates are smaller in magnitude and not statistically different from zero. We reject the hypothesis that the coefficients are equal across years (p-value 0.002), and the estimated linear time trend is negative and statistically significant (–0.058, p-value 0.000).

Table 8 reports the corresponding estimates for the more selective elite schools. Here, too, the effects are large and positive in the early cohorts, 0.242 and 0.207 standard deviations in 2005 and 2006, and then decline towards zero or negative values in later cohorts. The

Table 8: Math elite above median selectivity

	2005	2006	2007	2008	2009
RD_Estimate	0.242***	0.207***	0.042	-0.040	-0.078
	(0.065)	(0.055)	(0.064)	(0.043)	(0.078)
Optimal BW	10.749	13.123	12.345	10.940	12.582
Mean	0.227	0.212	0.265	0.229	0.279
N	1,607	2,879	2,681	4,933	3,115

H0: 2005=2006=2007=2008=2009, p-value: 0.001

Linear trend: coef -0.088, p-value 0.000

NOTE: This table shows RD estimates following the methods in Calonico et al. (2014) and the associated software package *rdrobust*. Each column indicates a different admission cohort. For each cohort, we also report the associated optimal bandwidth, the average outcome for the marginally rejected students, and the number of observations. Robust standard errors in parentheses.

joint test again rejects equality of coefficients across years (p-value 0.001), and the estimated linear time trend is negative and significant (–0.088, p-value 0.000).

These results show that the decline in the effect of marginal admission to elite schools on math scores is not driven by a small subset of institutions. Both less selective and more selective elite schools exhibit strong positive impacts in the first two cohorts and much smaller, statistically insignificant impacts by the end of the period. At the same time, the estimates remain too imprecise to support a fully disaggregated analysis by individual school or by detailed next-best alternative in the spirit of Kirkeboen et al. (2016).

6 Why did the effect change over time?

The math results in Section 5 are based on students who take the exit exam. Since we do not observe test scores for students who do not take the exam, changes in exam participation over time could, in principle, generate changes in the estimated effects on test scores. However, the results in Table 6 show that there is no clear trend in the effects on exit exam take-up, indicating that marginal admission to an elite school raises the dropout risk for marginal

Table 9: Peers admission exam

	2005	2006	2007	2008	2009
RD_Estimate	17.424***	17.303***	18.432***	18.341***	18.625***
	(0.923)	(0.999)	(0.966)	(0.917)	(1.002)
Optimal BW	10.334	10.614	10.206	11.509	14.063
Mean	63.906	66.282	65.576	66.457	62.585
N	7,429	8,155	9,534	11,871	13,753

H0: 2005=2006=2007=2008=2009, p-value: 0.853

Linear trend: coef 0.280, p-value 0.323

NOTE: This table shows RD estimates following the methods in Calonico et al. (2014) and the associated software package *rdrobust*. Each column indicates a different admission cohort. For each cohort, we also report the associated optimal bandwidth, the average outcome for the marginally rejected students, and the number of observations. Robust standard errors in parentheses.

students by a fairly stable amount over time, even as the math effects decline. We interpret this as evidence against differential selection over time as an explanation for the declining test-score effects. We highlight that under non-differential selection over time, correcting for selection into the exit exam would shift the estimated test-score effects by a similar amount in each cohort, without changing the time trend in the effect on mathematics test scores.

Another possibility is that the discontinuous jump in peer quality induced by marginal elite school admission has changed over time, generating time-varying effects on test scores. In fact, a common interpretation in the selective school literature is that the RDD-estimated parameter measures peer effects. The logic behind this interpretation is that marginally admitted students to selective schools experience higher peer quality relative to their counterfactual alternatives. For example, this occurs when selective schools admit students based on skill measures, as in Mexico City.

In Table 9, we show the jump in peer quality for students marginally admitted to an elite school relative to their next-best alternative. We measure peer quality using the average admission exam score of all the admitted students to a school in a given year. Our results show that admission to an elite school implies an increase in peer quality for each year in our sample. Nevertheless, the estimated effect on peer quality is roughly constant over time.

We do not reject the equality of the yearly parameters or find evidence of a linear trend in them. Therefore, changes in the average peer quality experienced by those admitted to elite schools are unlikely to be the driver of the decreasing effect on test scores.

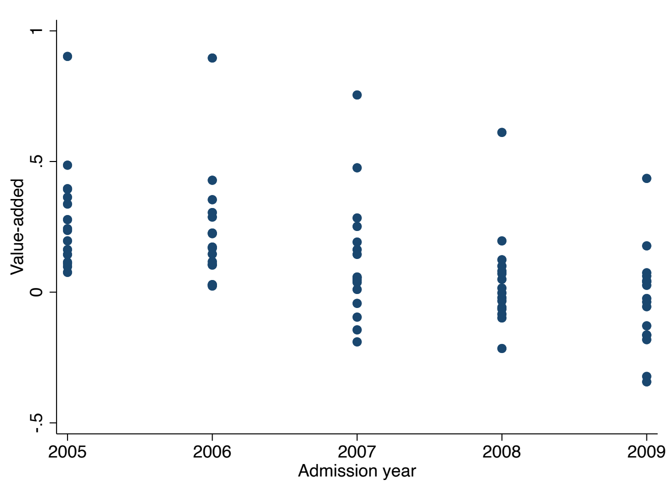
Since we have sixteen elite schools as preferred schools and multiple non-elite schools as next-best alternatives, another possible explanation for the time-varying effects could be that the weights of the different schools have been changing over time. This could be the case for the RDD estimates, as our parameter of interest compares a weighted average of elite schools against a weighted average of non-elite schools. Appendix D shows that for students marginally admitted to elite schools, the weights for different elite schools have been mostly constant over time. As there are many next-best alternatives for students marginally rejected, we summarize the information by grouping schools into subsystems. Appendix D also shows that most of the next-best subsystems have a constant weight over time, except the two most popular ones, which experience small changes. However, when we repeat the RDD analysis excluding students admitted to the two most popular next-best subsystems, we still find decreasing treatment effects on mathematics test scores with a similar negative time trend as in our main specification. We take this as evidence that the small changes in school/subsystem weights over time are unlikely to explain the time-varying effects.

We next consider changes in school quality as an explanation for our time-varying effects. Access to elite schools could have a positive effect on test scores if they provide students with access to higher-quality schools relative to their next-best alternatives. Therefore, if the relative school quality between the first and next-best alternatives has been changing over time, the RDD estimated parameter would also change over time. To measure individual school quality, we estimate the value-added model shown in Equation 2

$$Y_{it} = \sum_{j=0}^{J} \alpha_{jt} D_{ijt} + X'_{it}\Gamma_t + \nu_{it}, \tag{2}$$

where X_i is a vector that includes the standardized admission exam score and standardized middle school GPA. The outcome variable Y_{it} measures mathematics exam performance at the end of high school. Our parameters of interest are α_{jt}, which measure school quality for each school j and for each year t. For the value-added estimation, we do not restrict the sample to the RDD sample; instead, we use all available data.

Figure 5: Elite schools value-added ($\hat{\alpha}_{jt}$)



NOTE: This figure shows the evolution over time of elite-schools value-added estimated parameters. Each dot represents a school quality estimate $\hat{\alpha}_{jt}$.

In Figure 5, we show the value-added estimated parameters for the sixteen elite schools for each admission year. The value-added of elite schools consistently decreases every year. Note that the value-added parameter is estimated using the full population of students, not just those on the margin of admission. We take this as evidence of the average quality of elite schools decreasing over time.

We then use all the value-added estimated parameters as the outcome variable in our RDD estimations. We aim to capture whether there is a discontinuous jump in school value-added between students marginally admitted to elite schools and those marginally rejected. Additionally, we aim to determine if the effect on school value-added has changed over time. In Table 10, we show that marginal admission to elite schools results in gaining access to schools with higher value-added. In addition, the effect on school quality is decreasing over time. The RD estimates for value-added are 0.270 and 0.258 in 2005 and 2006, respectively, and remain positive but smaller in 2007 (0.133) and 2008 (0.055), before falling to a small and statistically insignificant value in 2009 (0.013). As we are using a two-step estimator, we include bootstrapped 95% confidence intervals for our estimates in addition to robust standard errors. We reject the equality of coefficients across years (p-value 0.000) and find a negative and statistically significant linear trend (-0.071, p-value 0.000).

Table 10: Value-added ($\hat{\alpha}_{jt}$)

	2005	2006	2007	2008	2009
RD_Estimate	0.270***	0.258***	0.133***	0.055***	0.013
	(0.025)	(0.021)	(0.026)	(0.015)	(0.024)
CI Bootstrap	[0.245, 0.296]	[0.233, 0.283]	[0.107, 0.158]	[0.030, 0.081]	[-0.014, 0.041]
Optimal BW	14.580	15.973	13.647	14.039	16.147
Mean	-0.045	-0.063	-0.057	-0.062	-0.085
N	9,151	10,333	11,274	13,737	14,937

H0: 2005=2006=2007=2008=2009, p-value: 0.000

Linear trend: coef -0.071, p-value 0.000

Note: This table shows RD estimates following the methods in Calonico et al. (2014) and the associated software package *rdrobust*. Each column indicates a different admission cohort. For each cohort, we also report bootstrapped 95% confidence intervals, the associated optimal bandwidth, the average outcome for the marginally rejected students, and the number of observations. Robust standard errors in parentheses.

6.1 Robustness

As a robustness check, we estimate value-added using a cross-fitting procedure described in Appendix E. To reduce the risk of overfitting bias, for each student, the value-added assigned to their school is always computed from a model estimated on other students, rather than on the same data used to evaluate the RD effect. We then re-estimate the RDD with these cross-fitted value-added measures. The estimates are presented in Table 11. The results show that the estimated RD effects on value-added are 0.268 and 0.251 in 2005 and 2006, and then decline to 0.132 in 2007, 0.062 in 2008, and 0.021 in 2009. As before, we reject the hypothesis that the coefficients are equal across years (p-value 0.000), and we find a negative and statistically significant linear trend (–0.068, p-value 0.000).

To further confirm the robustness of our results, in Appendix F, we also estimate school value-added using an empirical Bayes (shrinkage) procedure and repeat the RDD analysis with the shrunken value-added estimates as the outcome. Table F.1 reports the corresponding RD estimates. The pattern is very similar to Table 10: the effects on value-added are large and positive in 2005 and 2006 (0.261 and 0.252), decline in magnitude in 2007 and 2008 (0.129 and 0.052), and are close to zero in 2009 (0.011). We again reject the equality

Table 11: Value-added (Cross-fitting, 20 groups)

	2005	2006	2007	2008	2009
RD_Estimate	0.268***	0.251***	0.132***	0.062***	0.021
	(0.023)	(0.021)	(0.026)	(0.015)	(0.023)
Optimal BW	14.177	15.638	12.995	13.764	15.370
Mean	-0.028	-0.049	-0.046	-0.048	-0.068
N	9,151	10,333	10,740	13,174	14,388

H0: 2005=2006=2007=2008=2009, p-value: 0.000

Linear trend: coef -0.068, p-value 0.000

NOTE: This table shows RD estimates following the methods in Calonico et al. (2014) and the associated software package *rdrobust*. Each column indicates a different admission cohort. For each cohort, we also report the associated optimal bandwidth, the average outcome for the marginally rejected students, and the number of observations. Robust standard errors in parentheses.

of coefficients across years (p-value 0.000), and the estimated linear time trend is negative and statistically significant (–0.069, p-value 0.000).

Taken together, the results in Tables 10 and 11, and in Appendix F show that the school value-added advantage of elite schools over their next-best alternatives declines sharply over time and that this pattern is robust to alternative ways of estimating value-added. With this evidence, we conclude that the decreasing effects of marginal admission to elite schools on test scores are not due to changes in the composition of test takers over time or changes in peer quality. Instead, our results suggest that the difference in the quality of schools that treated and untreated students are exposed to is behind the time-varying effects.

7 Why did school quality change?

When estimating school value-added parameters, we estimate fixed effects that reflect school characteristics and their productivities. Changes in these characteristics or their productivity

Table 12: $Z'_{jt}\hat{\theta}_t$

	2005	2006	2007	2008	2009
RD_Estimate	0.018***	-0.101***	-0.082***	-0.089***	-0.069***
	(0.003)	(0.010)	(0.011)	(0.012)	(0.015)
CI Bootstrap	[-0.002, 0.038]	[-0.123, -0.080]	[-0.112, -0.052]	[-0.114, -0.065]	[-0.092, -0.046]
Optimal BW	11.075	12.909	12.133	11.754	11.371
Mean	-0.000	-0.021	0.008	-0.021	-0.007
N	7,812	8,973	10,621	11,676	11,538

H0: 2005=2006=2007=2008=2009, p-value: 0.000

Linear trend: coef -0.014, p-value 0.001

NOTE: This table shows RD estimates following the methods in Calonico et al. (2014) and the associated software package *rdrobust*. Each column indicates a different admission cohort. For each cohort, we also report bootstrapped 95% confidence intervals, the associated optimal bandwidth, the average outcome for the marginally rejected students, and the number of observations. Robust standard errors in parentheses.

over time will thus affect the value-added estimates. We model this relationship as:

$$\hat{\alpha}_{jt} = Z'_{jt}\theta_t + \eta_{jt}, \tag{3}$$

where Z_{jt} is a vector of school characteristics that includes average peers' admission score, average peers' GPA, teachers per pupil, female teachers per pupil, full-time teachers per pupil, highly qualified teachers per pupil, and classrooms per pupil. We estimate Equation 3 to separate the part of value-added explained by observable school characteristics and their productivities ($Z'_{jt}\hat{\theta}_t$) from unobservable school characteristics ($\hat{\eta}_{jt}$). We then use these estimates as outcome variables in our RDD framework.

In Table 12 we show that marginal admission to elite schools is associated with an increase in the fitted value $Z'_{jt}\hat{\theta}_t$ for the 2005 admission cohort. However, for later cohorts the effect becomes negative and decreasing. We reject equality of coefficients over time, and we find they have a statistically significant negative linear trend. The RD effect on $Z'_{jt}\hat{\theta}_t$ is 0.018 in 2005 and remains negative thereafter (between –0.101 and –0.069), with a negative and statistically significant linear time trend (–0.014, p-value 0.001).

The change in the effect on the fitted value can be due to time changes in school char-

Table 13: $\hat{\eta}_{jt}$

	2005	2006	2007	2008	2009
RD_Estimate	0.252***	0.357***	0.209***	0.142***	0.085***
	(0.024)	(0.025)	(0.024)	(0.016)	(0.023)
CI Bootstrap	[0.224, 0.280]	[0.328, 0.386]	[0.174, 0.244]	[0.108, 0.177]	[0.055, 0.114]
Optimal BW	16.921	16.149	14.300	11.189	14.465
Mean	-0.044	-0.044	-0.061	-0.038	-0.081
N	9,747	10,478	11,671	11,676	13,697

H0: 2005=2006=2007=2008=2009, p-value: 0.000

Linear trend: coef -0.057, p-value 0.000

NOTE: This table shows RD estimates following the methods in Calonico et al. (2014) and the associated software package *rdrobust*. Each column indicates a different admission cohort. For each cohort, we also report bootstrapped 95% confidence intervals, the associated optimal bandwidth, the average outcome for the marginally rejected students, and the number of observations. Robust standard errors in parentheses.

acteristics (Z_{jt}), time changes in productivities (θ_t), or both. In Appendix G, we use each school characteristic as the outcome variable in the RDD specification and find no effect pattern or time trend. In particular, we do not find systematic changes over time in the RD jumps for peer composition or for the main observable inputs such as teachers per pupil, highly qualified teachers per pupil, or classrooms per pupil. This implies that the same inputs (e.g., teacher ratios) became less productive in elite schools relative to non-elite schools during our study period.

We next consider the estimated residuals $\hat{\eta}_{jt}$, which we interpret as a measure of unobservable school characteristics that affect school quality. In Table 13, we show the results of our RDD estimations using $\hat{\eta}_{jt}$ as the outcome variable. We find that marginal admission to elite schools implies gains in unobservable school characteristics for all years between 2005-2009. However, this effect decreases over time. We reject equality of effects over time and find a statistically significant negative linear trend in coefficients. The RD effect on $\hat{\eta}_{jt}$ declines from 0.252 in 2005 to 0.085 in 2009, with a negative and statistically significant linear trend (–0.057, p-value 0.000). This suggests that both observed and unobserved aspects of school quality contributed to the declining time-varying effects on mathematics test scores.

To assess robustness of this decomposition to functional-form choices, we also estimate log-level specifications, using the logarithm of test score before estimating Equation 2. Tables H.1 and H.2 in Appendix H show that the RD effect on the fitted component goes from positive to negative over time, while the RD effect on the residual component remains positive but decreasing in this specification as well. The joint tests reject equality of coefficients across years, and the estimated linear trends are negative and statistically significant in both cases.

A possible explanation behind the change in the productivity of school characteristics for elite and non-elite schools is the *Reforma Integral de la Educación Media Superior* (RIEMS). RIEMS, launched in 2007, created a Sistema Nacional de Bachillerato and a Marco Curricular Común with the stated goal of articulating the different upper-secondary subsystems and defining common competency-based curricular standards across them. Such a reform would not necessarily change observable inputs such as teacher–student ratios or the number of classrooms, but it is designed to change how similar inputs are used and which competencies they are meant to produce in the classroom. In this sense, RIEMS provides a plausible mechanism for the decline in the relative contribution of observed inputs to value-added that we document for elite schools: as curricula and competency requirements become more similar across subsystems, differences in the productivity of given inputs may narrow. We cannot observe school-level curricula directly, however, so we interpret this link between RIEMS and our value-added results as suggestive rather than as a separately identified causal effect.

The results in this section highlight that school value-added captures differences in school characteristics and their associated productivities. Since school characteristics and productivities can change over time, value-added can also change. Thus, if the effect of elite schools on academic outcomes is due to students gaining access to higher value-added schools, then this effect does not need to be constant over time. The same applies when comparing effects across different contexts, as the estimated parameters may not measure the same treatments.

8 Conclusions

The results presented in this paper indicate that the effect of being marginally admitted to an elite high school is not constant over time and relates to time changes in relative school quality between elite schools and their next-best alternatives. In the case of Mexico City, we find that over five years (2005-2009), the effect of marginal admission to elite science schools on mathematics test scores monotonically decreased and went from positive and statistically significant to not significant. In contrast, the effect on the probability of taking the exit exam is negative, sizable, and relatively stable across cohorts, implying that marginal admission to elite schools increases dropout risk by a similar amount in each year of our sample.

We explain the time-varying effect by showing that the gains in school quality from marginal admission to elite schools monotonically decreased during our study period. The gains in peer quality due to marginal admission did not change over time. Also, there were no time changes in the effect of marginal admission on exit-exam test taking. In addition, we find that the decline in effects on mathematics test scores and on school value-added is present both for more selective and less selective elite schools, which shows that the pattern is not driven by a small subset of institutions.

Our decomposition of value-added indicates that both the part explained by observable school characteristics and the residual component decreased over time for students marginally admitted to elite schools relative to their next-best alternatives. At the same time, we do not find systematic time trends in the RD jumps of the main observable inputs themselves, such as peer composition, teacher–student ratios, or classrooms per pupil. Taken together, these findings suggest that changes in the productivity of school characteristics, rather than changes in their levels, played an important role in the decline of elite schools' relative value-added. The timing and stated objectives of the Reforma Integral de la Educación Media Superior (RIEMS), which sought to align curricular standards across upper-secondary subsystems, are consistent with a mechanism in which curricular alignment reduces differences in how similar inputs translate into achievement, although we cannot directly observe curricula or separately identify the causal effect of this reform.

Our results contribute to a growing understanding that the effects of elite schools are highly context-dependent. They suggest that the caution researchers apply when generalizing results across countries should also be extended to generalizations across time, even within the same institutional setting. More broadly, the findings underscore that estimates from internally valid designs need not be stable when the underlying school environment, value-added, or policy context changes.

References

Abdulkadiroğlu, Atila, Joshua Angrist, and Parag Pathak, "The elite illusion: Achievement effects at Boston and New York exam schools," *Econometrica*, 2014, *82* (1), 137–196.

_ , _ , Yusuke Narita, Parag Pathak, and Roman Zarate, "Regression discontinuity in serial dictatorship: Achievement effects at Chicago's exam schools," *American Economic Review, Papers and Proceedings*, 2017, *107* (5), 240–245.

Altonji, Joseph G and Richard K Mansfield, "Estimating group effects using averages of observables to control for sorting on unobservables: School and neighborhood effects," *American Economic Review*, 2018, *108* (10), 2902–2946.

Angrist, Joshua D, Parag A Pathak, and Roman A Zarate, "Choice and consequence: Assessing mismatch at Chicago exam schools," *Journal of Public Economics*, 2023, *223*, 104892.

Beuermann, Diether W and C Kirabo Jackson, "The short-and long-run effects of attending the schools that parents prefer," *Journal of Human Resources*, 2022, *57* (3), 725–746.

Calonico, Sebastian, Matias D Cattaneo, and Rocio Titiunik, "Robust nonparametric confidence intervals for regression-discontinuity designs," *Econometrica*, 2014, *82* (6), 2295–2326.

CENEVAL, "Manual Técnico ENLACE Media Superior 2008–2010," 2012. Centro Nacional de Evaluación para la Educación Superior.

— , "Manual Técnico ENLACE Media Superior 2011–2012," 2013. Centro Nacional de Evaluación para la Educación Superior.

— , "Manual Técnico ENLACE Media Superior 2013–2014," 2015. Centro Nacional de Evaluación para la Educación Superior.

Clark, Damon, "Selective schools and academic achievement," *The BE Journal of Economic Analysis & Policy*, 2010, *10* (1).

Dobbie, Will and Roland G Fryer Jr, "The impact of attending a school with high-achieving peers: Evidence from the New York City exam schools," *American Economic Journal: Applied Economics*, 2014, *6* (3), 58–75.

Dustan, Andrew, Alain De Janvry, and Elisabeth Sadoulet, "Flourish or fail? The risky reward of elite high school admission in Mexico City," *Journal of Human Resources*, 2017, *52* (3), 756–799.

Fort, Margherita, Andrea Ichino, Enrico Rettore, and Giulio Zanella, "Multi-cutoff rd designs with observations located at each cutoff: problems and solutions," 2022. CEPR Discussion Paper No. DP16974.

Groote, Olivier De and Koen Declercq, "Tracking and specialization of high schools: Heterogeneous effects of school choice," *Journal of Applied Econometrics*, 2021, *36* (7), 898–916.

Imbens, Guido W and Thomas Lemieux, "Regression discontinuity designs: A guide to practice," *Journal of econometrics*, 2008, *142* (2), 615–635.

Jackson, C Kirabo, "Do students benefit from attending better schools? Evidence from rule-based student assignments in Trinidad and Tobago," *The Economic Journal*, 2010, *120* (549), 1399–1429.

Kirkeboen, Lars J, Edwin Leuven, and Magne Mogstad, "Field of study, earnings, and self-selection," *The Quarterly Journal of Economics*, 2016, *131* (3), 1057–1111.

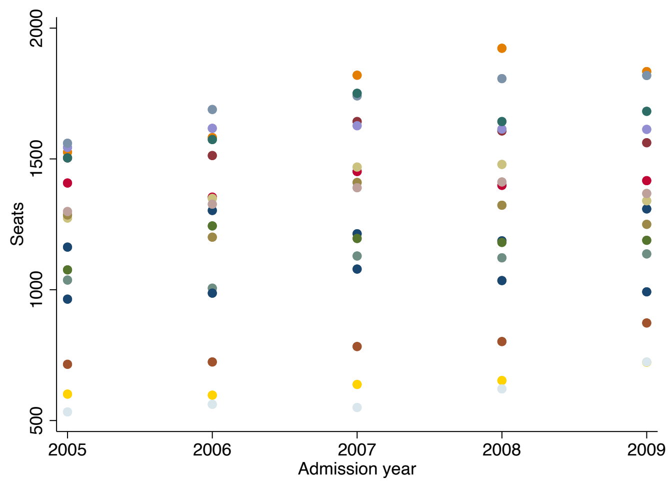
Lucas, Adrienne M and Isaac M Mbiti, "Effects of school quality on student achievement: Discontinuity evidence from kenya," *American Economic Journal: Applied Economics*, 2014, *6* (3), 234–263.

Pop-Eleches, Cristian and Miguel Urquiola, "Going to a better school: Effects and behavioral responses," *American Economic Review*, 2013, *103* (4), 1289–1324.

Todd, Petra E and Kenneth I Wolpin, "On the specification and estimation of the production function for cognitive achievement," *The Economic Journal*, 2003, *113* (485), F3–F33.

A Capacities, applicants, and schools

Figure A.1: Elite schools seats



Note: This figure shows the evolution over time of the number of seats in each elite high school between 2005 and 2009. Each dot represents the seat capacity of one elite school in a given admission year.

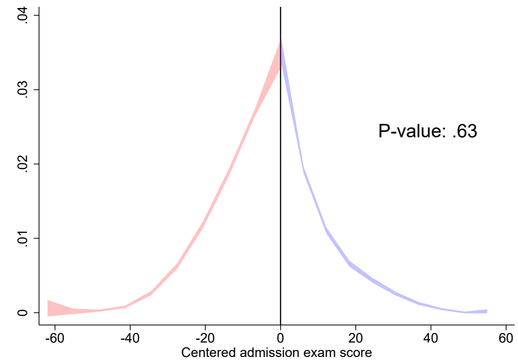
Table A.1: Number of applicants and schools, 2005–2009

	2005	2006	2007	2008	2009
Applicants	249,499	255,851	256,335	257,199	270,097
Schools	634	650	658	663	676

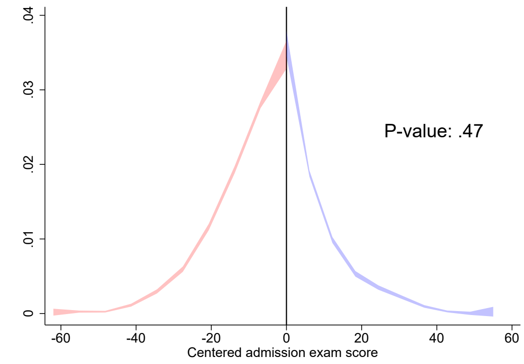
Note: This table shows the evolution over time of the number of applicants (who reach the matching-algorithm stage) and the number of available public high schools in the centralized admission market for each application cohort.

B Density by cohort

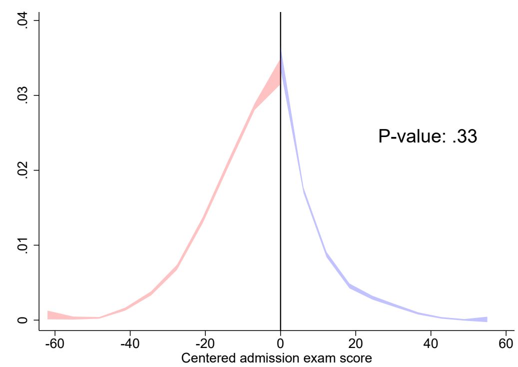
Figure B.1: Density test



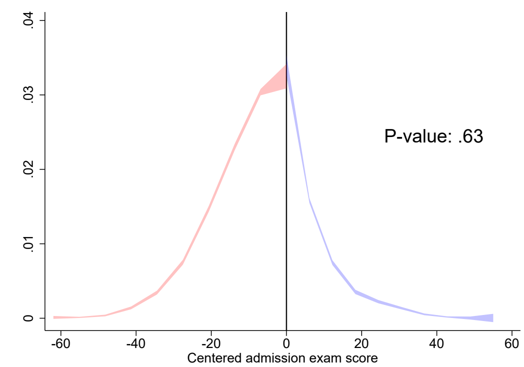
(a) 2005



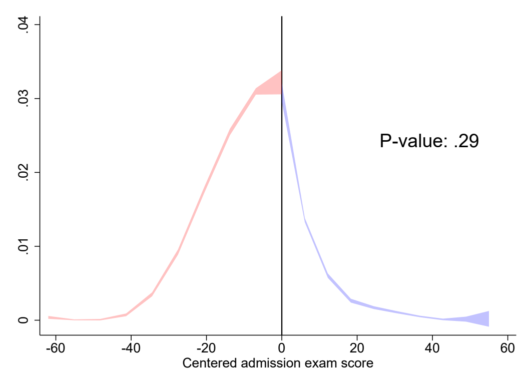
(b) 2006



(c) 2007



(d) 2008



(e) 2009

NOTE: This figure shows the density of the centered running variable for each admission cohort. The lines and shaded areas indicate the 95% confidence intervals for robust inference. The vertical line indicates the admission threshold.

C Covariates by cohort

Table C.1: Girl

	2005	2006	2007	2008	2009
RD_Estimate	-0.017	0.014	0.012	0.022	0.025
	(0.020)	(0.025)	(0.023)	(0.023)	(0.028)
Optimal BW	12.177	8.676	10.922	9.602	11.005
Mean	0.422	0.407	0.448	0.459	0.438
N	8,374	7,021	9,534	10,276	11,584

NOTE: This table shows RD estimates following the methods in Calonico et al. (2014) and the associated software package *rdrobust*. Each column indicates a different admission cohort. For each cohort, we also report the associated optimal bandwidth, the average outcome for the marginally rejected students, and the number of observations. Robust standard errors in parentheses.

Table C.2: GPA

	2005	2006	2007	2008	2009
RD_Estimate	0.009	-0.017	-0.014	0.042	-0.014
	(0.030)	(0.030)	(0.025)	(0.028)	(0.034)
Optimal BW	10.759	10.573	12.818	11.679	11.138
Mean	8.205	8.218	8.254	8.258	8.325
N	7,429	8,155	10,740	11,871	11,584

NOTE: This table shows RD estimates following the methods in Calonico et al. (2014) and the associated software package *rdrobust*. Each column indicates a different admission cohort. For each cohort, we also report the associated optimal bandwidth, the average outcome for the marginally rejected students, and the number of observations. Robust standard errors in parentheses.

Table C.3: Father's education

	2005	2006	2007	2008	2009
RD_Estimate	-0.010	0.026	0.020	0.019	-0.010
	(0.021)	(0.017)	(0.019)	(0.021)	(0.021)
Optimal BW	9.483	16.071	10.152	8.908	9.780
Mean	0.308	0.288	0.328	0.343	0.346
N	6,398	9,825	8,767	8,708	9,228

NOTE: This table shows RD estimates following the methods in Calonico et al. (2014) and the associated software package *rdrobust*. Each column indicates a different admission cohort. For each cohort, we also report the associated optimal bandwidth, the average outcome for the marginally rejected students, and the number of observations. Robust standard errors in parentheses.

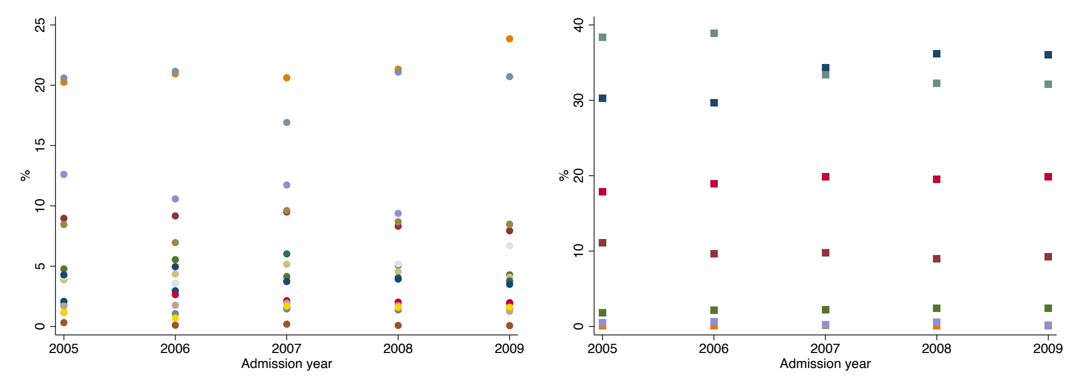
Table C.4: Siblings

	2005	2006	2007	2008	2009
RD_Estimate	0.027	-0.048	-0.025	-0.040	0.070
	(0.064)	(0.062)	(0.050)	(0.051)	(0.046)
Optimal BW	9.692	7.144	8.870	7.456	9.190
Mean	2.080	2.011	1.968	1.885	1.829
N	6,869	6,346	8,108	8,413	10,021

NOTE: This table shows RD estimates following the methods in Calonico et al. (2014) and the associated software package *rdrobust*. Each column indicates a different admission cohort. For each cohort, we also report the associated optimal bandwidth, the average outcome for the marginally rejected students, and the number of observations. Robust standard errors in parentheses.

D First and next-best schools

Figure D.1: Local school ranking (weights)



(a) Preferred elite schools (b) Next-best non-elite subsystems

NOTE: This figure shows the weights of preferred elite schools and next-best non-elite subsystems for students within ten points of their respective elite school admission cut-offs. Panel (a) shows the weights of the sixteen elite schools over time (circles). Panel (b) shows the weights of the six non-elite subsystems over time (squares).

Table D.1: Math restricted sample

	2005	2006	2007	2008	2009
RD_Estimate	0.195**	0.086	-0.051	-0.075	-0.115*
	(0.085)	(0.077)	(0.078)	(0.055)	(0.063)
Optimal BW	8.825	11.636	11.068	10.445	10.385
Mean	0.132	0.123	0.149	0.192	0.193
N	1,576	2,083	2,408	2,666	2,443

H0: 2005=2006=2007=2008=2009, p-value: 0.000

Linear trend: coef -0.073, p-value 0.000

NOTE: This table shows RD estimates following the methods in Calonico et al. (2014) and the associated software package *rdrobust*. We restrict the sample to exclude students assigned to the two most popular non-elite subsystems when marginally rejected from an elite school. Each column indicates a different admission cohort. For each cohort, we also report the associated optimal bandwidth, the average outcome for the marginally rejected students, and the number of observations. Robust standard errors in parentheses.

E Cross-fitting

To reduce the risk that our value-added estimates are driven by overfitting, we implement a simple cross-fitting procedure. We begin by randomly partitioning the full sample into 20 roughly equal groups. For each group $g = 1, \ldots, 20$, we (i) exclude all observations in group g; (ii) estimate the value-added model in Equation 2 using only the remaining 19 groups; and (iii) use the resulting coefficients to predict value-added for the schools attended by students in group g. This ensures that, for each student, the value-added assigned to their school is always computed from a model estimated on other students, rather than on the same data used to evaluate the RD effect. Repeating this procedure for every group yields cross-fitted value-added measures for all students, where each prediction is based on a model estimated on a sample that does not include that student. We then use these cross-fitted value-added estimates as the outcome variable in our RDD specifications in Section 6. The corresponding results are reported in Table 11.

F EB shrinkage

Table F.1: Value-added (Empirical Bayes estimates)

	2005	2006	2007	2008	2009
RD_Estimate	0.261***	0.252***	0.129***	0.052***	0.011
	(0.024)	(0.020)	(0.026)	(0.015)	(0.024)
Optimal BW	14.840	16.090	13.338	14.069	16.334
Mean	-0.030	-0.046	-0.041	-0.047	-0.073
N	9,151	10,637	11,274	13,737	14,937

H0: 2005=2006=2007=2008=2009, p-value: 0.000

Linear trend: coef -0.069, p-value 0.000

NOTE: This table shows RD estimates following the methods in Calonico et al. (2014) and the associated software package *rdrobust*. Each column indicates a different admission cohort. For each cohort, we also report the associated optimal bandwidth, the average outcome for the marginally rejected students, and the number of observations. Robust standard errors in parentheses.

G Other school characteristics

Table G.1: Peers middle school GPA

	2005	2006	2007	2008	2009
RD_Estimate	0.510***	0.494***	0.501***	0.475***	0.476***
	(0.023)	(0.024)	(0.021)	(0.023)	(0.020)
Optimal BW	14.134	13.333	14.278	16.773	13.920
Mean	7.829	7.867	7.900	7.912	7.949
N	9,151	9,545	11,806	14,817	13,084

H0: 2005=2006=2007=2008=2009, p-value: 0.881

Linear trend: coef -0.007, p-value 0.298

NOTE: This table shows RD estimates following the methods in Calonico et al. (2014) and the associated software package *rdrobust*. Each column indicates a different admission cohort. For each cohort, we also report the associated optimal bandwidth, the average outcome for the marginally rejected students, and the number of observations. Robust standard errors in parentheses.

Table G.2: Teachers per pupil

	2005	2006	2007	2008	2009
RD_Estimate	0.024***	0.017***	0.047***	0.017***	0.020***
	(0.005)	(0.002)	(0.004)	(0.003)	(0.003)
Optimal BW	15.559	13.150	11.558	20.510	13.734
Mean	0.056	0.051	0.026	0.053	0.050
N	9,415	9,297	9,927	15,963	12,745

H0: 2005=2006=2007=2008=2009, p-value: 0.000

Linear trend: coef -0.001, p-value 0.231

NOTE: This table shows RD estimates following the methods in Calonico et al. (2014) and the associated software package *rdrobust*. Each column indicates a different admission cohort. For each cohort, we also report the associated optimal bandwidth, the average outcome for the marginally rejected students, and the number of observations. Robust standard errors in parentheses.

Table G.3: Female teachers per pupil

	2005	2006	2007	2008	2009
RD_Estimate	0.004*	0.003*	0.014***	0.001	0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Optimal BW	14.191	16.980	13.065	16.280	14.020
Mean	0.021	0.019	0.010	0.021	0.020
N	9,045	10,349	11,010	14,311	13,390

H0: 2005=2006=2007=2008=2009, p-value: 0.000

Linear trend: coef -0.001, p-value 0.135

NOTE: This table shows RD estimates following the methods in Calonico et al. (2014) and the associated software package *rdrobust*. Each column indicates a different admission cohort. For each cohort, we also report the associated optimal bandwidth, the average outcome for the marginally rejected students, and the number of observations. Robust standard errors in parentheses.

Table G.4: Full time teachers per pupil

	2005	2006	2007	2008	2009
RD_Estimate	0.010***	0.011***	0.010***	0.012***	0.009***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Optimal BW	13.681	12.126	13.795	14.412	13.422
Mean	0.012	0.011	0.011	0.011	0.011
N	8,647	8,873	11,010	13,282	12,745

H0: 2005=2006=2007=2008=2009, p-value: 0.356

Linear trend: coef 0.000, p-value 0.875

NOTE: This table shows RD estimates following the methods in Calonico et al. (2014) and the associated software package *rdrobust*. Each column indicates a different admission cohort. For each cohort, we also report the associated optimal bandwidth, the average outcome for the marginally rejected students, and the number of observations. Robust standard errors in parentheses.

Table G.5: High education teachers per pupil

	2005	2006	2007	2008	2009
RD_Estimate	-0.000	0.001***	0.001**	0.000	0.004
	(0.000)	(0.000)	(0.001)	(0.000)	(0.002)
Optimal BW	9.806	16.151	16.671	16.190	12.456
Mean	0.002	0.002	0.002	0.002	0.002
N	6,831	10,349	12,394	14,311	12,029

H0: 2005=2006=2007=2008=2009, p-value: 0.014

Linear trend: coef 0.001, p-value 0.115

NOTE: This table shows RD estimates following the methods in Calonico et al. (2014) and the associated software package *rdrobust*. Each column indicates a different admission cohort. For each cohort, we also report the associated optimal bandwidth, the average outcome for the marginally rejected students, and the number of observations. Robust standard errors in parentheses.

Table G.6: Classrooms

	2005	2006	2007	2008	2009
RD_Estimate	-0.002	0.004***	0.001**	0.002**	0.001
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Optimal BW	13.817	10.645	14.069	12.807	12.602
Mean	0.023	0.022	0.022	0.021	0.021
N	8,647	7,955	11,525	12,126	12,029

H0: 2005=2006=2007=2008=2009, p-value: 0.023

Linear trend: coef 0.001, p-value 0.139

NOTE: This table shows RD estimates following the methods in Calonico et al. (2014) and the associated software package *rdrobust*. Each column indicates a different admission cohort. For each cohort, we also report the associated optimal bandwidth, the average outcome for the marginally rejected students, and the number of observations. Robust standard errors in parentheses.

H Log-level

We estimate Equation 2 using the logarithm of test score, and then calculate the fitted value and residuals.

Table H.1: $Z'_{jt}\hat{\theta}_t$, log-level

	2005	2006	2007	2008	2009
RD_Estimate	0.005***	-0.019***	-0.016***	-0.015***	-0.010***
	(0.001)	(0.002)	(0.002)	(0.002)	(0.003)
Optimal BW	10.481	13.123	13.165	11.844	11.193
Mean	6.223	6.229	6.260	6.285	6.316
N	7,349	9,405	11,147	11,676	11,538

H0: 2005=2006=2007=2008=2009, p-value: 0.000

Linear trend: coef -0.002, p-value 0.012

NOTE: This table shows RD estimates following the methods in Calonico et al. (2014) and the associated software package *rdrobust*. Each column indicates a different admission cohort. For each cohort, we also report the associated optimal bandwidth, the average outcome for the marginally rejected students, and the number of observations. Robust standard errors in parentheses.

Table H.2: $\hat{\eta}_{jt}$, log-level

	2005	2006	2007	2008	2009
RD_Estimate	0.041***	0.065***	0.038***	0.025***	0.013***
	(0.004)	(0.004)	(0.004)	(0.003)	(0.004)
Optimal BW	16.117	15.494	14.302	11.154	14.780
Mean	-0.007	-0.007	-0.012	-0.006	-0.014
N	9,747	10,182	11,671	11,676	13,697

H0: 2005=2006=2007=2008=2009, p-value: 0.000

Linear trend: coef -0.010, p-value 0.000

NOTE: This table shows RD estimates following the methods in Calonico et al. (2014) and the associated software package *rdrobust*. Each column indicates a different admission cohort. For each cohort, we also report the associated optimal bandwidth, the average outcome for the marginally rejected students, and the number of observations. Robust standard errors in parentheses.

I Replication

In order to have an initial reference for our estimates, we replicate the results obtained by Dustan et al. (2017) using application cohorts 2005-2006. Table I.1 presents the results of this exercise.

Table I.1: Effects of elite assignment, 2005-2006

	Dropout	Math	Spanish
admit	0.094***	0.197***	0.028
	(0.017)	(0.030)	(0.031)
N	17,850	11,959	11,216

The effect of elite assignment on the probability of dropout is identical to the result they obtained, and it is also statistically significant at 99%. The effect on the mathematics test score is slightly smaller than their result (their point estimate is 0.246) but is also statistically significant at 99%. Lastly, the effect on the Spanish test score is similar in magnitude, and it is also not statistically significant.