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

Salvador Navarro † Marco Pariguana ‡

January 21, 2025

Abstract

We study whether the academic effects of being marginally admitted to an elite science school depend on the admission year as a reflection of how school characteristics change over time. We take advantage of five years (2005-2009) of administrative data on a large centralized high school admission system. We find that the effect on mathematics test scores at the end of high school decreases each year, starting positive and statistically significant in 2005 and ending not significant by 2009. We show that the discontinuous jumps in peer quality and other school characteristics induced by elite school admission have not systematically changed. However, the gains in school quality decreased, affecting the treatment definition. Varying relative school quality limits the external validity of otherwise internally valid estimates.

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

JEL codes: I21, I24, I28, J24.

*We thank seminar participants at 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.

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

‡School of Economics, University of Edinburgh. E-mail: mparigua@ed.ac.uk.

1 Introduction

In many centralized education systems, a considerable proportion of students apply to attend high-performing schools (hereafter referred to as elite schools). They hope that these schools' characteristics, such as a more stimulating peer environment, superior infrastructure, highly qualified teachers, and better curriculum, will lead to greater academic success, and ultimately to better outcomes overall. However, it remains unclear whether their expected benefits materialize and which school characteristics contribute to or potentially inhibit these anticipated outcomes.

Furthermore, the benefits could depend on the admission year. To the extent that school characteristics and their productivity are not fixed over time, the effects of elite schools may also vary. In this context, even internally valid estimates may lack external validity and be of limited policy use. For example, consider the case of a policymaker interested in increasing the seats offered by elite schools. Whether past estimates of the effect of elite schools on academic outcomes are informative for such a policy ultimately depends on whether the school characteristics behind the effects differ from those in place when obtaining the estimates.

In this paper, we study whether and why the estimated effects of elite schools on academic outcomes change over time for a fixed set of elite schools and admission policies. The answer to this question could also help explain differing effects across education markets to the extent that they may differ in what the estimated effects are measuring.

We focus on the case of Mexico City over five years, from 2005 to 2009, and work with yearly administrative data from its centralized high school admission system. To measure academic outcomes, we combine the admission records with the records of an exit exam students take during the last quarter of high school. To measure school inputs, we use yearly school census data. This rich dataset presents three advantages for the analysis. First, the admission process creates discontinuities around unpredictable cut-offs that randomize students between admission to elite and non-elite schools, ensuring internally valid estimates. Second, the centralized market is large, with around 250,000 applicants yearly. As the admission process is the same throughout our study period, the large number of applicants allows us to estimate effects *for each* admission year without running into statistical precision problems. Third, data on yearly school inputs allows us to assess their role in the estimated effects of elite schools.

Estimating the effects of elite schools on academic outcomes is challenging because gaining admission to an elite school may be correlated with unobservable student characteristics, such as ability. Previous literature has addressed this problem by exploiting how centralized educational systems operate and using Regression Discontinuity Designs (RDDs) to estimate causal effect(s). The intuition behind this identification strategy is that oversubscribed schools in centralized systems generate admission cut-offs that allow for comparing academic outcomes between marginally admitted and marginally rejected students who are ex-ante equivalent other than their admitted or rejected status. However, as non-parametric estimation only uses data on individuals around admission cut-offs, a common practice is pooling different cohorts of applicants to obtain an average estimate. As we show, this practice may mask substantial heterogeneity when the relative quality of elite and non-elite schools varies over time.

We take advantage of the large centralized system in Mexico City and estimate separate RDDs for each admission cohort between 2005-2009. Our main results indicate that the effect of being marginally admitted to an elite high school on end-of-high school math test scores depends on the admission year and decreases across time. In particular, the effect monotonically decreases yearly and changes from positive and statistically significant in 2005 to insignificant by 2009. We also find that marginal admission to an elite school increases the probability of dropping out, but this effect is roughly constant over time. This alleviates concerns regarding time changes in sample selection as an explanation for the time-varying effects on test scores. In addition, we show that the change in peer quality induced by marginal admission to an elite school is roughly constant over time.

To explain the trend of the effects on test scores, we explore changes over time in the relative academic quality of elite and non-elite high schools. We first estimate individual high school value-added on mathematics for each year and show that the gains in school value-added induced by marginal admission to an elite school decreased over time. We then decompose school value-added into observable school characteristics, parameters that map school characteristics into outputs, and unobservable school characteristics. Although elite school admission implies a drastic change in several school characteristics, such as better peers and more qualified teachers, these changes have stayed the same over time. Most of the changes over time come from how these school characteristics map into school value-added and changes in unobservable school characteristics. A potential explanation behind our results is that during our study period, the government initiated a reform of curricular alignment between high schools, including elite and non-elite schools. Such a policy could affect how school characteristics map into school quality and affect schools in unobservable ways.

Regarding our research question, the results tell us that the effect of marginal admission to an elite school depends on the change in school quality induced by crossing their admission thresholds. Furthermore, school quality is not time-invariant as schools may change in observable and unobservable ways. Therefore, RDD parameter estimates of the sort studied here are time-specific, constraining their external validity.

Our work contributes to two strands of the literature. First, it contributes to the extensive literature that studies the academic effects of elite/selective schools.[1] For the Mexico City context Dustan et al. (2017) find large positive and statistically significant gains in mathematics test scores from marginal admission to elite science schools. However, as we show in this paper, this is only the case for the two cohorts of applicants they study. Beuermann and Jackson (2022) review the results from this literature and finds that estimates go from positive and statistically significant to not significant. The heterogeneity of these results could be due to overall institutional differences across the markets being studied (Hanushek, 2021). Yet, we find that estimates for a single market over time range over the previously obtained estimates across markets. Thus, differences in school quality between first and next-best schools across markets could also explain the diversity of results.

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

[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).

estimate production function parameters. Thus, when studying the effects of elite schools, even if a study convincingly deals with selection bias and obtains some effect, it still needs to be determined why there is an effect. Some prior literature has interpreted the effect of marginal admission to a selective school as the effect of a discontinuous change in peer quality (Jackson, 2010; Abdulkadiroğlu et al., 2014; Dobbie and Fryer Jr, 2014). However, in our setting, the effect of elite schools on test scores changes even though the discontinuous change in peer quality remains constant over time. More generally, following Altonji and Mansfield (2018) framework, the effects of schools on individual academic outcomes depend on individual inputs, group-level inputs, and inputs productivity. Thus, as group-level inputs or their productivity may vary across time and context, there is no reason why a given set of estimates should apply to any other period or context.

The remainder of the paper proceeds as follows. Section 2 describes the institutional background of the market under study. Section 3 describes the data, samples, and outcomes used for the analysis. Section 4 describes the empirical strategy and offers evidence on the validity of the research design. Section 5 presents the main results of the paper. Sections 6 and 7 explore mechanisms. 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.

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 admission process remained unchanged during our study period (2005-2009).

The admission exam evaluates students on the material covered during middle school and in verbal and mathematical reasoning. The exam score takes integer values from 31 to 128. Students who score less than 31 are excluded from the admission process.

The matching process follows the serial dictatorship mechanism. This is a particular case of the student proposing a deferred acceptance mechanism where all schools use the same ranking of students. In Mexico City, all applicants are ranked based on their admission exam scores. The highest-scoring students are assigned to their first choices. Then, in admission score descending order, students are assigned to their highest-ranked schools with available seats.

Every high school in Mexico City belongs to one of nine subsystems. Subsystems are administrative units that manage a subset of schools. Two of the subsystems are considered elite: the IPN and the UNAM. High schools in the IPN and UNAM subsystems are affiliated with the country's two most prestigious public universities. This group of high schools are commonly referred as elite schools.

As an additional constraint, admission to any school affiliated with IPN and UNAM requires students to have a middle school GPA higher than 7/10. In practice, this constraint is not binding as the GPA requirement is satisfied by more than 90% of students each year. Also, the minimum GPA to obtain a middle school certificate is 6/10.

From 2007 until 2014, during the last quarter of high school, students from public and private schools took a standardized exam that evaluated them on mathematics and Spanish. The government mainly used this test to assess high school-level performance. Students from UNAM-affiliated high schools did not participate in this exam. For the rest of the paper, we refer to this test as the exit exam.

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 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.

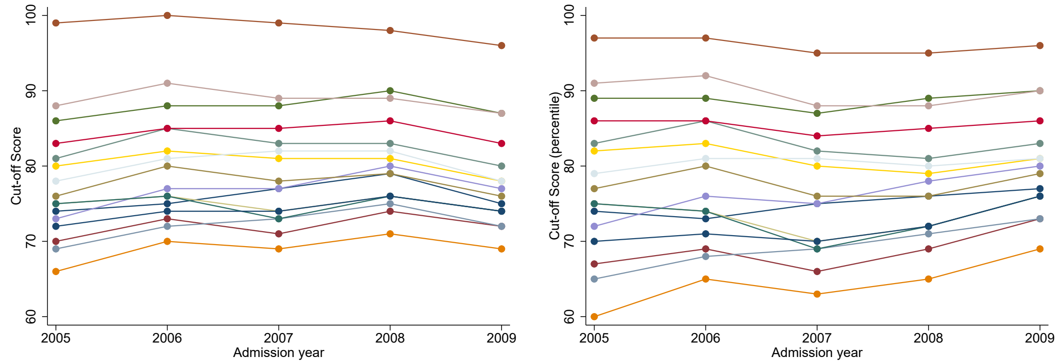
Students admitted to UNAM-affiliated schools do not participate in the exit exam, so we do not have outcomes for these students. In the analysis we compare students admitted to IPN schools with students admitted to non-IPN and non-UNAM schools. Throughout the analysis period, the IPN subsystem has 16 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 1 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 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 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 2. 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. The admission system creates stratification across schools by initial ability, and this has not changed much over time. However, this does not mean elite schools have a causal effect on exit exam test scores for any particular

Figure 1: Elite cut-offs

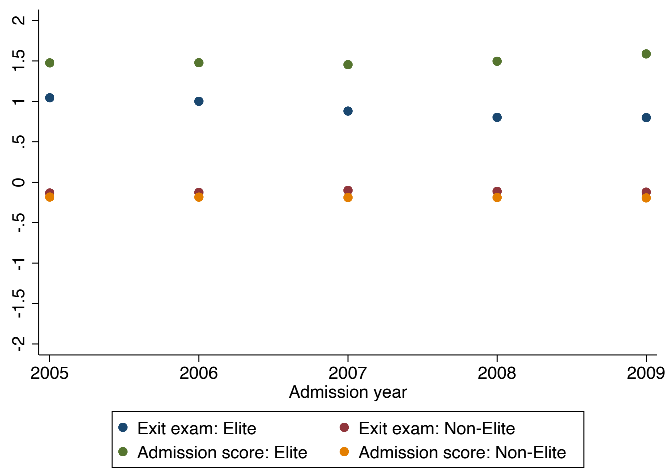


(a) Cut-offs levels

(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.

Figure 2: Exams scores



Note: This figure shows the admission and exit exam scores for students assigned to elite and non-elite schools. We standardize each score within the distribution of scores for each admission cohort.

year. We use the empirical strategy outlined in the next section to separate the effect of elite schools on academic outcomes from what simply reflects the selection of better students.

Regarding school-level information, we obtain school characteristics from each year's school census (Formato 911). The school census tracks information on all schools in Mexico at the campus level. The close to 600 schools in our region of analysis are distributed into around 300 campuses, each belonging to one of the nine subsystems. The school census data includes 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 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 compare the outcomes of students with the same local institution ranking (i.e., elite$\succ$non-elite), some of whom are admitted to their first best while others gain admission to their next best institution. 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 the local institution ranking. We use a Regression

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.

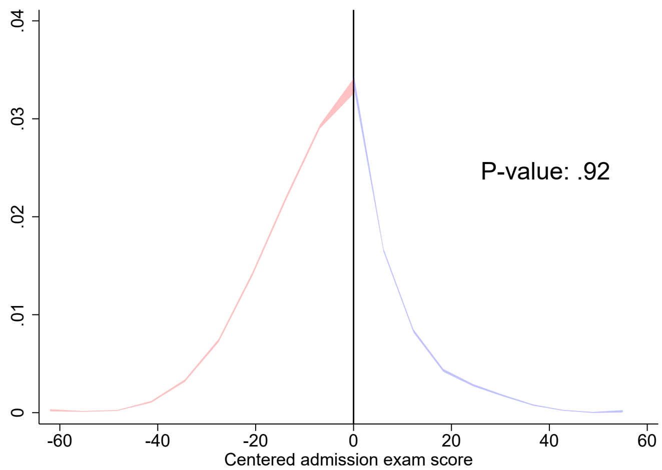
Discontinuity Design (RDD) and focus on students close to the elite school admission cut-offs. Since there are 16 elite schools, we have 16 admission cut-offs. The intuition behind the identification strategy is that marginally admitted and rejected students from elite high schools have similar observable and unobservable characteristics.

Our empirical model is:

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

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. We include a dummy variable for elite school admission $admit_i$. We denote the admission exam score S_i and $\underline{s}_k$ is the elite school admission cut-off relevant to student i. In this specification $admit_i = 1$ when $S_i - \underline{s}_k \geq 0$.

Figure 3: Density



NOTE: This figure shows the density of the centered running variable. The vertical line indicates the admission threshold.

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, 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 guarantee the validity of an RDD. Figure 3 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.

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	40254.000	58423.000	51427.000	36187.000

NOTE: This table shows RD estimates following the methods in Calonico et al. (2014) and the associated software package *rdrobust*. Each column indicate 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 parenthesis.

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 family income, gender, parents' education, or students' GPA in middle school. We report the estimates on predetermined covariates for each admission cohort in Appendix C.

Overall, we find no evidence of manipulation in the pooled sample or for each separate admission cohort, which we take as evidence 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). However, pooling multiple years together may be masking substantial heterogeneity in the effects by year. Importantly, our sample sizes are large enough to allow us to implement the same empirical design on a yearly basis.

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	31070.000	27532.000	53088.000

NOTE: This table shows RD estimates following the methods in Calonico et al. (2014) and the associated software package *rdrobust*. Each column indicate 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 parenthesis.

To study if 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 goes from being 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. Dustan et al. (2017) only study admission cohorts 2005 and 2006, finding similar effects on mathematics performance for those two cohorts.[2] However, this effect is not time invariant and disappears for later admission cohorts.

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

[2] Their sample selection criteria are different from ours. Appendix D 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 indicate 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 parenthesis.

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 indicate 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 parenthesis.

and find no statistically significant linear time trend in them.

Overall, the results show the time specificity of the estimated effects on mathematics performance. The next step is to understand why this effect experiences such a dramatic change over time.

6 Why did the effect change over time?

We only observe end of high school test scores for those who participate in the exit exam. Therefore, time varying effects on exam participation could potentially explain time varying effects on test scores. To asses this possibility, we estimate the year-specific effect of marginal admission to an elite school on the probability of taking the exit exam. We follow our empirical specification in Equation 1.

The results in Table 6 indicate that the estimated coefficients are negative, between 8 to 10 percentage points, and always statistically significant. However, there is no clear pattern in the effects. Furthermore, we cannot reject the equality of coefficients over time, and their linear trend is not statistically significant. We take this as evidence against differential selection over time. Under non-differential selection over time, methodologies, such as the one in Lee (2009), to correct for bias would shift the estimates on test scores in the same direction without affecting the time trend on the effect on mathematics test scores.

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 increased peer quality relative to their counterfactual alternatives. For example, this occurs when selective schools admit students based on skill measures, as in Mexico City. Therefore, if the discontinuous jump in peer quality has changed over time, then we would expect time-varying effects on test scores.

In Table 7, 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 average peer quality experienced by those admitted to elite

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 indicate 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 parenthesis.

Table 7: Peers 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 indicate 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 parenthesis.

schools cannot explain the decreasing effect on test scores.

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

We then use the value-added estimated parameters as the outcome variable in our RDD estimations. We aim to capture if there is a discontinuous jump in school value-added between students marginally admitted to elite schools and those marginally rejected. In addition, we aim to measure if the effect on school value-added has changed over time. In Table 7 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. We reject the equality of coefficients across time and find a negative and statistically significant linear trend in the estimated coefficients.

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 capture school characteristics and their productivities. Therefore, if the school characteristics or their productivities change over time, the value-added estimated parameters will also change over

Table 8: 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)
Optimal BW	14.580	15.973	13.647	14.039	16.147
Mean	-0.033	-0.050	-0.044	-0.051	-0.076
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 indicate 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 parenthesis.

time. We consider the following equation relating school characteristics to school value-added:

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

where Z_j 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 specification.

In Table 9 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 monotonically decreasing. We reject equality of coefficients over time and we find they have a statistically significant negative linear trend.

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

Table 9: $Z'_{jt}\hat{\theta}_t$

	2005	2006	2007	2008	2009
RD_Estimate	0.018***	-0.101***	-0.063***	-0.070***	-0.073***
	(0.003)	(0.010)	(0.014)	(0.012)	(0.019)
Optimal BW	11.075	13.817	12.180	11.817	11.603
Mean	0.012	-0.008	0.016	-0.005	0.004
N	7,812	9,297	10,488	11,500	11,296

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

Linear trend: coef -0.013, p-value 0.006

Note: This table shows RD estimates following the methods in Calonico et al. (2014) and the associated software package *rdrobust*. Each column indicate 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 parenthesis.

acteristics (Z_{jt}), time changes in productivities (θ_t), or both. In Appendix A, we use each school characteristic as the outcome variable in the RDD specification and find no effect pattern or time trend. Therefore, we argue that the time-varying effect on the fitted value is due mainly to changes in the combined productivities of school characteristics.

A possible explanation behind the change in the productivities of school characteristics for elite and non-elite schools is that a curriculum alignment policy was in place during our study period. Notice that such a policy would not likely change the school characteristics levels but would change how productive the same levels are in creating school quality. Therefore, a curriculum alignment policy that imposes similar curriculums at elite and non-elite schools would align the productivities of school characteristics across sectors and decrease the treatment effect of marginal admission to an elite school.

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 10, 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

Table 10: $\hat{\eta}_{jt}$

	2005	2006	2007	2008	2009
RD_Estimate	0.252***	0.358***	0.190***	0.123***	0.087***
	(0.024)	(0.026)	(0.023)	(0.015)	(0.024)
Optimal BW	16.921	16.579	14.010	11.627	14.730
Mean	-0.044	-0.046	-0.057	-0.043	-0.081
N	9,747	10,349	11,525	11,500	13,390

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

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 indicate 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 parenthesis.

statistically significant positive linear trend in coefficients. Therefore, time changes in unobserved school characteristics could also be behind our time-varying effects on mathematics test scores.

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.

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. A plausible explanation for the changes in school quality gains is a curriculum alignment policy in place during our study period. Such policy affected how school inputs mapped into school value-added. In addition, there were also changes in the school value-added part that were unexplained by observed school characteristics.

Our results highlight that when studying the effects of elite/selective schools it is important to understand what is the estimated parameter measuring in each particular context. In this sense, caution when generalizing results from within-country analyses to other countries should also be extended to generalizing within-period studies to different time periods.

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.

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.

Hanushek, Eric A, "Addressing cross-national generalizability in educational impact evaluation," *International Journal of Educational Development*, 2021, *80*, 102318.

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.

Lee, David S, "Training, wages, and sample selection: Estimating sharp bounds on treatment effects," *The Review of Economic Studies*, 2009, *76* (3), 1071–1102.

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 Other school characteristics

Table 11: Change in average peers 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)
b-c CI	[.457 ; .555]	[.451 ; .551]	[.461 ; .545]	[.435 ; .526]	[.441 ; .522]
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=2009, p-value: 0.445

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

Linear trend: coef -0.007, p-value 0.298

Table 12: 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)
b-c CI	[.015 ; .034]	[.012 ; .021]	[.039 ; .053]	[.012 ; .024]	[.014 ; .026]
Optimal BW	15.072	13.150	11.558	20.510	13.734
Mean	0.056	0.051	0.026	0.053	0.050
N	9,343	9,297	9,927	15,963	12,745

Standard errors in parentheses

H0: 2005=2009, p-value: 0.395

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

Linear trend: coef -0.001, p-value 0.215

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 13: 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)
b-c CI	[-.001 ; .008]	[-.001 ; .006]	[.009 ; .018]	[-.003 ; .005]	[-.003 ; .005]
Optimal BW	13.835	16.980	13.065	16.280	14.020
Mean	0.021	0.019	0.010	0.021	0.020
N	8,581	10,349	11,010	14,311	13,390

Standard errors in parentheses

H0: 2005=2009, p-value: 0.411

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

Linear trend: coef -0.001, p-value 0.138

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 14: 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)
b-c CI	[.007 ; .012]	[.008 ; .014]	[.007 ; .013]	[.01 ; .016]	[.006 ; .012]
Optimal BW	13.553	12.126	13.795	14.412	13.422
Mean	0.012	0.011	0.011	0.011	0.011
N	8,581	8,873	11,010	13,282	12,745

Standard errors in parentheses

H0: 2005=2009, p-value: 0.828

H0: 2005=2006=2007=2008=2009, p-value: 0.344

Linear trend: coef 0.000, p-value 0.854

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 15: 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)
b-c CI	[-.001 ; 0]	[0 ; .002]	[0 ; .002]	[-.001 ; .001]	[-.001 ; .008]
Optimal BW	9.735	16.151	16.671	16.190	12.456
Mean	0.001	0.002	0.002	0.002	0.002
N	6,781	10,349	12,394	14,311	12,029

Standard errors in parentheses

H0: 2005=2009, p-value: 0.108

H0: 2005=2006=2007=2008=2009, p-value: 0.015

Linear trend: coef 0.001, p-value 0.118

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 16: Classrooms

	2005	2006	2007	2008	2009
RD_Estimate	-0.001	0.004***	0.001**	0.002**	0.001
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Optimal BW	13.854	10.645	14.069	12.807	12.602
Mean	0.023	0.022	0.022	0.021	0.021
N	8,581	7,955	11,525	12,126	12,029

Standard errors in parentheses

H0: 2005=2009, p-value: 0.310

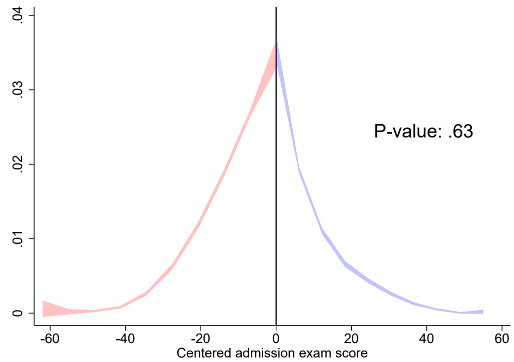
H0: 2005=2006=2007=2008=2009, p-value: 0.026

Linear trend: coef 0.001, p-value 0.142

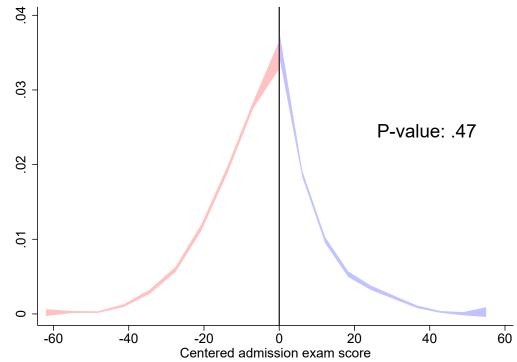
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

B Density by cohort

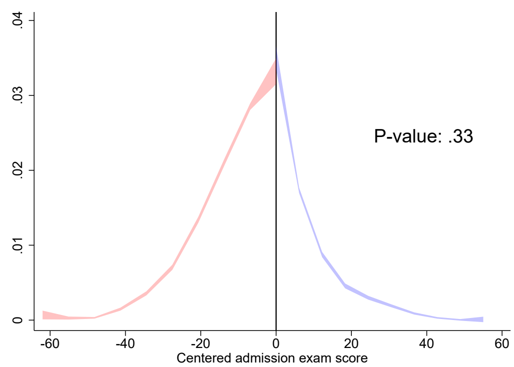
Figure 4: Density test



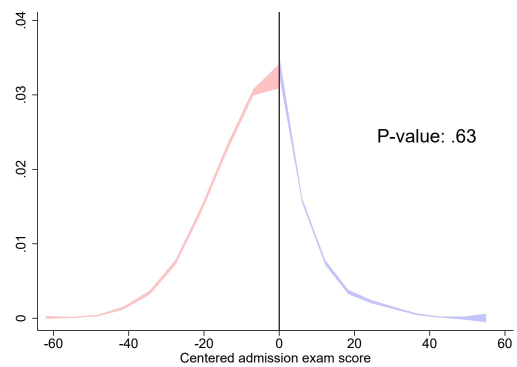
(a) 2005



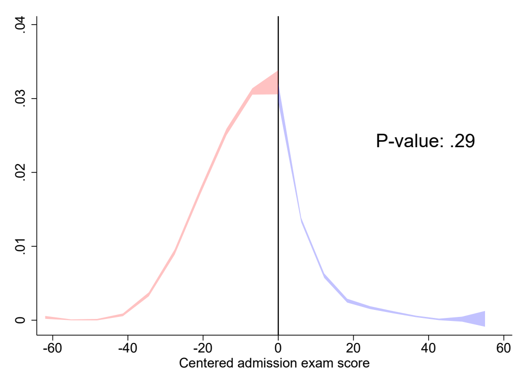
(b) 2006



(c) 2007



(d) 2008



(e) 2009

C Covariates by cohort

Table 17: 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)
b-c CI	[-.061 ; .031]	[-.042 ; .075]	[-.04 ; .066]	[-.033 ; .075]	[-.04 ; .086]
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

Standard errors in parentheses

H0: 2005=2009, p-value:

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

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 18: 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)
b-c CI	[-.061 ; .083]	[-.092 ; .051]	[-.078 ; .035]	[-.012 ; .113]	[-.096 ; .059]
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

Standard errors in parentheses

H0: 2005=2009, p-value:

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

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 19: Father 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)
b-c CI	[-.067 ; .031]	[-.01 ; .066]	[-.026 ; .066]	[-.025 ; .073]	[-.064 ; .038]
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

Standard errors in parentheses

H0: 2005=2009, p-value:

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

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 20: 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)
b-c CI	[-.127 ; .174]	[-.214 ; .066]	[-.157 ; .077]	[-.177 ; .062]	[-.024 ; .186]
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

Standard errors in parentheses

H0: 2005=2009, p-value:

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

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

D Replication

In order to have an initial reference for my estimates, I replicate the results obtained by Dustan et al. (2017) using COMIPEMS data for 2005-2006. Table 21 presents the results of this exercise.

Table 21: Effects of elite assigment, 2005-2006

	(1)	(2)	(3)
	dropout	math	spanish
admit	0.094***	0.197***	0.028
	(0.017)	(0.030)	(0.031)
N	17850	11959	11216

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 it is slightly smaller than their result (their point estimate is 0.246) but is also statistically significant at 99%. This difference comes from the way we merge the COMIPEMS and ENLACE datasets, since for some students that did not have the unique identifier (mostly elite students) they imposed the condition that they finished high school at the same high school were they ended up assigned. I do not impose this condition because I consider it creates selection problems when calculating the ITT effect. Instead, I solve the issue of missing identifiers by creating my own identifier for all the students (based on their names).

Lastly, the effect on the Spanish test score is similar in magnitude and it is also not statistically significant.