School Choice, Skill Measures, and Graduation*

Marco Pariguana[†] María Elena Ortega-Hesles [‡]

Abstract

This paper studies the effects of using both one-shot exam scores and GPAs to construct the priority order of a centralized education market. We use data from Mexico City, where seat rationing relies solely on a one-shot exam score. We first show that marginal admission to the most over-subscribed high schools decreases graduation for students with low GPAs and boys and has no effect for students with high GPAs and girls. We then study the effects of counterfactual priority orders that combine the one-shot exam score and GPA with different weights. The larger the weight on GPA, the larger the share of girls and low-SES students that get access to the most over-subscribed schools. However, using roughly equal weight on both skill measures maximizes the treatment effects on graduation.

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

JEL codes: I21, I24, I28, J24.

*We thank Salvador Navarro, Matteo Bobba, Todd Stinebrickner, Chao Fu, Maia Güell, Caterina Calsamiglia, Emily Cook, and Sylvia Blom for their helpful comments and suggestions. We are also grateful to the Executive Committee of COMIPEMS, the heads of the high school sub-systems, as well as to Ana María Aceves and Roberto Peña of the Mexican Ministry of Education (SEP) for making this study possible.

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

[‡]VIA Educación. E-mail: elenaortega@viaeducacion.org.

1 Introduction

In all centralized education systems some schools experience excess demand. Consequently, centralized systems need a way to ration the available seats [Shi, 2022]. Since using prices as a rationing mechanism is not feasible for K-12 public schools, policymakers define priority orders that solve the excess demand problem by determining who gets access to over-subscribed schools.[1]

Many centralized systems use a one-shot exam score as their priority order.[2] Understanding the consequences of this practice is important for several reasons. First, performance on a one-shot exam may be a noisy or incomplete measure of academic preparation. This could lead to a mismatch between students' preparation and some schools' academic requirements, affecting educational outcomes. Second, subpopulations with the same academic preparation may perform differently on a one-shot exam. This could lead to unequal access to highly demanded schools among well prepared students. For example, women and low-SES students tend to perform worse on one-shot exams than men or high-SES students but have otherwise similar (or better) preparation under alternate measures [Azmat et al., 2016; Arenas et al., 2021].

In this paper, we explore mismatch and equity of access by studying the centralized high school admission system in Mexico City. In this system, students' priority order is solely based on their scores in a one-shot admission exam. We study the following question: Can combining the one-shot exam with middle school GPA improve equity of access without adversely affecting graduation rates? We focus on GPA as a potential channel to improve student-school matches because previous literature shows that grades measure non-cognitive skills (e.g., conscientiousness) to a higher degree than one-shot exams do and that non-cognitive skills are important determinants of desirable educational outcomes [Stinebrickner and Stinebrickner, 2006; Duckworth et al., 2012; Borghans et al., 2016; Jackson, 2018]. However, as grades may have their own biases [Lavy, 2008; Hanna and Linden, 2012;

[1]Typical components of priority orders are siblings, residential zones, lotteries, standardized exams, and GPAs.

[2]For example, the centralized systems in Romania, Kenya, Trinidad and Tobago, Ghana, Barbados, and Mexico City use one-shot exams. In the US, selective schools in NYC rely solely on a standardized exam. In contrast, selective schools in Chicago and Boston combine standardized exams and GPA.

Lavy and Sand, 2018], we consider policies that combine standardized skill measures with non-standardized skill measures.

We use participants' administrative records from the centralized high school admission process in Mexico City. We complement these data with official high school graduation records (i.e., three to five years after admission) for students assigned to a school through the centralized process. For students not assigned and those who re-apply or transfer to other public or private schools, we use participation in a high school exit exam as a proxy for graduation. This unique dataset features two advantages for the analysis. First, it has information on the application and graduation of more than 250,000 students, allowing us to explore rich heterogeneity without losing much precision. Second, our dataset includes applicants' skill measures beyond the admission exam score, including their middle school GPAs and scores on a standardized exam used for school accountability.

We first shed light on the importance of the skills captured by GPA and their influence on students' probability of graduation when admitted to the most over-subscribed schools in the system (i.e., elite schools). Using a regression discontinuity design (RDD), we show that marginal admission to an elite school decreases the probability of graduation by six percentage points. However, students at the margin of admission to an elite school are heterogeneous with respect to their middle school GPAs. The correlation between the admission exam score and middle school GPA is 0.4. To study heterogeneity by GPA, we estimate effects separately for students with above- and below-median GPAs. We find that marginal admission to an elite school decreases the probability of graduation by twelve percentage points for students with low GPAs. For students with high GPAs, marginal admission to an elite school does not affect their probability of graduation. Notably, high and low GPA students experience a similar jump in peer quality when marginally admitted to elite schools, yet they experience considerably different outcomes.

We also estimate effects separately for boys and girls and find heterogeneous effects by gender. We find that boys experience a decrease in their graduation probability (ten percentage points), while girls are unaffected. This is consistent with previous findings showing that selective schools affect the educational attainment of boys and girls differently [Jackson, 2010; Clark, 2010; Deming et al., 2014]. We further show that a potential explanation behind these results is that girls have higher GPAs than boys at all levels of the admission exam score, including at the elite schools' admission cutoffs.

Our first set of findings imply that, for students at the margin of admission to the most over-subscribed schools, an assignment mechanism that relies on a single measure of skills affects educational outcomes by excluding important information about a student's academic potential, such as the information contained in GPA. In addition, given the pattern of heterogeneous results by GPA and gender, there is scope for increasing equality of access without affecting the graduation rate by taking GPAs into account.

Although policy relevant for certain policy counterfactuals (e.g., small increases in the number of offered seats), our RDD parameters may not be informative for policies that change the priority order for two reasons. First, changes in the priority order may lead to placement and displacement effects across all schools in the market (i.e., equilibrium effects). Second, who is affected by a change in the priority order ultimately depends on the interaction between the priority order, students' preferences, and school capacities. Thus, we rely on models of school choices and graduation outcomes to study the effects of counterfactual priority orders that combine the one-shot admission exam with GPA (or within school ranking by GPA) using different weights.

We estimate student preferences under the stability of the market equilibrium assumption [Fack et al., 2019].[3] This approach is robust to students potentially deviating from truth-telling behavior in their rank-ordered lists (ROLs) by omitting schools that they consider infeasible. For example, a high GPA student that performs poorly on one-shot exams may omit very selective schools in her ROL if she knows that the priority order gives no weight to her grades, but change this behavior as the priority order adds weight to GPA. Our counterfactuals take into consideration student preferences for schools that were infeasible in the status quo but became feasible under the counterfactual priority orders [Artemov et al., 2023].

To quantify the effects on graduation across the market, we estimate a graduation value-

[3]The matching algorithm is the Serial Dictatorship which incentivizes truthful revelation of preferences [Svensson, 1999]. However, in practice, some students may not reveal their preferences in their rank-ordered lists as there is a constraint on the number of schools they can list [Haeringer and Klijn, 2009; Calsamiglia et al., 2010] or due to application mistakes [Artemov et al., 2023; Hassidim et al., 2017].

added model. To deal with the non-random sorting of students across schools, we exploit the fact that school assignments only depend on students' ROLs and admission exam scores. To deal with sorting on preferences and application strategies, we control for students' ROLs. To deal with sorting on skills, we control for a set of skill measures other than the admission exam score. Intuitively, our empirical strategy assumes that conditional on skill measures and ROLs, students get assigned to schools for idiosyncratic reasons unrelated to potential outcomes. Our approach follows Angrist and Rokkanen [2015] method to extrapolate treatment effects for inframarginal applicants in an RDD design. For extrapolation, we assume that the admission exam score becomes ignorable once we condition on a set of covariates. We calculate the treatment effects of our policies by combining our model's estimated parameters with the characteristics of the students affected by changes in the priority order.

There are two important findings from the counterfactual analysis. First, the higher the weight on GPA, the higher the share of girls and low-SES students assigned to elite schools. We observe an increase in the share of girls because they have higher GPAs than boys and because they also prefer selective schools, so the counterfactual provides them with greater access to their preferred schools. We observe an increase in the share of low-SES students because the admission exam score is highly correlated with family income, whereas GPA is not. Second, the weights on the one-shot exam and GPA matter. Too little weight on GPA negatively affects the graduation rate of students reallocated from non-elite to elite schools, while too much weight on GPA diminishes the gains in graduation of students reallocated from elite to non-elite schools. For a central planner interested in equity in access and graduation, the optimal priority order puts roughly equal weights on the admission exam score and GPA.[4]

Our paper contributes to three strands of literature. First, it contributes to the literature on centralized education systems. Much of the previous literature considers school priorities as given and studies the consequences of using different matching mechanisms to allocate students to schools [Pathak, 2011; Agarwal and Somaini, 2020]. Yet, defining a priority structure is an integral part of the design of a centralized system. Shi [2022] and Abdulkadiroğlu et al. [2021] are the closest papers to ours. Their focus is on finding optimal priority

[4]Our findings are robust to alternately using within-school rankings by GPA instead of GPA.

structures in centralized education systems. We complement their work by also looking at students' downstream outcomes, such as graduation rates, which are crucial to assess the impact of mismatch within an assignment system. As Agarwal et al. [2020] and Larroucau and Rios [2020] highlights, it is essential to understand how assignment mechanisms perform when evaluated on outcomes of policymakers' concern beyond efficiency measures based on revealed preferences. Also, our counterfactual analysis follows some recent literature showing the importance of taking into account the congestion effects inherent in centralized markets when studying large-scale policy changes [Bobba et al., 2023; Larroucau et al., 2024].

Second, we contribute to the extensive literature studying the effects of elite/selective schools on educational outcomes.[5] Dustan et al. [2017] find that marginal admission to a subset of science schools in Mexico City increases dropout and that this effect is decreasing in GPA. They exclude from their analysis the most over-subscribed schools in the market, which are requested as a first choice by more than 50% of students. We complement their work in three ways. First, we study the effect of admission to the most over-subscribed schools in the market as these are the schools for which the priority order is most relevant in terms of seat rationing. Second, we explore heterogeneous results by gender and their connection with the heterogeneity by GPA. Third, we show that in equilibrium, the pattern of heterogeneous effects by GPA and gender allows for some policies to increase equity of access without negatively affecting the graduation rate.

Lastly, we contribute to the literature on using one-shot exams and grades in admission policies. Arenas and Calsamiglia [2022] study the effects of a policy change that increased the weight on standardized exams relative to high school grades in a university admission index. The change decreased the share of females at selective degrees and affected the females who were likely to do better in college than the males who benefited from the change. We complement their work by showing that over-reliance on a one-shot exam can also affect academically prepared, low-SES students. Bleemer [2021] shows that a grade-based top-percent policy for university admission in California promoted economic mobility without efficiency losses. Borghesan [2022] estimates a model that allows for endogenous responses

[5]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].

by students and universities and finds that banning a standardized exam for university admissions in the US does not improve diversity and affects the graduation rate. Our results are consistent with these findings to the extent that, as we show, using one skill measure or another is not better than combining them.

The remainder of the paper proceeds as follows. Section 2 describes the education system in Mexico City. Section 3 details the administrative data we use for the analysis. Section 4 contains the implementation and results of our RDDs. Section 5 contains the implementation and results of our counterfactuals. Section 6 concludes.

2 Education in Mexico City

The school system in Mexico has three levels: elementary, middle and high school. Elementary school is six years long, and middle and high school are three years each. The centralized high school education system in Mexico City encompasses the Federal District and 22 nearby urban municipalities in the State of Mexico. Most of the high school admission process participants are middle school students who reside in Mexico City and are in their last semester of middle school. Additional participants (less than 25%) attend middle schools outside of Mexico City, already have a middle school certificate, or are enrolled in adult education. In total, about 300,000 students participate in the admission system.

Public high schools in Mexico City belong to one of nine subsystems (Table 1). Each subsystem manages a different number of schools and offers its own curriculum. Two subsystems, SUB 1 and SUB 2 in Table 1, enjoy a high reputation, are affiliated with the two most prestigious public universities in Mexico City, and offer a more advanced curriculum. For the rest of the paper, we refer to the schools belonging to these subsystems as elite schools.

The first column of Table 1 shows the number of schools affiliated with each subsystem. The second column indicates that elite schools offer only 23% of the total number of seats in the system. The third column shows a high demand for elite schools; 63% of students list an elite school as their first option. Since elite schools are heavily over-subscribed, admission to elite schools is very competitive, which leads to these schools having high admission cut-off

Table 1: Subsystems in 2007

	Number of Schools	Seats (%)	First in ROL (%)	Admission Cut-Off
SUB 1	14	14.1	47.7	86.3
SUB 2	16	8.7	13.9	79.6
SUB 3	1	0.4	0.7	74.0
SUB 4	2	0.9	0.6	60.5
SUB 5	40	16.9	6.4	49.2
SUB 6	215	22.8	16.2	47.0
SUB 7	186	17.6	8.0	44.5
SUB 8	179	18.4	6.3	35.8
SUB 9	5	0.3	0.2	32.4
Total	658	100.0	100.0	45.0

NOTE: This table shows the aggregate supply, demand, and equilibrium cut-offs for the high school subsystems in Mexico City. The fourth column shows the average admission cut-offs of the schools in a given subsystem.

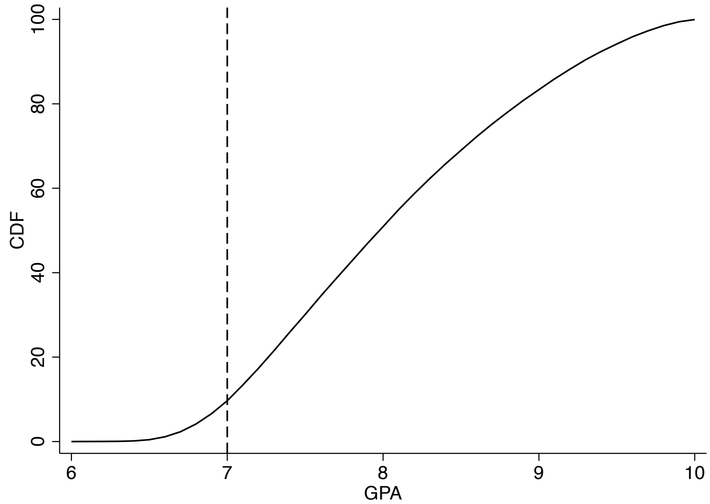
scores. We define an admission cut-off as the lowest score obtained by the students assigned to a given school in the previous admission cycle. The admission exam scores range from 31 to 128 points. The fourth column of Table 1 shows that elite schools' average admission cut-offs are the highest in the market.

The timeline of the application process is as follows. In February, students receive an information booklet describing the steps they need to follow. The information booklet also lists all available schools, their specializations, addresses, and previous years' admission cut-offs. The government also provides a website where students can download additional information about each school and use a mapping tool to see each school's location. In March, students submit a rank-ordered list (ROL) listing up to 20 schools. In June, all students take a system-wide admission exam. We include a more detailed description of the admission exam in Appendix A.

All schools prioritize students based on the admission exam score. Elite schools exclude from consideration students with a middle school GPA lower than 7 out of 10. However, most of the students meet this requirement. To obtain a middle school certificate, students must have a GPA of at least 6 out of 10. In 2007, 91 percent of students met the GPA requirement for elite school admission (Figure 1).

Before implementing the matching algorithm, schools decide the number of seats to offer. During the matching process, some students may have the same admission exam score and

Figure 1: Elite schools minimum GPA requirement



Note: This figure shows the cumulative distribution function of middle school GPA. The minimum GPA for middle school graduation and participation in the centralized high school admission system is six. To be considered for admission to an elite school, students must have a GPA greater or equal to seven (dashed line).

compete for the last available seats at a given school. In this case, schools either admit or reject all tied students. For example, if a school has ten seats remaining during the matching process, but 20 tied students compete for them, the school must decide between admitting all 20 or rejecting them all.

The matching algorithm is the serial dictatorship. The serial dictatorship algorithm ranks students by the admission exam score and, proceeding in order, matches each applicant to her most preferred school among the schools with available seats. We provide a more detailed explanation of the serial dictatorship algorithm in Appendix B.

Some students may be left unmatched at the end of the matching process. There are two reasons why some students are unmatched. First, some students do not clear the cut-off for any schools they list in their ROLs. Second, some students only apply to elite schools and do not meet the minimum GPA requirement. Unmatched students can register at schools with available seats after the matching process is over.

Table 2: Students' characteristics by assignment group

	All	Elite	Non-Elite	Unmatched
Exam Score	65.24	90.16	60.27	51.20
	(19.21)	(10.87)	(14.88)	(12.80)
GPA	8.03	8.56	7.88	7.89
	(0.84)	(0.81)	(0.80)	(0.72)
Female	0.51	0.45	0.51	0.61
	(0.50)	(0.50)	(0.50)	(0.49)
Age	15.82	15.56	15.90	15.88
	(1.60)	(1.23)	(1.72)	(1.55)
Length of ROL	9.32	9.62	9.53	8.03
	(3.75)	(3.92)	(3.71)	(3.41)
Position assigned	3.32	1.94	3.79	
	(2.94)	(1.72)	(3.11)	
Graduation	0.58	0.71	0.58	0.43
	(0.49)	(0.45)	(0.49)	(0.50)
Observations	256,335	54,654	162,063	39,618

Note: This table shows the characteristics of the middle school students participating in the assignment process. The length of ROL is the number of schools a student includes in her application list. The position assigned is where she ends up assigned in the ranking submitted by a student. Graduation indicates if a student graduated or not within five years. Standard deviations are in parenthesis.

3 Administrative Data

We use individual-level administrative data from the 2007 high school admission process in Mexico City. In that year, 256,335 students applied to 658 high schools. We observe each student's admission exam score, ROL, GPA, assigned school, and socio-demographic characteristics, such as gender and parental income. In Table 2, we include descriptive statistics of the applicant population. Students assigned to elite schools have higher admission exam scores, higher GPAs, and a larger share of them are male.

On the high school side, we have information on the number of seats each school offers, the subsystem to which each school belongs, and previous years' admission cut-offs for each school. With this information, we use the Serial Dictatorship algorithm and fully replicate the assignments we observe in the data (Table 3). Being able to reproduce the student-school matches observed in the data gives us confidence in the transparency of the admission system.

Table 3: Matching outcomes in 2007

		N	%
Matched		216,717	73.02
Unmatched		39,618	13.35
Subtotal		256,335	
Ineligible	< 31 in exam	5,841	1.97
	No exam	6,353	2.14
	No middle school	28,249	9.52
Total		296,778	100

NOTE: This table shows the results of running the serial dictatorship algorithm using the administrative data. A student is ineligible if she obtains a score lower than 31 in the admission exam, does not show up for the exam, or does not obtain a middle school degree.

We define high school graduation as high school completion between 3 to 5 years after participating in the 2007 admission process. Expected high school duration is three years for all high schools. To measure graduation, we combine two sources of data. First, we rely on administrative graduation records to measure whether a student graduates from the assigned subsystem. Second, we use participation in a high school exit exam to measure the graduation of students who were unmatched during the admission process, switched schools across subsystems, or moved to the private sector. Combining these two data sources, we obtain an unconditional measure of high school graduation that splits students into those who complete any high school and those who are high school dropouts. Appendix C includes a more detailed description of how we construct the graduation variable.

In Table 2, we show that the system-wide graduation rate is 58%, and students assigned to elite schools have a thirteen percentage points higher average graduation rate than those assigned to non-elite schools. This difference likely reflects the selection of more skilled students into elite schools. Not all unmatched students become high school dropouts, 43% of them finish high school within the next five years. Unmatched students can still complete high school by re-applying the following year or attending private schools.

In addition to the admission and graduation information, we observe students' scores on a standardized, low-stakes exam that they take during the last semester of middle school. The exam evaluates students in two subjects: mathematics and Spanish. The government designed and implemented this exam for school accountability purposes. We refer to it as the low-stakes exam.

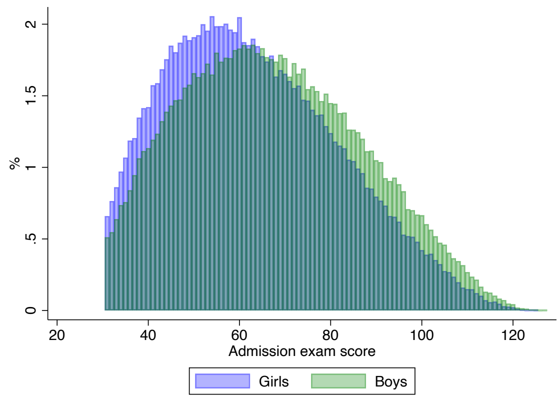
Previous literature shows that females tend to perform worse in standardized tests than males [Niederle and Vesterlund, 2010]. This gap in performance does not mean that females have lower skills than males but that there are gender differences in performance under competitive pressure. In Figure 2, we show some descriptive statistics regarding gender differences in our available skill measures. Panel (a) shows that boys score higher than girls in the admission exam score. In contrast, Panel (b) shows that girls have higher GPAs than boys. Furthermore, Panel (c) shows that girls have higher GPAs than boys at every decile of the admission exam score distribution. In this context, rationing over-subscribed schools seats based only on performance in an admission exam could limit girls' access to them. Further, if GPA is a strong predictor of graduation, then such an admission rule could increase mismatch by restricting the access of high-GPA students to the most academically demanding schools.

4 Regression Discontinuity Evidence

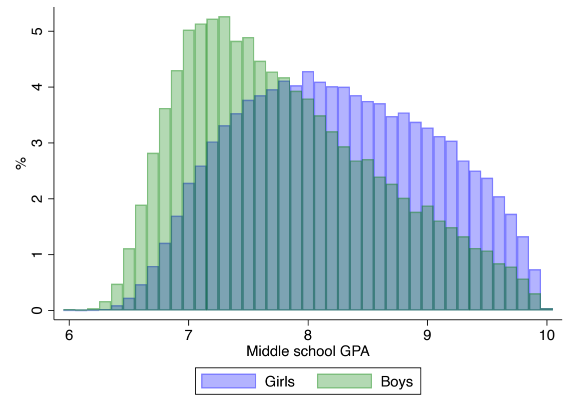
Elite schools are the most over-subscribed schools in the system, and admission to them requires clearing their admission cut-offs. We exploit these cut-offs to identify the effect of marginal admission to an elite school on the probability of graduation. We treat admission as equal to enrollment because enrollment at elite schools is almost universal. The average enrollment rate for students admitted to an elite school is 97.42%.

We follow Dustan et al. [2017] and construct a sample of students who would be assigned to an elite school if they meet the cut-off and assigned to a non-elite school otherwise. Our definitions of elite schools differ because they only consider as elite schools a subset of them that specialize in science education. Notably, the schools in their analysis sample are not the most over-subscribed in the market. We impose three sample restrictions. First, we exclude all ineligible students for admission to an elite school. To be eligible for admission to an elite school, students must have a GPA higher than 7/10 during middle school. Second, we only include students who have applied to at least one elite and non-elite school. Third, we only include students who rank elite schools higher than non-elite ones. The purpose of the last

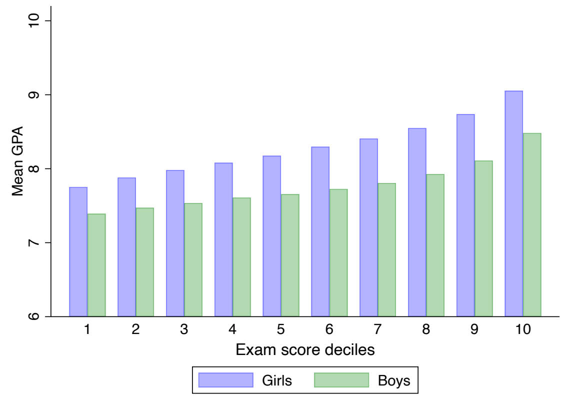
Figure 2: Skill measures by gender



(a) Admission exam score



(b) GPA



(c) GPA and admission exam score

NOTE: Panel (a) in this figure shows the distribution of admission exam scores for girls and boys. Panel (b) in this figure shows the distribution of GPA for girls and boys. Panel (c) in this figure shows the average GPA for girls and boys at each decile of the exam score.

restriction is to select students with similar preferences in that they prefer elite schools to non-elite schools.

Our strategy to estimate the effect of admission to a particular institution follows the same intuition as in Kirkeboen et al. [2016]. In our case, we consider only two institutions, elite and non-elite. In the estimation sample, we have students whose first best is an elite school and whose second best is a non-elite school in the local institution ranking (i.e., same ordinal preferences around their admission score). However, in addition to students having the same preferences in the local institution ranking, we only consider students who prefer elite to non-elite schools in their full ranking. We can impose this last restriction because most students who apply to both types of schools rank elite schools higher than non-elite schools. The previous restriction only excludes 815 (0.76%) students.

In our estimation sample, each student has a minimum cut-off for elite admission, c_k, that depends on her preferences. For example, if a student applied to multiple elite schools, her admission cut-off would be the lowest cut-off of the elite schools she included in her application. There are $k = 30$ groups of students that share the same c_k, corresponding to the cut-offs of the 30 elite schools. Within each group k, the following condition is satisfied:

$$\begin{cases} s_i \geqslant c_k \text{ admitted to some elite school,} \\ s_i < c_k \text{ admitted to some non-elite school,} \end{cases}$$

where s_i indicates student i score in the admission exam.

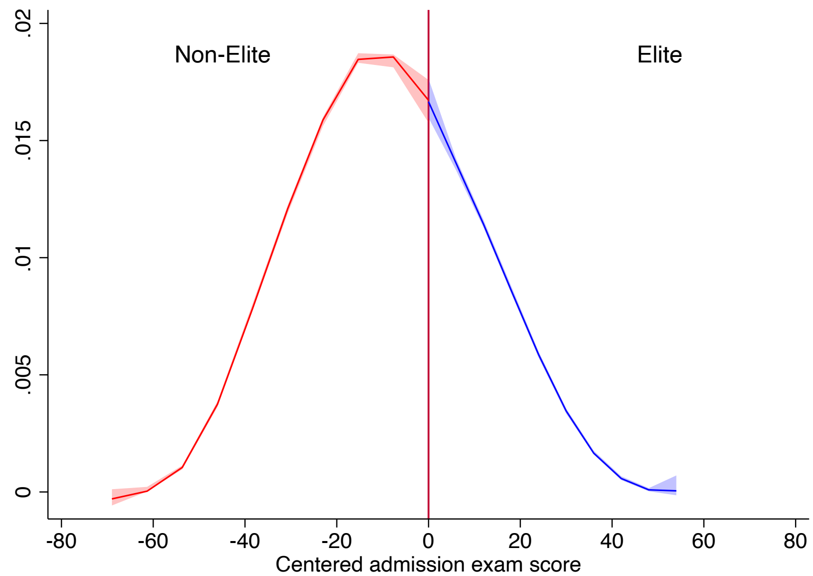
Our empirical specification follows Equation 1, where we stack our previously defined k groups. In this equation, y_{ik} is a dummy variable that denotes whether student i in group k graduates from high school. We center the running variable s_i by the group-specific admission cut-offs c_k such that a positive value of $s_i - c_k$ indicates admission to an elite school. The dummy variable $admit_i$ takes a value of one when a student is admitted to an elite school and zero otherwise.

$$y_{ik} = \mu_k + \gamma admit_i + \delta(s_i - c_k) + \tau(s_i - c_k) \times admit_i + \epsilon_{ik}. \tag{1}$$

Our parameter of interest γ indicates the effect of marginal admission to an elite school on graduation. For estimation, we follow the non-parametric robust estimator proposed by Calonico et al. [2014]. We also follow their method to calculate the mean squared error optimal bandwidth. For robustness, we estimate three additional specifications. First, we add a polynomial of degree two of the running variable. Second, we include k group fixed effects (i.e., cut-off fixed effects). Third, since our running variable only takes integer values, we follow Kolesár and Rothe [2018] approach for estimation and inference with a discrete running variable. All of our estimation results are not affected but the specification changes.

Regarding the validity of the design [Imbens and Lemieux, 2008], we show that there

Figure 3: Running variable



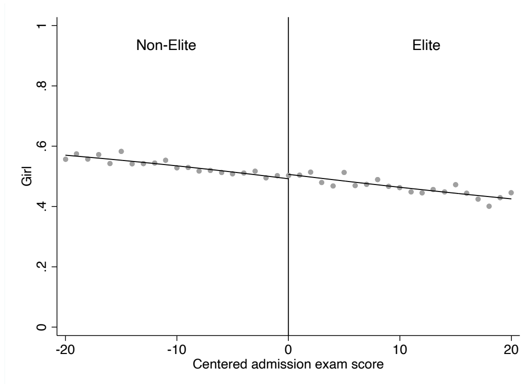
NOTE: This figure shows the density of the centered running variable. The vertical red line indicates the admission threshold.

is no evidence of manipulation of the running variable around the admission cut-offs. If students could manipulate the running variable, they could sort themselves to be above an elite school admission cut-off. This type of sorting is unlikely in our context for two reasons. First, admission cut-offs are determined in equilibrium after students submit their applications and take the admission exam. Second, students do not know their score in the admission exam until the end of the admission process. If there were manipulation, we would expect to observe bunching of the running variable just above the admission cut-offs. Figure 3 shows the density of the running variable. The density does not show any bunching, and a continuity test [McCrary, 2008] does not reject its continuity at the admission cut-offs (p-value=0.958). Our findings are consistent with the absence of manipulation.

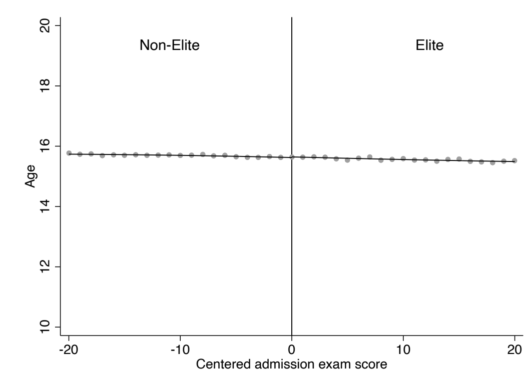
Figure 4 shows that other predetermined covariates such as gender, age, GPA, family income, and number of siblings also do not vary discontinuously at the cut-offs. This is further evidence supporting the validity of the design. The estimates and standard errors are in Appendix D.

Figure 5 shows a graphical representation of the effect of marginal admission to an elite school on graduation. Elite schools decrease the graduation rate of marginally admitted

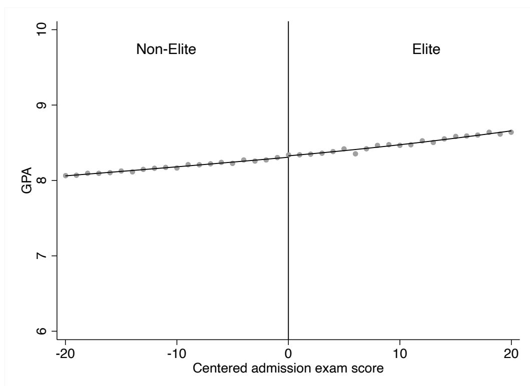
Figure 4: Predetermined covariates



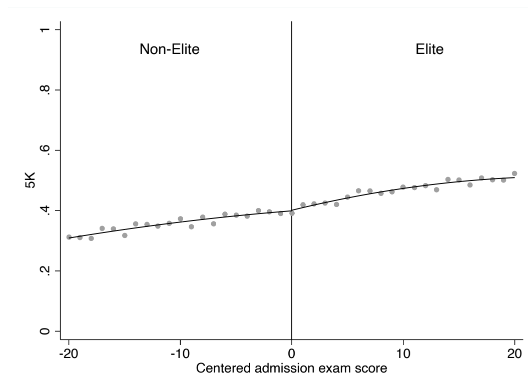
(a) Girl



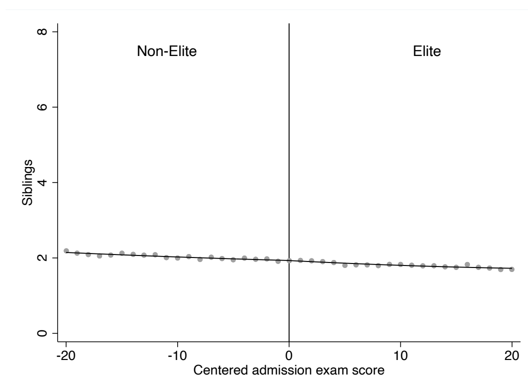
(b) Age



(c) GPA



(d) Income

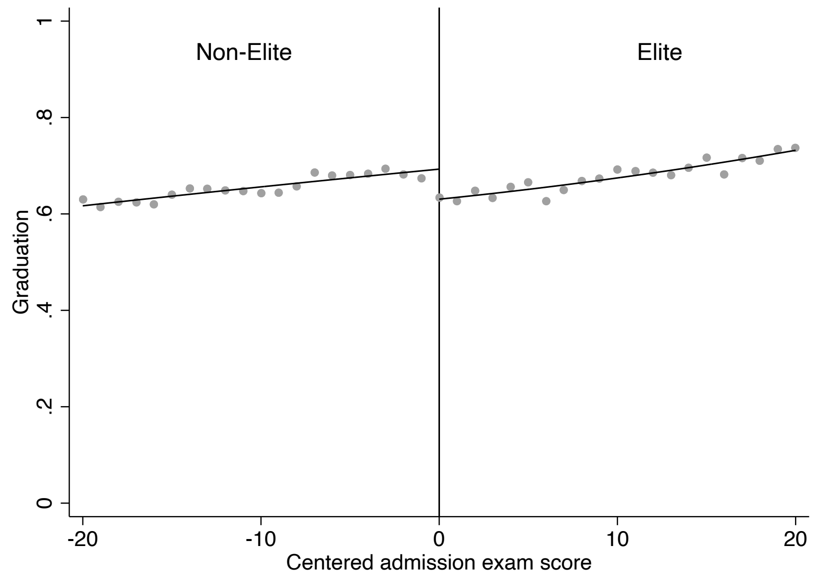


(e) Siblings

NOTE: This figure shows binned means of predetermined covariates around the elite admission thresholds. Income is a dummy variable indicating if the family monthly income is higher or lower that 5000 pesos (458 USD).

students (six percentage points). We show the estimated parameter $\hat{\gamma}$ and its standard error in Appendix G. Elite schools have a more demanding curriculum and better quality peers, and students marginally admitted using a single standardized exam may not be prepared

Figure 5: The effect of elite schools on graduation



NOTE: This figure shows binned means of graduation around the elite admission thresholds.

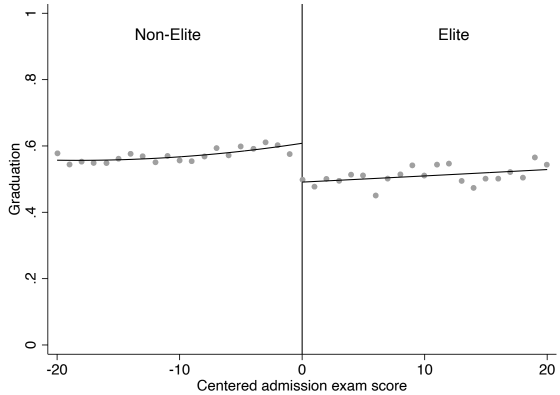
enough for what these schools offer. However, this does not mean that all students marginally admitted experience a negative effect from elite schools. Since the correlation between the admission exam score and middle school GPA is 0.4, some students at the margin have high and low middle school GPAs. In the next section, we explore if the effects are different for these two subgroups of students.

Before we analyze the effect for students with high and low middle school GPAs, we show that the design is also valid for each subgroup. There is no evidence of manipulation of the running variable for our samples of high- and low-GPA students. In addition, the pre-determined covariates are also continuous at the cut-offs. We include these results in Appendixes E and F.

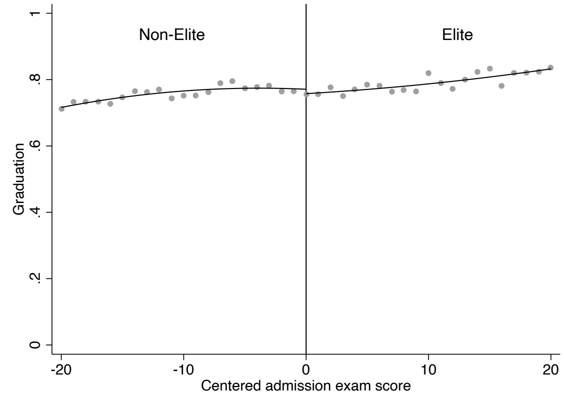
4.1 Heterogeneity by GPA

Students at the elite school admission cut-offs can be heterogeneous in other characteristics that affect graduation. For example, they may have high or low GPAs. Borghans et al. [2016] show that grades and achievement tests capture IQ and personality traits, but grades weigh personality traits more heavily. Since personality traits such as self-control or conscientiousness could matter for graduation when admitted to an elite school, we next explore

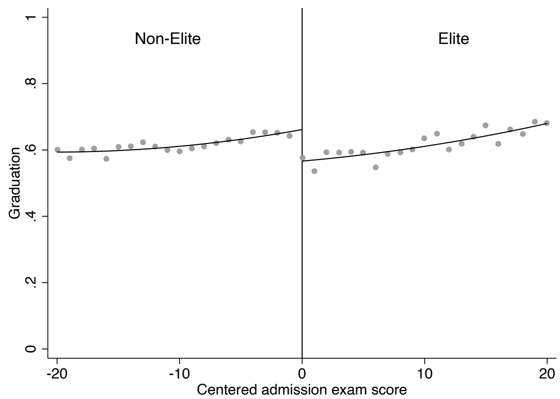
Figure 6: Elite school admission and graduation by GPA and gender



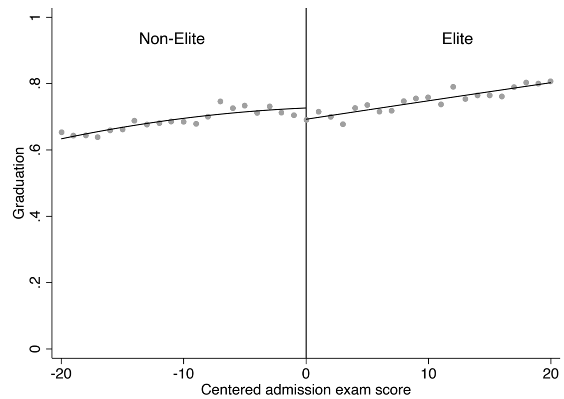
(a) Low middle school GPA



(b) High middle school GPA



(c) Boys



(d) Girls

NOTE: This figure shows binned means of graduation around the elite admission thresholds for boys and girls, and students with high- and low-GPA.

if the effect is different for students with above and below-median GPAs.

In an extreme example, consider the case where the admission exam only captures IQ while GPA only captures self-control. Then, exploring our heterogeneity of interest would be equivalent to differentiating between the effect of elite schools on high-ability, low-self-control students and high-ability, high-self-control students. In this example, to gain admission to an elite school, a student needs to perform well in the admission exam (high-ability), but she need not have high self-control. To the extent that graduation when admitted to an elite school requires you not only to have high ability but also have high self-control, we would expect differentiated effects.

The panels (a) and (b) in Figure 6 shows that the effect of marginal admission to an elite school on graduation is heterogeneous by middle school GPA. It is negative (twelve percentage points) and significant for students with below-median GPA and it does not affect the graduation of students with above-median GPA. We include point estimates and standard errors in Appendix G. We take these results as evidence of both skill measures being important for high school graduation when admitted to an elite school.

Admission to an elite school implies being exposed to a more advanced curriculum and experiencing higher quality peers, among other factors. It could be the case that high GPA students experience a different change in peers than low GPA students and that this is driving our heterogeneous results. We discard this possibility in Appendix H, where we show that the change in peer quality when admitted to an elite school for the full sample, for the high GPA sample and for the low GPA sample is approximately the same. We measure peer quality by the average admission exam score at the admitted school. In other words, better prepared students as measured by GPA respond differently to a similar peer quality shock when marginally admitted to an elite school.

In Appendix I, instead of separating students as having above- or below-median GPAs in the entire distribution of GPAs, we define above- and below-median GPA students relative to the distribution of GPAs within their middle schools. We do this to control for middle school effects and ensure that our results are not driven by attending particular subgroups of middle schools. Our heterogeneous results by GPA are robust to this alternative definition of high and low GPA.

In Appendix J, we include an additional robustness check showing that the heterogeneity by GPA does not depend on elite schools having relatively higher or lower admission cut-offs. We separate elite schools into two groups, high- and low-cut-offs, among our thirty elite school cut-offs. We then show that the negative effect for low GPA students and the null effect for high GPA students is present in both groups of elite schools.

In Appendix K, we show that our heterogeneous results are not just the product of using multiple measures of the same skill (i.e., noise reduction). To do so, instead of GPA, we explore heterogeneity by performance in the low-stakes standardized exam. Our results in K shows negative effects on graduation for both the high and low performers in the low-stakes standardized exam.

In Appendix L, to isolate the skills that GPA measures from those already accounted for by standardized exams, we use the residuals from regressing GPA on the admission exam score and the low-stakes standardized exam to define high and low GPA students. Our results show that our heterogeneous results in Figure 6 remain almost identical. We interpret this as evidence that the additional skills that GPA better captures are driving our heterogeneous results by GPA.

4.2 Heterogeneity by gender

In the last section, we showed that the effect of elite schools on the graduation probability of marginally admitted students depends on their previous GPA. Since in Section 3, we showed that girls have higher GPAs than boys and, arguably, are better prepared for elite schools, we would also expect to observe heterogeneous effects by gender.

The panels (c) and (d) in Figure 6 shows the results of estimating RDDs separately for girls and boys. The effect for boys is almost identical (decrease of ten percentage points) to that for students with below-median GPA. In contrast, the effect for girls replicates the null effect for students with above-median GPA. We include point estimates and standard errors in Appendix G. Our results can partially be explained by girls having higher GPAs than boys throughout the support of the admission exam score (Panel c in Figure 2).

To understand the source of heterogeneity in treatment effects, we follow Gerardino et al. [2017] and use propensity score weighting to keep one characteristic balanced while doing subgroup analysis for the other. In our case, we keep gender balanced while doing heterogeneity by GPA and keep GPA balanced while doing heterogeneity by gender. We show the main results of this exercise in Appendix M. When we hold gender balanced, we still observe heterogeneous results between high and low-GPA students, although the difference in effect sizes is smaller than before. However, when we hold GPA balanced, we no longer observe differences in the effect between girls and boys. We interpret this as evidence that what drives our heterogeneous results are the skills being captured by GPA, and what is behind the gender results is that girls have higher GPAs than boys at the elite admission cut-offs.

Overall, the results of our RDD analysis tell us two facts. First, marginal admission to elite schools only affects the graduation for students without enough of the skills needed to face their higher academic standards and better quality peers. Second, a combination of the admission exam and GPA is better at capturing these skills than the admission exam alone.

5 Counterfactual Priority Orders

Motivated by the RDD results, we examine the effects of counterfactual priority orders that may better match students to schools. We combine the admission exam score (s_i) and GPA (g_i) with different weights (ω) to define new priority orders. Our priority orders follow Equation 2. Since the matching algorithm is the Serial Dictatorship, all schools j give the same priority to student i. Notice that when $\omega = 0$, we are in the baseline case where schools rank students using only their admission exam scores.

$$priority_{ij}^{\omega} = (1 - \omega)s_i + \omega g_i, \tag{2}$$

where $\omega \in [0, 1]$.

We create a grid of weights ω that go from zero to one in 0.1 increments for our counterfactuals. We run the Serial Dictatorship algorithm for each grid point to find the stable equilibrium allocation μ^{ω}. Equation 3 defines f^{SD} as a matching function that has as inputs the priorities, students' preferences U_{ij}, and the available seats. In our counterfactuals, we keep preferences and seats fixed while changing $priority_{ij}^{\omega}$ through changes in ω.

$$\mu^{\omega} = f^{SD}(priority_{ij}^{w}, U_{ij}, seats_j). \tag{3}$$

We use GPA in levels in our counterfactual analysis. However, a potential concern about using levels is that middle schools may react by inflating grades. Thus, we also study a counterfactual that combines the admission exam score with within middle school percentile ranking by GPA. Since within-school rankings are unaffected by grade inflation, such a policy could help prevent this response. As shown in Appendix N, our counterfactual results are not sensitive to the implementation option. Another potential concern could be that students may respond by transferring between middle schools. However, in the Mexican context, middle school mobility is restricted since middle school admissions are also centralized [Fabregas, 2023].

Students could also react by changing their efforts from studying for the admission exam to working on their middle school coursework. Such a behavioral response is not necessarily negative. Suppose students move more of their effort toward coursework and away from studying for the admission exam. In that case, we might expect a larger positive effect on graduation, assuming that studying for middle school coursework is more productive in building knowledge/skills associated with future academic success than studying for the entrance exam. In this case, we would expect our results to be a lower bound for the total effects on graduation. Other potential ways to increase GPA, such as private tutoring, are less likely to occur given that we are considering a measure of overall GPA during three years of middle school.

5.1 Preferences

We observe students' ROLs, but there are some reasons why they may not reflect student preferences. The matching algorithm is the serial dictatorship, which incentivizes truthful revelation of preferences when students can rank all schools in the market [Svensson, 1999]. However, in the Mexican system, there is a constraint on the number of schools students can list (a maximum of 20), which may lead to some students not revealing their preferences in their ROLs [Haeringer and Klijn, 2009; Calsamiglia et al., 2010]. Also, some students may misreport their preferences due to strategic mistakes [Artemov et al., 2023; Hassidim et al., 2017]. Critically, some students may omit infeasible (under the status quo priority order) over-subscribed schools in their ROLs, which may become feasible under alternative priority orders.

We estimate student preferences for schools under the stability of the market equilibrium assumption [Fack et al., 2019]. The stability of the market implies that students are assigned to their preferred ex-post feasible schools. Feasibility is determined by students' admission exam scores and schools' equilibrium cut-offs. We define the indirect utility of student i for school j as follows.

$$U_{ij} = \delta_j + \gamma'_{s(j)} x_i + \psi' x_i k_j + \lambda d_{ij} + \epsilon_{ij}, \tag{4}$$

where δ_j denotes average taste for school j. Each school belongs to a sub-system, denoted by the index $s(j)$. We allow students to have heterogenous tastes for different sub-systems through the vector of parameters $\gamma'_{s(j)}$. Individual heterogeneity is captured by vector x_i, which contains the low-stakes exam score (known at the application stage), middle school GPA, and gender. k_j indicates the selectivity of school j measured by its previous year's admission cut-off. We also allow for heterogeneous tastes for selectivity through parameters ψ'. Parameter λ captures preferences for distance to school j in kilometers. ϵ_{ij} measures the unobservables, which we assume to be i.i.d and come from a type I extreme value distribution.

Under the stability assumption, we define the individual choice sets as follows:

$$F(s_i, K(\mu)), \tag{5}$$

where s_i indicates student i admission exam score and $K(\mu)$ indicates the equilibrium cut-offs associated with matching μ. F denotes the set of feasible schools for student i given equilibrium cut-offs $K(\mu)$.

We estimate preference parameters by MLE using a conditional logit with heterogeneous choice sets. Our model has an outside option that indicates preferences for remaining unmatched. We normalized the mean utility of the outside option to zero. We use our estimated preference parameters and a draw of ϵ_{ij} to simulate preferences in the market. We then use our simulated preferences to approximate stable equilibriums as suggested by Artemov et al. [2023].

In Table 4 we show $\hat{\psi}'$ estimates. Since equilibrium cut-offs depend on demand and supply in the market, the previous year's cut-offs measure how over-subscribed schools are. Parameters $\hat{\psi}'$ capture heterogeneous tastes for selectivity. For ease of interpretation, we

Table 4: WTT

	Estimates
Selectivity× GPA	0.260
	(0.107)
Selectivity× LS exam	-0.005
	(0.126)
Selectivity× Girl	0.950
	(0.196)

Note: This table shows parameters that capture heterogeneous preferences for school selectivity. Selectivity is measured by the previous year's admission cut-offs. Individual heterogeneity considers the low-stakes exam score, GPA, and gender. Standard errors in parenthesis.

Table 5: Model fit

	Elite		Non-Elite	
	Data	Model	Data	Model
Admission exam	90.16	90.01	60.27	60.40
GPA	8.56	8.57	7.88	7.88
Girl	0.45	0.46	0.51	0.51

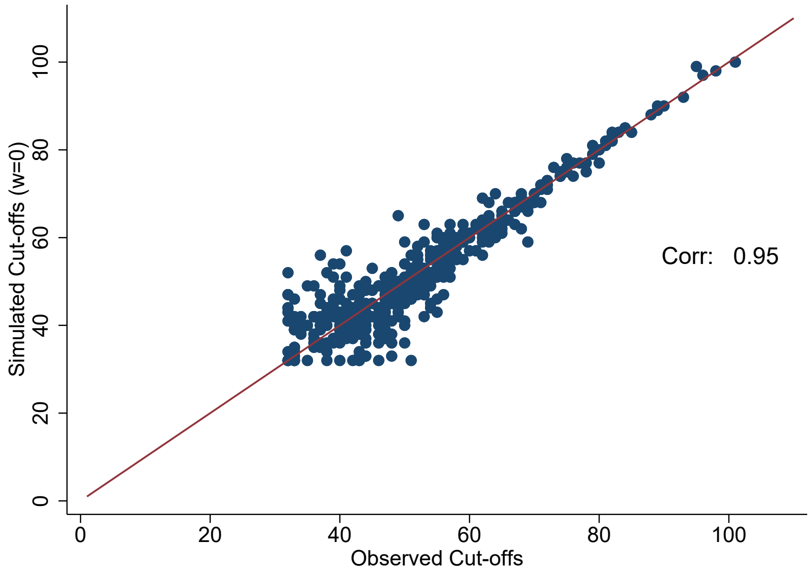
Note: This table shows the average skill measures and the share of girls at elite and non-elite schools. The table compares the observed data and the simulation using estimated preferences and setting $\omega = 0$.

divide our parameters by the distance parameter $\hat{\lambda}$ such that they are measured in willingness to travel (WTT). Students with higher GPAs would be willing to travel 0.26 kilometers farther in order to gain access to more selective schools. Girls would be willing to travel 0.95 kilometers farther to gain access to more selective schools.

To assess the fit of our model, we compare the observed equilibrium cut-offs in the data with the equilibrium cut-offs generated by using simulated preferences and a priority order that only considers the admission exam score ($\omega = 0$). We plot the observed cut-offs against the simulated cut-offs in Figure 7. The model fits the equilibrium cut-offs remarkably well. The correlation between observed and simulated cut-offs is 0.95. Furthermore, in Table 5, we show that we reproduce the average skills measures and gender composition of students allocated to elite and non-elite schools.

We assume that students' preferences do not depend on equilibrium outcomes. Consider

Figure 7: Cut-offs fit



NOTE: This figure shows a comparison between the observed and simulated cut-offs using estimated preferences and a priority order with $\omega = 0$.

the case where students' preferences for schools depend on the average skills of their future peers, and students have rational expectations. Then, the change in priorities could affect the average skills of students assigned to different schools, changing students' preferences for schools. A common assumption in the school choice literature is that preferences do not depend on equilibrium outcomes [Agarwal and Somaini, 2020]. We also work under this assumption.

5.2 Graduation

We define potential outcomes as:

$$Y_{ij} = \alpha_j + \beta_j' x_i + \nu_{ij}, \tag{6}$$

where Y_{ij} indicates the graduation status of student i if matched to school j. α_j measures school j effect and β_j' is a vector that capture match effects between student covariates x_i

and school j. ν_{ij} captures unobservables. In order to quantify treatment effects, we would like to obtain consistent estimates of $\theta_j = (\alpha_j,\, \beta_j')$.

Students are not randomly matched to schools, but sorting depends on two observable student variables: the ROLs and the admission exam score. We take two steps to deal with the non-random sorting of students to schools and obtain consistent estimates of θ_j. First, we control for students' application behavior by including ROLs fixed effects. This is a non-parametric way to control preference heterogeneity, similar to the parametric control function approach in Abdulkadiroğlu et al. [2020]. To reduce the number of unique ROLs, we take advantage of the fact that under the serial dictatorship, a student will never be assigned to a school that is not ranked in cut-off descending order. Therefore, a version of the ROLs where we remove the schools not ranked in cut-off descending order results in the same equilibrium. We control for this version of the ROLs.

Second, we follow Angrist and Rokkanen [2015] method to move beyond the margin of admission in an RDD and control for predictors of graduation other than the admission exam score. The matching algorithm guarantees that students with the same ROLs and admission exam scores are matched to the same schools. Yet, students with the same ROLs, low-stakes exam scores, and middle school GPAs could be matched to different schools due to idiosyncratic factors. We exploit this variation for identification. In other words, we assume that once we control for the ROLs and the skill measures in a vector X_i, the admission exam score becomes ignorable.

We work under the following restriction:

$$E[Y_{ij} \mid X_i, S_i, R_i] = E[Y_{ij} \mid X_i, R_i], \tag{7}$$

where X_i is a vector that includes the low-stakes standardized exam and GPA, S_i is the admission exam score, and R_i denotes ROL. We use the following parametric form:

$$E[Y_{ij} \mid X_i, R_i] = \alpha_{c(j)} + \rho' \bar{x}_j + \beta'_{c(j)} x_i + \Omega_r, \quad \text{for } j \text{ in } \{0, \cdots, J\}. \tag{8}$$

Table 6: Subsystem level match effects ($\hat{\beta}'_{s(j)}$)

	SUB 1	SUB 2	SUB 3	SUB 4	SUB 5	SUB 6	SUB 7	SUB 8	SUB 9
Low-stakes exam×	0.028	0.025	0.044	0.011	0.032	0.024	0.022	0.023	0.032
	(0.004)	(0.007)	(0.025)	(0.024)	(0.007)	(0.005)	(0.008)	(0.007)	(0.048)
GPA×	0.130	0.134	0.103	0.138	0.115	0.123	0.130	0.122	0.090
	(0.004)	(0.007)	(0.024)	(0.018)	(0.005)	(0.005)	(0.007)	(0.006)	(0.051)

NOTE: This table shows subsystem and skill measures interaction effects on graduation (i.e., match effects). The model is the same as in Equation 8, but using subsystem level dummy variables instead of the school level dummy variables. Standard errors in parenthesis.

To reduce the number of parameters to estimate, we group schools j into campuses that we index by $c(j)$. The $J = 658$ schools belong to 311 campuses. The physical location of the schools defines a campus, and each campus is part of a unique subsystem. In our notation, $j = 0$ indicates that a student is unassigned by the matching algorithm. To allow school j peer composition to affect graduation, we include average peer quality at the school level $\bar{x}_j$ in our specification. We denote ROL fixed effects by $\Omega_r = \mathbb{1}[R_i = r]$. Under this specification, our parameters of interest become $\theta_j = (\alpha_{c(j)}, \beta'_{c(j)}, \rho')$.

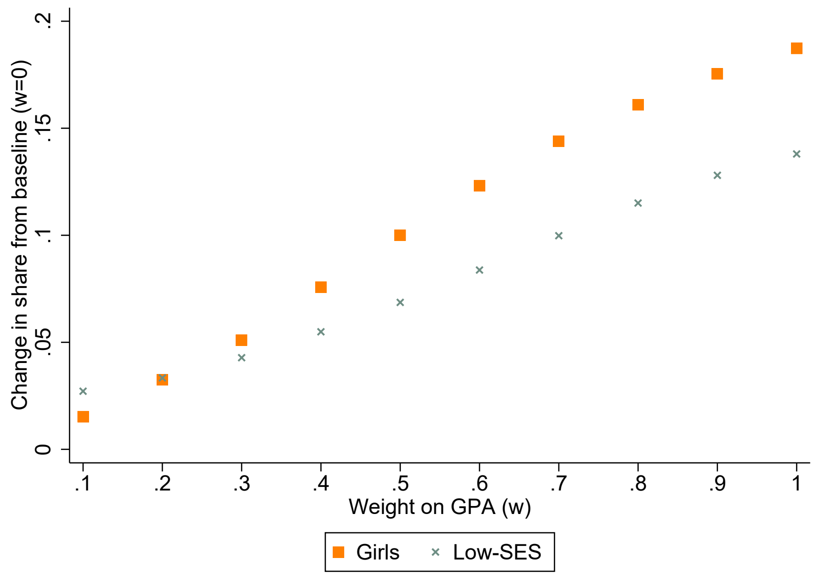
We use our counterfactual simulations to define the subgroups of students affected by the priority order changes. Thus, we combine the characteristics of the students affected in our equilibrium simulations with the parameters $\hat{\theta}_j$ to quantify the treatment effects of our proposed policies.

We use campus-level estimates for the counterfactual analysis, but for ease of exposition, we show selected estimates at the subsystem level in Table 6. The same as in Table 1, SUB 1 and SUB 2 denote the two elite subsystems. We highlight two results from this table. First, standardized preparation and GPA are important determinants of graduation for most of the subsystems. Second, GPA is a stronger predictor of graduation than the low-stakes exam.

5.3 Results

Our counterfactual exercises result in different equilibrium allocations of students across schools. We first analyze the changes in the composition of students allocated to elite schools and then explore the treatment effects of such reallocations for three subgroups of students: the placed, the displaced, and the changed. A placed student moves from a

Figure 8: Changes in the composition of students at elite schools



NOTE: This figure shows the change in the share of girls and low-SES students admitted to elite schools in each counterfactual equilibrium ($\omega > 0$) with respect to the shares in the baseline ($\omega = 0$). The x-axis indicates the weight on GPA (ω). The weight on the admission exam score is $(1 - \omega)$. We define a low-SES student as one whose family income is lower than 5000 Mexican pesos per month (458 USD).

non-elite school (or unassigned) in the baseline to an elite school in the counterfactuals. A displaced student goes from an elite school in the baseline to a non-elite school (or unassigned) in the counterfactuals. A changed student is any student who changes her allocation from the baseline to the counterfactuals.

Figure 8 shows that the higher the weight in GPA, the higher the share of girls assigned to elite schools. This change occurs because girls prefer selective schools (Table 4), but the one-shot exam priority order limits their access. By adding weight to GPA, a measure in which girls outperform boys (Figure 2), more girls gain access to elite schools.

Figure 8 also shows that the higher the weight in GPA, the higher the share of low-SES students assigned to elite schools. Income is highly correlated with the admission exam score but less correlated with GPA. The correlation between income and the admission exam score can partially be explained by high-SES students accessing costly private exam preparation institutions. Adding weight to GPA makes the admission exam score relatively less important and increases low-SES students' access to elite schools.

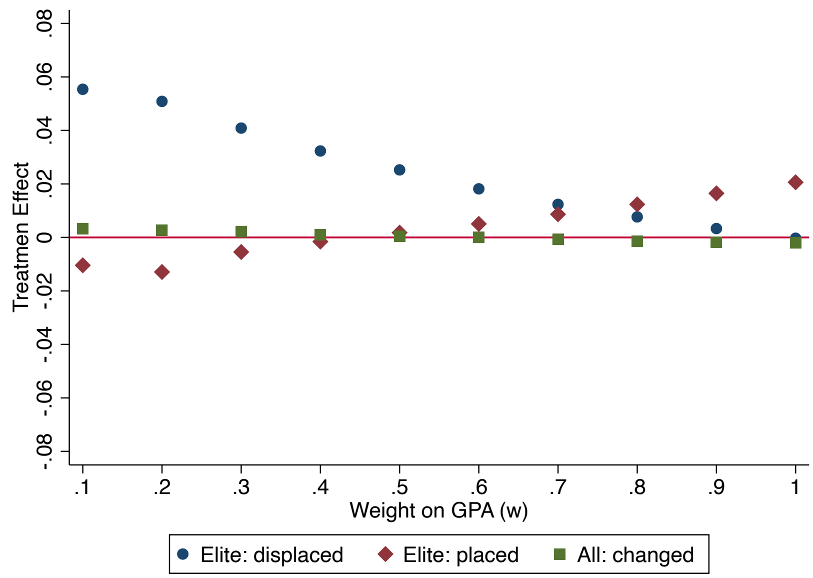
In Figure 9, we show the treatment effects on graduation induced by the changes to the priority order. The average treatment effect on the displaced students is positive for all $\omega > 0$ and decreasing in ω. Displaced students are not prepared enough for elite schools, so when displaced from them, they experience an increase in graduation probability. The treatment effect is decreasing in ω because when the weight on GPA is too high, many of the highly prepared students sort into elite schools, and preparation also matters for non-elite school graduation. The average treatment effect on the placed students is negative for low values of ω and is increasing in ω. Given student preferences in the market, too little weight on GPA still does not sort prepared enough students into elite schools. As we increase ω, more highly prepared students sort into elite schools allowing for them not to be negatively affected by elite admission. In addition, more high GPA students unassigned in the baseline gain access to elite schools, increasing their graduation chances. The average treatment effect on changed students is always close to zero for all values of ω. We take this as evidence that our reallocations do not have (on average) a negative effect on other students in the market through placement and displacement effects.

Overall, our counterfactuals give us insight into the optimal weights in skill measures. For example, consider that policymakers' objectives are to increase the share of girls at elite schools, increase the share of low-SES students at elite schools, and maximize the treatment effects on graduation. In this case, the optimal weights on the admission exam score and GPA are roughly equal. Too much weight on GPA and the positive treatment effect on graduation for the displaced students disappears. Too little weight on GPA and the system does not take advantage of reallocating well prepared students into to elite schools.

6 Conclusions

How a central planner chooses to ration school seats in a centralized education system can affect the equity of access and graduation rates. The relevance of this choice is highlighted when a system priority ordering includes skill measures, and students have diverse latent skills. In this case, using only a one-shot exam as the priority order could match underprepared students with the most academically demanding schools, affecting their graduation

Figure 9: Treatment effects on graduation



NOTE: This figure shows the average treatment effects on graduation for three subgroups of students and different priority orders indexed by ω. The x-axis indicates the weight on GPA ω. The weight on the admission exam score is $(1-\omega)$.

rate. Furthermore, this practice could affect the equity of access when some subgroups of students underperform in one-shot exams while outperforming their peers in other skill measures. Thus, priority orderings play an essential role when centralized education systems evaluations go beyond efficiency measures based on revealed preferences and consider additional policy-relevant outcomes such as equity of access and graduation rates.

We use administrative data from the centralized high school admission system in Mexico City, where all schools share a priority order that relies solely on a one-shot admission exam. We study the effects of adding the information in middle school overall GPA to the priority order. We focus on GPA because previous literature shows that grades measure non-cognitive skills to a higher degree than achievement tests and that non-cognitive skills are a strong predictor of educational success. We first show that boys and low middle school GPA students marginally admitted to the most oversubscribed schools (i.e., elite schools) experience a decrease in their graduation probability whereas high middle school GPA students and girls are unaffected. Our first set of results motivates the importance of using the informational content of grades when considering what skill measures to use in the priority order.

Guided by these results, we then study the effects of counterfactual admission policies where the central planner increasingly adds weight to GPA (or within school ranking by GPA) in the priority order. We have two important findings. First, the higher the weight on GPA, the higher the share of girls and low-SES students admitted to elite schools. Behind this result is that girls have higher GPAs than boys, and family income is less correlated with GPA than the admission exam score. Second, the choice of weights matter for the effects on graduation. Too little weight on GPA negatively affects the graduation rate of students reallocated from non-elite to elite schools, while too much weight on GPA diminishes the gains in graduation of students reallocated from elite to non-elite schools. Both standardized skill measures and non-standardized skill measures are important determinants of graduation. It is not optimal in terms of graduation to only use one skill measure or another as the priority order. Instead, the optimal policy gives roughly equal weight to both skill measures.

A limitation of our study is that our counterfactual admission policies could induce behavioral responses that we are not currently considering.[6] For example, they could affect students' effort allocation between exam preparation and middle school coursework by increasing the effort allocated to coursework. In this paper, we assume that study effort does not change. However, if increased study effort in middle school coursework leads to higher study effort in high school coursework and time spent studying for coursework is more productive than time spent studying for an admission exam, then our effect on the elite schools' graduation rate would be a lower-bound.[7]

From a policy perspective, our results indicate that *combining* the informational content of GPA and the admission exam score in its priority ordering can benefit the centralized system in Mexico City. More broadly, other centralized systems that rely on a one-shot exam score to define school priorities could also benefit from adding some weight to GPA. Examples of such systems are the centralized education systems in Romania, Kenya, Trinidad and Tobago, Ghana, Barbados, and the college admission system in China.

[6]For the case of university admissions, Arenas and Calsamiglia [2022] study a reform in Spain that increased the weight of a standardized exam from 40% to 57% while decreasing the weight on prior GPA. They show that behavioral responses explain 25% of the total effect of the change.

[7]Stinebrickner and Stinebrickner [2006] show that coursework study effort is strongly correlated across time between high school and college.

References

Abdulkadiroğlu, A., Angrist, J., Narita, Y., Pathak, P. and Zarate, R. [2017], 'Regression discontinuity in serial dictatorship: Achievement effects at chicago's exam schools', *American Economic Review, Papers and Proceedings* **107**(5), 240–245.

Abdulkadiroğlu, A., Angrist, J. and Pathak, P. [2014], 'The elite illusion: Achievement effects at boston and new york exam schools', *Econometrica* **82**(1), 137–196.

Abdulkadiroğlu, A., Dur, U. M. and Grigoryan, A. [2021], School assignment by match quality, Technical report, National Bureau of Economic Research.

Abdulkadiroğlu, A., Pathak, P. A., Schellenberg, J. and Walters, C. R. [2020], 'Do parents value school effectiveness?', *American Economic Review* **110**(5), 1502–1539.

Agarwal, N., Hodgson, C. and Somaini, P. [2020], Choices and outcomes in assignment mechanisms: The allocation of deceased donor kidneys, Technical report, National Bureau of Economic Research.

Agarwal, N. and Somaini, P. [2020], 'Revealed preference analysis of school choice models', *Annual Review of Economics* **12**, 471–501.

Angrist, J. D., Pathak, P. A. and Zarate, R. A. [2023], 'Choice and consequence: Assessing mismatch at chicago exam schools', *Journal of Public Economics* **223**, 104892.

Angrist, J. D. and Rokkanen, M. [2015], 'Wanna get away? regression discontinuity estimation of exam school effects away from the cutoff', *Journal of the American Statistical Association* **110**(512), 1331–1344.

Arenas, A. and Calsamiglia, C. [2022], Gender differences in high-stakes performance and college admission policies, Technical report, IZA Discussion Papers.

Arenas, A., Calsamiglia, C. and Loviglio, A. [2021], 'What is at stake without high-stakes exams? students' evaluation and admission to college at the time of covid-19', *Economics of education review* **83**, 102143.

Artemov, G., Che, Y.-K. and He, Y. [2023], 'Stable matching with mistaken agents', *Journal of Political Economy Microeconomics* **1**(2), 270–320.

Azmat, G., Calsamiglia, C. and Iriberri, N. [2016], 'Gender differences in response to big stakes', *Journal of the European Economic Association* **14**(6), 1372–1400.

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

Bleemer, Z. [2021], Top percent policies and the return to postsecondary selectivity, *in* 'Top Percent Policies and the Return to Postsecondary Selectivity: Bleemer, Zachary', [Sl]: SSRN.

Bobba, M., Frisancho, V. and Pariguana, M. [2023], 'Perceived ability and school choices: Experimental evidence and scale-up effects'.

Borghans, L., Golsteyn, B. H., Heckman, J. J. and Humphries, J. E. [2016], 'What grades and achievement tests measure', *Proceedings of the National Academy of Sciences* **113**(47), 13354–13359.

Borghesan, E. [2022], 'The heterogeneous effects of changing sat requirements in admissions: An equilibrium evaluation'.

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

Calsamiglia, C., Haeringer, G. and Klijn, F. [2010], 'Constrained school choice: An experimental study', *American Economic Review* **100**(4), 1860–1874.

Clark, D. [2010], 'Selective schools and academic achievement', *The BE Journal of Economic Analysis & Policy* **10**(1).

Deming, D. J., Hastings, J. S., Kane, T. J. and Staiger, D. O. [2014], 'School choice, school quality, and postsecondary attainment', *American Economic Review* **104**(3), 991–1013.

Dobbie, W. and Fryer Jr, R. G. [2014], 'The impact of attending a school with high-achieving peers: Evidence from the new york city exam schools', *American Economic Journal: Applied Economics* **6**(3), 58–75.

Duckworth, A. L., Quinn, P. D. and Tsukayama, E. [2012], 'What no child left behind leaves behind: The roles of iq and self-control in predicting standardized achievement test scores and report card grades.', *Journal of educational psychology* **104**(2), 439.

Dustan, A., De Janvry, A. and Sadoulet, E. [2017], 'Flourish or fail? the risky reward of elite high school admission in mexico city', *Journal of Human Resources* **52**(3), 756–799.

Fabregas, R. [2023], 'Trade-offs of attending better schools: Achievement, self-perceptions and educational trajectories', *The Economic Journal* p. uead042.

Fack, G., Grenet, J. and He, Y. [2019], 'Beyond truth-telling: Preference estimation with centralized school choice and college admissions', *American Economic Review* **109**(4), 1486–1529.

Gerardino, M. P., Litschig, S. and Pomeranz, D. [2017], 'Can audits backfire? evidence from public procurement in chile', *NBER Working Papers* (23978).

Haeringer, G. and Klijn, F. [2009], 'Constrained school choice', *Journal of Economic theory* **144**(5), 1921–1947.

Hanna, R. N. and Linden, L. L. [2012], 'Discrimination in grading', *American Economic Journal: Economic Policy* **4**(4), 146–168.

Hassidim, A., Marciano, D., Romm, A. and Shorrer, R. I. [2017], 'The mechanism is truthful, why aren't you?', *American Economic Review* **107**(5), 220–224.

Imbens, G. W. and Lemieux, T. [2008], 'Regression discontinuity designs: A guide to practice', *Journal of econometrics* **142**(2), 615–635.

Jackson, C. K. [2010], 'Do students benefit from attending better schools? evidence from rule-based student assignments in trinidad and tobago', *The Economic Journal* **120**(549), 1399–1429.

Jackson, C. K. [2018], 'What do test scores miss? the importance of teacher effects on non–test score outcomes', *Journal of Political Economy* **126**(5), 2072–2107.

Kirkeboen, L. J., Leuven, E. and Mogstad, M. [2016], 'Field of study, earnings, and self-selection', *The Quarterly Journal of Economics* **131**(3), 1057–1111.

Kolesár, M. and Rothe, C. [2018], 'Inference in regression discontinuity designs with a discrete running variable', *American Economic Review* **108**(8), 2277–2304.

Larroucau, T. and Rios, I. [2020], Dynamic college admissions and the determinants of students' college retention, Technical report, Technical Report 2020. and,"Do "Short-List" Students Report Truthfully.

Larroucau, T., Rios, I., Fabre, A. and Neilson, C. [2024], 'College application mistakes and the design of information policies at scale'.

Lavy, V. [2008], 'Do gender stereotypes reduce girls' or boys' human capital outcomes? evidence from a natural experiment', *Journal of public Economics* **92**(10-11), 2083–2105.

Lavy, V. and Sand, E. [2018], 'On the origins of gender gaps in human capital: Short-and long-term consequences of teachers' biases', *Journal of Public Economics* **167**, 263–279.

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

McCrary, J. [2008], 'Manipulation of the running variable in the regression discontinuity design: A density test', *Journal of econometrics* **142**(2), 698–714.

Niederle, M. and Vesterlund, L. [2010], 'Explaining the gender gap in math test scores: The role of competition', *Journal of economic perspectives* **24**(2), 129–44.

Pathak, P. A. [2011], 'The mechanism design approach to student assignment', *Annu. Rev. Econ.* **3**(1), 513–536.

Pop-Eleches, C. and Urquiola, M. [2013], 'Going to a better school: Effects and behavioral responses', *American Economic Review* **103**(4), 1289–1324.

Shi, P. [2022], 'Optimal priority-based allocation mechanisms', *Management Science* **68**(1), 171–188.

Stinebrickner, R. and Stinebrickner, T. R. [2006], 'What can be learned about peer effects using college roommates? evidence from new survey data and students from disadvantaged backgrounds', *Journal of public Economics* **90**(8-9), 1435–1454.

Svensson, L.-G. [1999], 'Strategy-proof allocation of indivisible goods', *Social Choice and Welfare* **16**(4), 557–567.

Appendices

A The admission exam

Table 7: Exam sections

	Questions
Math	12
Physics	12
Chemistry	12
Biology	12
Spanish	12
History	12
Geography	12
Civics and Ethics	12
Verbal ability	16
Math ability	16
Total	128

NOTE: This table shows the number of questions in different subjects that are part of the admission exam.

The admission exam is a multiple choice exam with 128 questions and five choices per question. Each correct answer is worth 1 point, and there are no negative points for wrong answers. Table 7 shows the different sections of the admission exam. The total score is calculated by adding up all the correct answers. Students must obtain a score no lower than 31 points in the admission exam to participate in the assignment process.

B Serial dictatorship

All schools share a unique priority ordering, and each student defines her ROL. Then, the matching algorithm is as follows:

- Step 1: The first ranked student is assigned to the first school on her ROL.

- Step (k+1): For any $k \geq 1$, once the k^{th} student in the priority ranking has been assigned, the student ranked $(k+1)^{th}$ is assigned to the highest-ranked element of her ROL that still has a vacancy. If all of the schools in her ROL are full at that point, she is left unassigned, and the algorithm proceeds to the next student.
- Stop: The algorithm stops after all students have been processed.

Notice that this algorithm is a special case of the Student Proposing Deferred Acceptance algorithm in which all schools share the same ranking of students.

C High school graduation

Table 8: High school graduation

	Step 1	Step 2	Step 3
Unmatched		41.5	43.0
SUB 1	67.0	72.4	72.4
SUB 2	60.6	68.4	68.7
SUB 3	63.2	76.4	76.6
SUB 4	51.1	59.5	60.2
SUB 5	40.2	51.6	52.6
SUB 6	38.5	64.1	64.3
SUB 7	43.0	57.2	57.5
SUB 8	37.1	52.0	52.1
SUB 9	48.8	62.2	62.2
Total	41.6	57.7	58.2

NOTE: This table shows graduation rates at each construction step.

We construct our graduation outcome variable following three steps. In the first step, we collect each subsystem's administrative graduation records 3-5 years after admission. Notice that this measures graduation from the assigned subsystem. We obtained this type of graduation records for all subsystems except SUB 6, for which we could only obtain records for half the admitted students. In the second step, we match the students with an exit exam students take at the end of high school. This allows us to improve our graduation measure by capturing the graduation of students who switched schools across subsystems, reapply in the next admission cycle, or enrolled in private schools. The exit exam helps us determine the graduation status of all matched and unmatched students except those admitted to SUB 1 because this subsystem does not participate in the exit exam. In the third step, we search for the students not matched to SUB 1 during the 2007 admission cycle in the administrative admission and graduation records of SUB 1 for the following admission cycle. The purpose of the last step is to capture the graduation status of students who reapplied after being rejected by SUB 1 in 2007 and were admitted to it during the next admission cycle.

In other words, each step improves our graduation measure compared to the previous one. The graduation variable we use for analysis is the one we obtained after the third step.

D Predetermined covariates

Table 9: Female

	CCT	CCT P2	CCT FE	KR
RD Estimate	0.011	0.004	0.010	0.011
	(0.012)	(0.015)	(0.012)	(0.010)

NOTE: Standard errors in parenthesis. The first three columns show RD estimates following the methods in Calonico et al. [2014] and the associated software package *rdrobust*. The first column shows the estimates of a local linear regression. The second column uses a polynomial of degree two for the running variable. The third column uses local linear regression with cut-off fixed effects. The fourth column shows RD estimates following the methods in Kolesár and Rothe [2018] and the associated software package *rdhonest*.

Table 10: Age

	CCT	CCT P2	CCT FE	KR
RD Estimate	0.022	0.044	0.030	0.010
	(0.027)	(0.036)	(0.029)	(0.043)

NOTE: Standard errors in parenthesis. The first three columns show RD estimates following the methods in Calonico et al. [2014] and the associated software package *rdrobust*. The first column shows the estimates of a local linear regression. The second column uses a polynomial of degree two for the running variable. The third column uses local linear regression with cut-off fixed effects. The fourth column shows RD estimates following the methods in Kolesár and Rothe [2018] and the associated software package *rdhonest*.

Table 11: GPA

	CCT	CCT P2	CCT FE	KR
RD Estimate	0.026	0.028	0.024	0.021
	(0.019)	(0.021)	(0.017)	(0.023)

NOTE: Standard errors in parenthesis. The first three columns show RD estimates following the methods in Calonico et al. [2014] and the associated software package *rdrobust*. The first column shows the estimates of a local linear regression. The second column uses a polynomial of degree two for the running variable. The third column uses local linear regression with cut-off fixed effects. The fourth column shows RD estimates following the methods in Kolesár and Rothe [2018] and the associated software package *rdhonest*.

Table 12: Income

	CCT	CCT P2	CCT FE	KR
RD Estimate	-0.003	-0.004	-0.002	-0.001
	(0.014)	(0.015)	(0.014)	(0.014)

NOTE: Standard errors in parenthesis. Income is a dummy variable indicating if the family monthly income is higher or lower that 5000 pesos (458 USD). The first three columns show RD estimates following the methods in Calonico et al. [2014] and the associated software package *rdrobust*. The first column shows the estimates of a local linear regression. The second column uses a polynomial of degree two for the running variable. The third column uses local linear regression with cut-off fixed effects. The fourth column shows RD estimates following the methods in Kolesár and Rothe [2018] and the associated software package *rdhonest*.

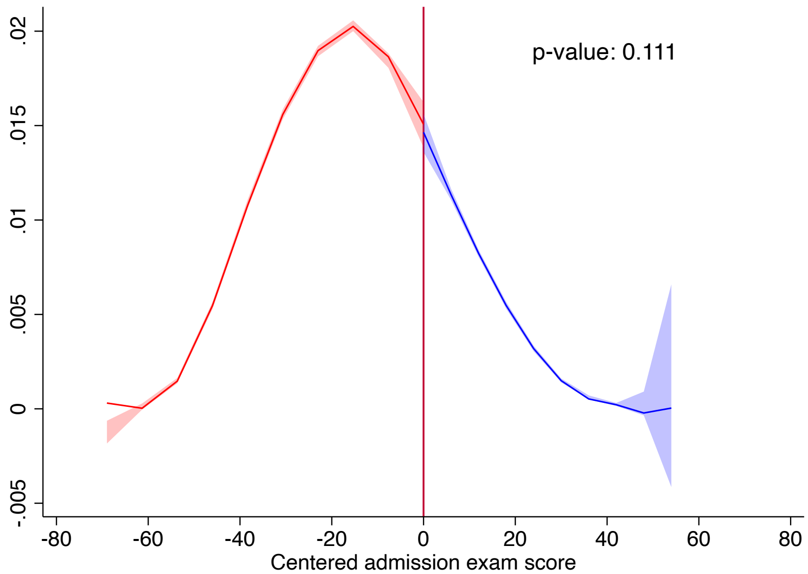
Table 13: Siblings

	CCT	CCT P2	CCT FE	KR
RD Estimate	0.013	0.019	0.010	0.025
	(0.037)	(0.043)	(0.037)	(0.037)

NOTE: Standard errors in parenthesis. The first three columns show RD estimates following the methods in Calonico et al. [2014] and the associated software package *rdrobust*. The first column shows the estimates of a local linear regression. The second column uses a polynomial of degree two for the running variable. The third column uses local linear regression with cut-off fixed effects. The fourth column shows RD estimates following the methods in Kolesár and Rothe [2018] and the associated software package *rdhonest*.

E RDD validity: low GPA

Figure 10: Density of the running variable, low GPA



Note: This figure shows the density of the centered running variable for low GPA students. The shaded regions are 95% confidence intervals.

Table 14: Female

	CCT	CCT P2	CCT FE	KR
RD Estimate	0.002	-0.008	0.005	-0.001
	(0.016)	(0.021)	(0.015)	(0.017)

Note: Standard errors in parenthesis. The first three columns show RD estimates following the methods in Calonico et al. [2014] and the associated software package *rdrobust*. The first column shows the estimates of a local linear regression. The second column uses a polynomial of degree two for the running variable. The third column uses local linear regression with cut-off fixed effects. The fourth column shows RD estimates following the methods in Kolesár and Rothe [2018] and the associated software package *rdhonest*.

Table 15: Age

	CCT	CCT P2	CCT FE	KR
RD Estimate	0.060	0.096	0.061	0.064
	(0.050)	(0.068)	(0.050)	(0.048)

NOTE: Standard errors in parenthesis. The first three columns show RD estimates following the methods in Calonico et al. [2014] and the associated software package *rdrobust*. The first column shows the estimates of a local linear regression. The second column uses a polynomial of degree two for the running variable. The third column uses local linear regression with cut-off fixed effects. The fourth column shows RD estimates following the methods in Kolesár and Rothe [2018] and the associated software package *rdhonest*.

Table 16: GPA

	CCT	CCT P2	CCT FE	KR
RD Estimate	0.010	0.013	0.010	0.012
	(0.014)	(0.016)	(0.014)	(0.014)

NOTE: Standard errors in parenthesis. The first three columns show RD estimates following the methods in Calonico et al. [2014] and the associated software package *rdrobust*. The first column shows the estimates of a local linear regression. The second column uses a polynomial of degree two for the running variable. The third column uses local linear regression with cut-off fixed effects. The fourth column shows RD estimates following the methods in Kolesár and Rothe [2018] and the associated software package *rdhonest*.

Table 17: Income

	CCT	CCT P2	CCT FE	KR
RD Estimate	-0.001	0.001	-0.000	0.010
	(0.021)	(0.027)	(0.021)	(0.022)

NOTE: Standard errors in parenthesis. Income is a dummy variable indicating if the family monthly income is higher or lower that 5000 pesos (458 USD). The first three columns show RD estimates following the methods in Calonico et al. [2014] and the associated software package *rdrobust*. The first column show the estimates of a local linear regression. The second column uses a polynomial of degree two for the running variable. The third column uses local linear regression with cut-off fixed effects. The fourth column shows RD estimates following the methods in Kolesár and Rothe [2018] and the associated software package *rdhonest*.

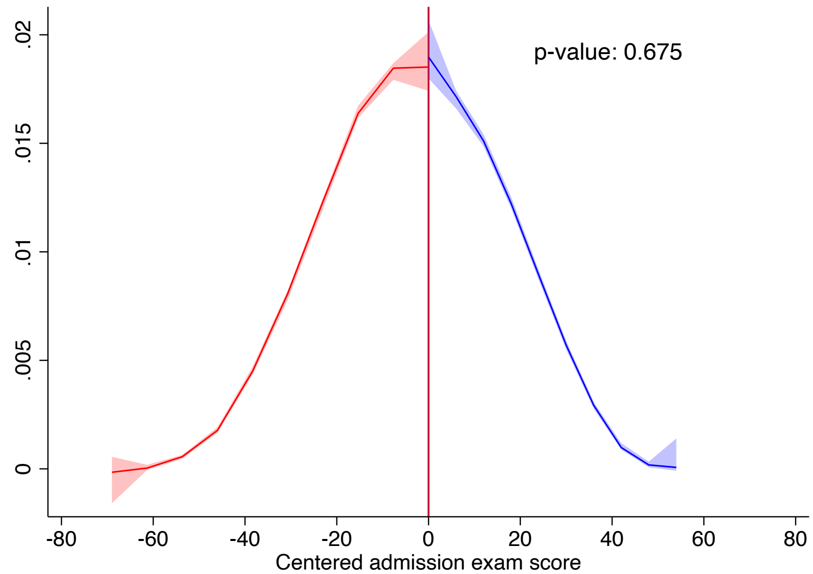
Table 18: Siblings

	CCT	CCT P2	CCT FE	KR
RD Estimate	0.032	0.056	0.026	0.104
	(0.056)	(0.069)	(0.055)	(0.071)

NOTE: Standard errors in parenthesis. The first three columns show RD estimates following the methods in Calonico et al. [2014] and the associated software package *rdrobust*. The first column shows the estimates of a local linear regression. The second column uses a polynomial of degree two for the running variable. The third column uses local linear regression with cut-off fixed effects. The fourth column shows RD estimates following the methods in Kolesár and Rothe [2018] and the associated software package *rdhonest*.

F RDD validity: high GPA

Figure 11: Density of the running variable, high GPA



NOTE: This figure shows the density of the centered running variable for high GPA students. The shaded regions are 95% confidence intervals.

Table 19: Female

	CCT	CCT P2	CCT FE	KR
RD Estimate	0.014	0.008	0.012	0.006
	(0.016)	(0.021)	(0.016)	(0.021)

NOTE: Standard errors in parenthesis. The first three columns show RD estimates following the methods in Calonico et al. [2014] and the associated software package *rdrobust*. The first column shows the estimates of a local linear regression. The second column uses a polynomial of degree two for the running variable. The third column uses local linear regression with cut-off fixed effects. The fourth column shows RD estimates following the methods in Kolesár and Rothe [2018] and the associated software package *rdhonest*.

Table 20: Age

	CCT	CCT P2	CCT FE	KR
RD Estimate	0.010	0.017	0.009	-0.009
	(0.031)	(0.042)	(0.031)	(0.045)

NOTE: Standard errors in parenthesis. The first three columns show RD estimates following the methods in Calonico et al. [2014] and the associated software package *rdrobust*. The first column shows the estimates of a local linear regression. The second column uses a polynomial of degree two for the running variable. The third column uses local linear regression with cut-off fixed effects. The fourth column shows RD estimates following the methods in Kolesár and Rothe [2018] and the associated software package *rdhonest*.

Table 21: GPA

	CCT	CCT P2	CCT FE	KR
RD Estimate	0.011	0.015	0.010	0.013
	(0.015)	(0.016)	(0.014)	(0.018)

NOTE: Standard errors in parenthesis. The first three columns show RD estimates following the methods in Calonico et al. [2014] and the associated software package *rdrobust*. The first column shows the estimates of a local linear regression. The second column uses a polynomial of degree two for the running variable. The third column uses local linear regression with cut-off fixed effects. The fourth column shows RD estimates following the methods in Kolesár and Rothe [2018] and the associated software package *rdhonest*.

Table 22: Income

	CCT	CCT P2	CCT FE	KR
RD Estimate	-0.006	-0.005	-0.007	-0.004
	(0.018)	(0.022)	(0.018)	(0.021)

NOTE: Standard errors in parenthesis. Income is a dummy variable indicating if the family monthly income is higher or lower that 5000 pesos (458 USD). The first three columns show RD estimates following the methods in Calonico et al. [2014] and the associated software package *rdrobust*. The first column shows the estimates of a local linear regression. The second column uses a polynomial of degree two for the running variable. The third column uses local linear regression with cut-off fixed effects. The fourth column shows RD estimates following the methods in Kolesár and Rothe [2018] and the associated software package *rdhonest*.

Table 23: Siblings

	CCT	CCT P2	CCT FE	KR
RD Estimate	-0.008	-0.005	-0.009	0.007
	(0.041)	(0.060)	(0.042)	(0.045)

NOTE: Standard errors in parenthesis. The first three columns show RD estimates following the methods in Calonico et al. [2014] and the associated software package *rdrobust*. The first column shows the estimates of a local linear regression. The second column uses a polynomial of degree two for the running variable. The third column uses local linear regression with cut-off fixed effects. The fourth column shows RD estimates following the methods in Kolesár and Rothe [2018] and the associated software package *rdhonest*.

G Main Estimates

Table 24: Graduation

	CCT	CCT P2	CCT FE	KR
RD Estimate	-0.060	-0.048	-0.061	-0.045
	(0.011)	(0.015)	(0.011)	(0.015)

NOTE: Standard errors in parenthesis. The first three columns show RD estimates following the methods in Calonico et al. [2014] and the associated software package *rdrobust*. The first column shows the estimates of a local linear regression. The second column uses a polynomial of degree two for the running variable. The third column uses local linear regression with cut-off fixed effects. The fourth column shows RD estimates following the methods in Kolesár and Rothe [2018] and the associated software package *rdhonest*.

Table 25: Graduation by GPA

	CCT	CCT P2	CCT FE	AK
High GPA	-0.011	-0.001	-0.009	-0.008
	(0.014)	(0.019)	(0.015)	(0.014)
Low GPA	-0.117	-0.115	-0.118	-0.091
	(0.016)	(0.020)	(0.017)	(0.022)
Difference	0.106	0.114	0.109	0.083
	(0.022)	(0.027)	(0.022)	(0.026)

NOTE: Standard errors in parenthesis. The first three columns show RD estimates following the methods in Calonico et al. [2014] and the associated software package *rdrobust*. The first column shows the estimates of a local linear regression. The second column uses a polynomial of degree two for the running variable. The third column uses local linear regression with cut-off fixed effects. The fourth column shows RD estimates following the methods in Kolesár and Rothe [2018] and the associated software package *rdhonest*.

Table 26: Graduation by gender

	CCT	CCT P2	CCT FE	AK
Girls	-0.021	0.008	-0.020	-0.015
	(0.016)	(0.022)	(0.016)	(0.016)
Boys	-0.098	-0.097	-0.097	-0.083
	(0.017)	(0.019)	(0.017)	(0.022)
Difference	0.077	0.105	0.076	0.068
	(0.024)	(0.029)	(0.024)	(0.027)

NOTE: Standard errors in parenthesis. The first three columns show RD estimates following the methods in Calonico et al. [2014] and the associated software package *rdrobust*. The first column shows the estimates of a local linear regression. The second column uses a polynomial of degree two for the running variable. The third column uses local linear regression with cut-off fixed effects. The fourth column shows RD estimates following the methods in Kolesár and Rothe [2018] and the associated software package *rdhonest*.

H Change in peer quality

We measure peer quality by the average admission exam score of the students admitted to a given school. The admission exam score takes integer values between 31 and 128.

Table 27: Change in peers

	CCT	CCT P2	CCT FE	KR
RD Estimate	19.216	19.187	19.167	18.863
	(0.266)	(0.235)	(0.250)	(0.413)

NOTE: The outcome for all columns is peer quality measured by the average admission exam score at the assigned school. The first three columns show RD estimates following the methods in Calonico et al. [2014] and the associated software package *rdrobust*. The first column shows the estimates of a local linear regression. The second column uses a polynomial of degree two for the running variable. The third column uses local linear regression with cut-off fixed effects. The fourth column shows RD estimates following the methods in Kolesár and Rothe [2018] and the associated software package *rdhonest*.

Table 28: Change in peers by GPA

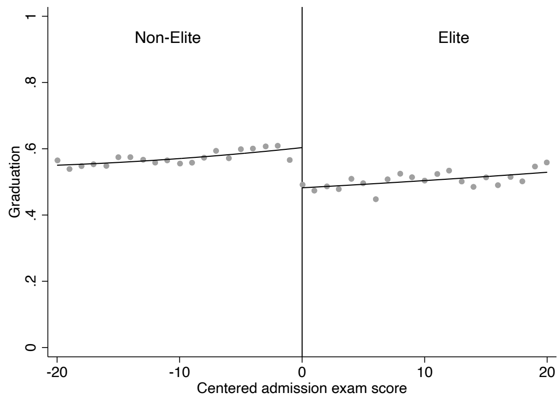
	CCT	CCT P2	CCT FE	KR
High GPA	19.363	19.344	19.217	19.079
	(0.348)	(0.313)	(0.319)	(0.357)
Low GPA	19.207	19.033	19.276	19.031
	(0.324)	(0.323)	(0.327)	(0.361)
Difference	0.155	0.311	-0.059	0.048
	(0.476)	(0.450)	(0.457)	(0.507)

NOTE: The outcome for all columns is peer quality measured by the average admission exam score at the assigned school. The first three columns show RD estimates following the methods in Calonico et al. [2014] and the associated software package *rdrobust*. The first column shows the estimates of a local linear regression. The second column uses a polynomial of degree two for the running variable. The third column uses local linear regression with cut-off fixed effects. The fourth column shows RD estimates following the methods in Kolesár and Rothe [2018] and the associated software package *rdhonest*.

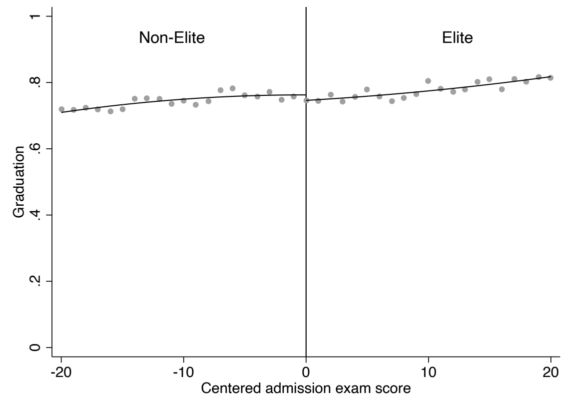
I Above and below-median relative to within middle school GPA distribution

Instead of separating students as having above or below median GPAs in the entire distribution of GPAs, we define above and below median GPA students relative to the distribution of GPAs within their middle schools. We do this to control for middle school fixed-effects and ensure that our results are not driven by attending particular subgroups of middle schools. In Figure 12, we show that our previous results are unchanged by this alternative definition of high and low GPA students.

Figure 12: Graduation by GPA ranking



(a) Graduation: low GPA ranking



(b) Graduation: high GPA ranking

NOTE: This figure shows binned means of graduation around the elite admission thresholds for students above and below median within middle school percentile ranking by GPA.

Table 29: Graduation by GPA ranking

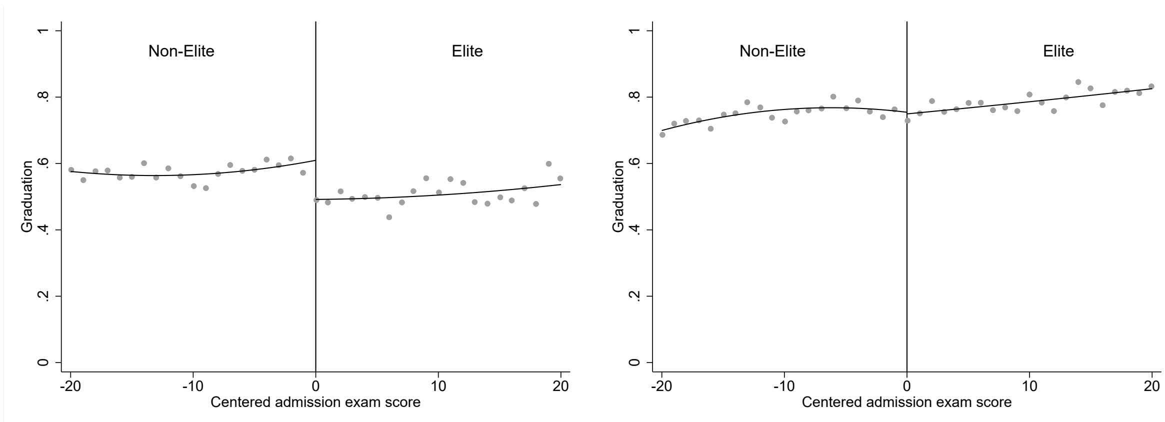
	CCT	CCT P2	CCT FE	AK
High Resid	-0.013	-0.003	-0.013	-0.005
	(0.013)	(0.018)	(0.013)	(0.015)
Low Resid	-0.122	-0.117	-0.122	-0.094
	(0.016)	(0.021)	(0.016)	(0.022)
Difference	0.109	0.114	0.109	0.089
	(0.021)	(0.028)	(0.021)	(0.027)

NOTE: Standard errors in parenthesis. The first three columns show RD estimates following the methods in Calonico et al. [2014] and the associated software package *rdrobust*. The first column shows the estimates of a local linear regression. The second column uses a polynomial of degree two for the running variable. The third column uses local linear regression with cut-off fixed effects. The fourth column shows RD estimates following the methods in Kolesár and Rothe [2018] and the associated software package *rdhonest*.

J Elite schools with high and low cut-offs

For the RDD analysis we pool k groups of students that share a common elite school cut-off c_k. In this Appendix we show that the effects on graduation do not depend on elite schools having high or low cut-offs. Instead of pooling together our k groups, we separate these groups into low and high elite school cut-offs and repeat the analysis for each sub-sample.

Figure 13: Elite school admission and graduation: low elite cut-offs



(a) Graduation: low GPA (b) Graduation: high GPA

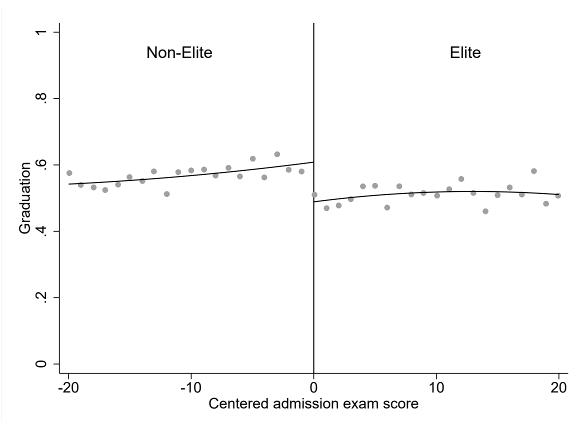
NOTE: This figure shows binned means of graduation around low elite school admission thresholds.

Table 30: Graduation, low elite cut-offs

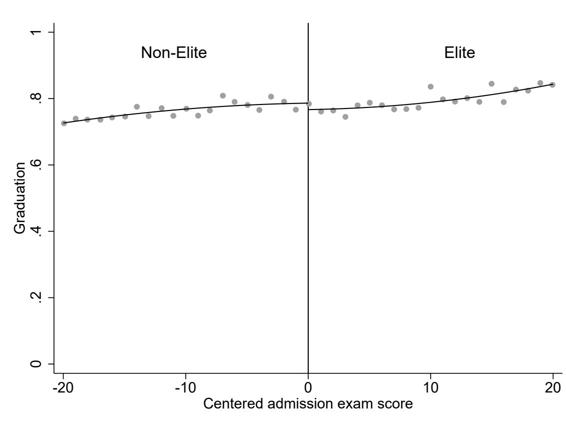
	CCT	CCT P2	CCT FE	AK
High GPA	-0.000	-0.003	0.001	-0.007
	(0.024)	(0.027)	(0.024)	(0.026)
Low GPA	-0.124	-0.112	-0.124	-0.094
	(0.023)	(0.029)	(0.024)	(0.026)
Difference	0.124	0.109	0.125	0.087
	(0.034)	(0.040)	(0.034)	(0.037)

NOTE: Standard errors in parenthesis. The first three columns show RD estimates following the methods in Calonico et al. [2014] and the associated software package *rdrobust*. The first column shows the estimates of a local linear regression. The second column uses a polynomial of degree two for the running variable. The third column uses local linear regression with cut-off fixed effects. The fourth column shows RD estimates following the methods in Kolesár and Rothe [2018] and the associated software package *rdhonest*.

Figure 14: Elite school admission and graduation on time: high elite cut-offs



(a) Graduation: low GPA



(b) Graduation: high GPA

NOTE: This figure shows binned means of graduation around high elite school admission thresholds.

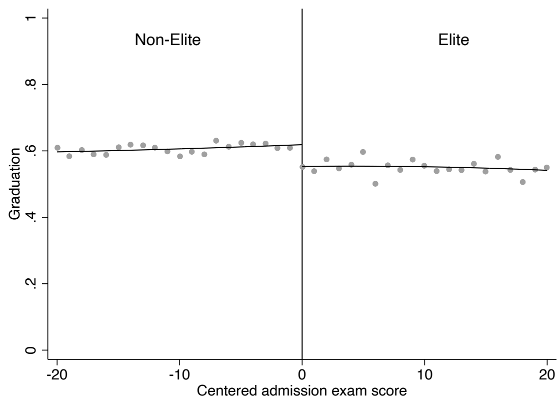
Table 31: Graduation, high elite cut-offs

	CCT	CCT P2	CCT FE	AK
High GPA	-0.010	0.001	-0.011	-0.004
	(0.021)	(0.027)	(0.021)	(0.023)
Low GPA	-0.103	-0.092	-0.102	-0.087
	(0.029)	(0.036)	(0.029)	(0.036)
Difference	0.092	0.093	0.091	0.083
	(0.036)	(0.045)	(0.036)	(0.042)

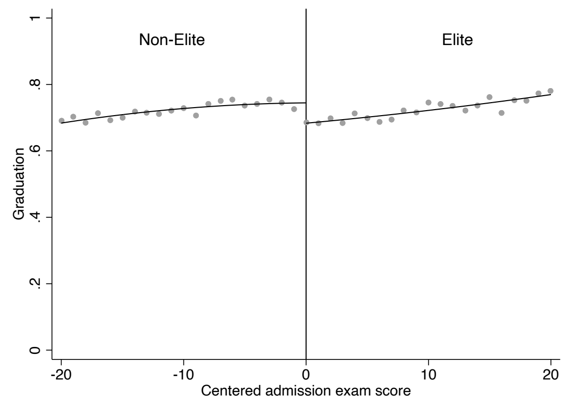
NOTE: Standard errors in parenthesis. The first three columns show RD estimates following the methods in Calonico et al. [2014] and the associated software package *rdrobust*. The first column shows the estimates of a local linear regression. The second column uses a polynomial of degree two for the running variable. The third column uses local linear regression with cut-off fixed effects. The fourth column shows RD estimates following the methods in Kolesár and Rothe [2018] and the associated software package *rdhonest*.

K RDD by low-stakes exam score

Figure 15: Elite school admission and on-time graduation by low-stakes exam



(a) Graduation: low low-stakes exam score



(b) Graduation: high low-stakes exam score

Note: This figure shows binned means of graduation around the elite admission thresholds for students with high and low scores in the low-stakes standardized exam.

Table 32: Graduation by low-stakes exam

	CCT	CCT P2	CCT FE	AK
High LS	-0.043	-0.033	-0.041	-0.045
	(0.017)	(0.020)	(0.017)	(0.017)
Low LS	-0.064	-0.060	-0.065	-0.057
	(0.017)	(0.025)	(0.017)	(0.023)
Difference	0.021	0.027	0.023	0.012
	(0.024)	(0.032)	(0.024)	(0.029)

NOTE: Standard errors in parenthesis. The first three columns show RD estimates following the methods in Calonico et al. [2014] and the associated software package *rdrobust*. The first column shows the estimates of a local linear regression. The second column uses a polynomial of degree two for the running variable. The third column uses local linear regression with cut-off fixed effects. The fourth column shows RD estimates following the methods in Kolesár and Rothe [2018] and the associated software package *rdhonest*.

L RDD by residuals

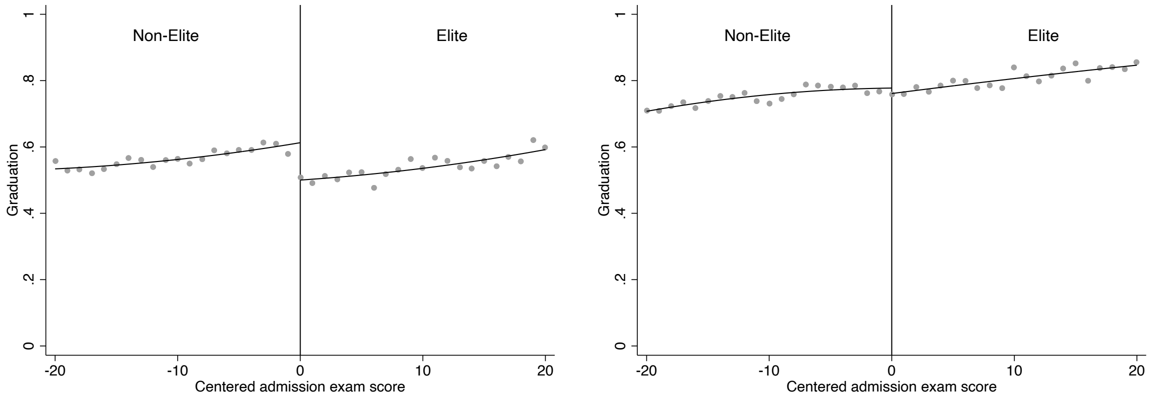
Define:

$$GPA_i = \alpha_0 + \alpha_1 s_i + \alpha_2 ls_i + \epsilon_i, \tag{9}$$

where s_i is the score in the admission exam score, ls_i is the score in the low-stakes exam, and GPA_i is middle school GPA.

We estimate equation 9 and use $\hat{\epsilon}_i$ to define high and low residuals (above and below median).

Figure 16: Graduation by GPA residuals



(a) Graduation: low $\hat{\epsilon}_i$

(b) Graduation: high $\hat{\epsilon}_i$

NOTE: This figure shows binned means of graduation around the elite admission thresholds for students with high and low GPA residuals.

Table 33: Graduation by GPA residuals

	CCT	CCT P2	CCT FE	AK
High Resid	-0.017	-0.005	-0.016	-0.004
	(0.013)	(0.019)	(0.013)	(0.017)
Low Resid	-0.111	-0.109	-0.111	-0.089
	(0.016)	(0.019)	(0.016)	(0.021)
Difference	0.094	0.104	0.095	0.085
	(0.021)	(0.026)	(0.021)	(0.027)

NOTE: Standard errors in parenthesis. The first three columns show RD estimates following the methods in Calonico et al. [2014] and the associated software package *rdrobust*. The first column shows the estimates of a local linear regression. The second column uses a polynomial of degree two for the running variable. The third column uses local linear regression with cut-off fixed effects. The fourth column shows RD estimates following the methods in Kolesár and Rothe [2018] and the associated software package *rdhonest*.

M GPA and gender

To compare two subgroups while holding another observable characteristic constant, we follow the approach proposed by Gerardino et al. [2017]. For example, in our case, the subgroup of students with high GPAs has a higher share of girls than those with low GPAs. Thus, the method allows us to reweight the observations to keep gender balanced across subgroups while studying heterogeneous effects between high- and low-GPA students.

Table 34: RDD estimates using propensity score weighting

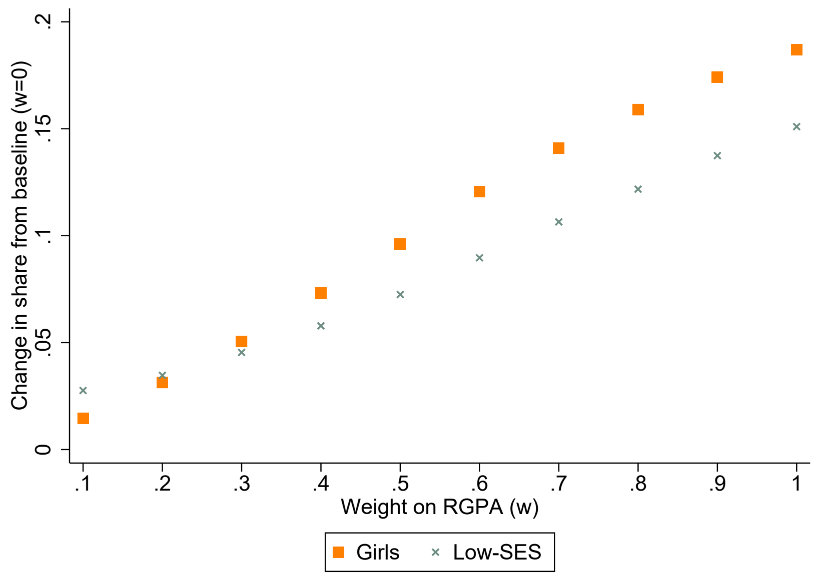
	Gender balanced	GPA balanced
Low GPA	-0.095	
	(0.017)	
High GPA	-0.019	
	(0.012)	
Boys		-0.032
		(0.019)
Girls		-0.007
		(0.013)
Difference	0.076	0.024
	(0.020)	(0.023)

NOTE: The outcome for all columns is graduation. The first column shows RDD estimates for low and high GPA students while holding gender balanced across subgroups. The second column shows RDD estimates for boys and girls while holding GPA balanced across subgroups. The last row shows the difference in treatment effects across subgroups. Standard errors in parenthesis.

N Within school ranking by GPA

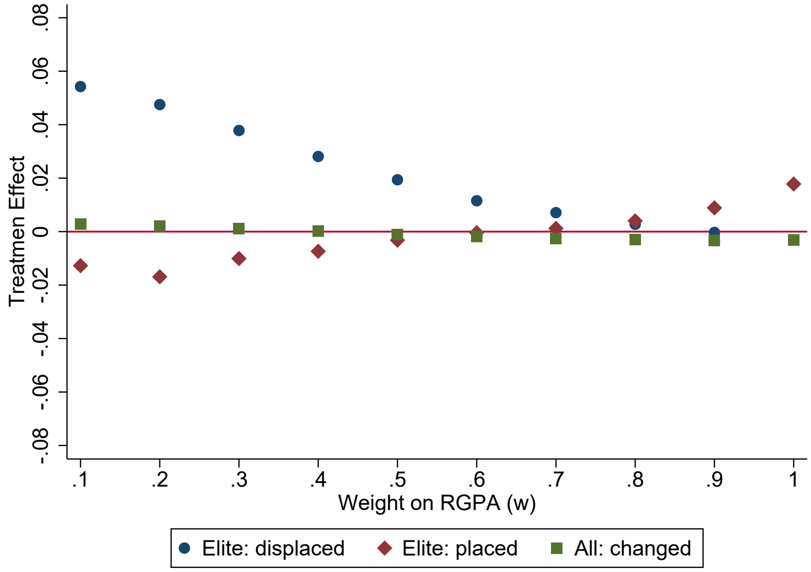
In this section, we include the results of a counterfactual analysis that increasingly adds weight to applicants' within-middle school percentile ranking by GPA.

Figure 17: Changes in the composition of students at elite schools (RGPA)



Note: This figure shows the change in the share of girls and low-SES students admitted to elite schools in each counterfactual equilibrium ($\omega > 0$) with respect to the shares in the baseline ($\omega = 0$). The x-axis indicates the weight on RGPA (ω). The weight on the admission exam score is $(1-\omega)$. We define a low-SES student as one whose family income is lower than 5000 Mexican pesos per month (458 USD).

Figure 18: Treatment effects on graduation (RGPA)



NOTE: This figure shows the average treatment effects on graduation for three subgroups of students and different priority orders indexed by ω. The x-axis indicates the weight on RGPA ω. The weight on the admission exam score is $(1-\omega)$.