

# Perceived Ability and School Choices: Experimental Evidence and Scale-up Effects\*

Matteo Bobba<sup>†</sup>

Veronica Frisancho<sup>‡</sup>

Marco Parigi<sup>§</sup>

May 2023

## Abstract

This paper studies an information intervention designed and implemented in the context of a school assignment mechanism in Mexico City. We find that providing students from socio-economically disadvantaged backgrounds with feedback about their academic performance contributes to placing applicants in schools that better fit their skills, allowing them to graduate on time from high school at a higher rate. We also quantify the effect of a counterfactual and yet feasible implementation of the information intervention at a much larger scale. Simulation results demonstrate substantial heterogeneity in the demand-side responses, which trigger sorting and displacement patterns within the assignment mechanism. The equilibrium effects of the intervention may possibly hinder the subsequent academic trajectories of high-achieving and socio-economically disadvantaged students.

**Keywords:** Subjective expectations, Information provision, School choice, Upper-secondary education, Scaling up experiments, Spillover and equilibrium effects.

**JEL codes:** D83, I21, I24, J24.

---

\*This draft builds on and supersedes our earlier working papers “Perceived Ability and School Choices” and “Learning About Oneself: The Effect of Performance Feedback on School Choices” (Bobba & Frisancho 2016). We are grateful to the Executive Committee of COMIPEMS, as well as to Ana Maria Aceves and Roberto Peña of the Mexican Ministry of Education (SEP) for making this study possible, to *Fundación IDEA*, C230/SIMO, the Inter-American Development Bank, and Maria Elena Ortega for the support with the field work, as well as to Jose Guadalupe Fernandez Galarza for invaluable help with the administrative data. Orazio Attanasio, Pascaline Dupas, Anais Fabre, Yinghua He, Thierry Magnac, Christopher Neilson, Imran Rasul, and Basit Zafar provided us with helpful comments and suggestions. We also thank Matias Morales, Jonathan Karver, and Nelson Oviedo for excellent research assistance. Matteo Bobba acknowledges financial support from the ANR under grant ANR-17-EURE-0010 (Investissements d’Avenir program). This study is registered in the AEA RCT Registry and the unique identifying number is: AEARCTR-0003429.

<sup>†</sup>Toulouse School of Economics, University of Toulouse Capitole. E-mail: matteo.bobba@tse-fr.eu.

<sup>‡</sup>CAF, Development Bank of Latin America. E-mail: vfrisancho@caf.com.

<sup>§</sup>Department of Economics and the Murphy Institute, Tulane University. E-mail: mpariguana@tulane.edu.

# 1 Introduction

One of the key drivers behind the increasing adoption of randomized evaluations in economics and other social sciences has been a genuine ambition to directly inform policy making. The compelling evidence provided by field experiments has contributed to the government implementation of effective programs or policies in various countries (Duflo 2017, 2020). Such a laudable goal has, however, been undermined by a “scale-up problem”. This is the tendency for the size effect of an intervention to diminish, if not vanish, when that intervention is scaled up to reach a larger (and more diverse) population of recipients (List 2022).

Several studies have focused on the technological challenges associated with scaling up a given program. Indeed, one of the most common threats to scalability, with respect to the protocol of an experiment, is to ensure the fidelity of the implementation at scale.<sup>1</sup> This paper is, instead, concerned with situations where such scaling up is quite straightforward. This is the case of information interventions in which it is often feasible to follow the original blueprint of the program. Conversely, the large-scale implementation of these interventions is prone to generating a variety of spillover and equilibrium effects, which may either offset or sustain the effectiveness of the policy. In markets for educational services, for instance, information provision has been shown to alter the equilibrium price and quality of schools through supply-side responses (Andrabi et al. 2017, Neilson et al. 2019).

We design and evaluate an intervention that provides students with individualized feedback on their academic skills in the context of a centralized school assignment mechanism in Mexico City. Students apply to the system during the second-to-last term while in the ninth grade (i.e., the last year of middle school) by submitting rank-ordered lists of their preferred high school programs. At the end of the school year, all applicants take a unique standardized admission test, which assesses curricular knowledge as well as verbal and analytical aptitude. The somewhat unusual timing of these events implies that high-stake decisions regarding school choice are not able to incorporate all relevant information about an applicant’s academic skills. We administer a mock version of the admission test among a sample of approximately one percent of the applicants (N=2,493) across 90 middle schools and communicate individual score results to a randomly chosen subset of students before the school rankings are submitted. In this setting, the score in the mock exam provides students with a signal about their own academic potential that is easy to interpret and contains relevant information on the individual-specific returns of attending different high school programs.

---

<sup>1</sup>See, e.g., Al-Ubaydli et al. (2020), Banerjee et al. (2017), Muralidharan & Niehaus (2017) for the recent literature on scaling up randomized experiments and Agostinelli, Avitable & Bobba (2021), Caron et al. (2021) on the issues of accurately replicating the exact program designs at a larger scale. Eligible participants in the experimental evaluation may also be positively selected and not representative of the general population (Allcott 2015, Davis et al. 2021).

Results from the experiment show that providing individual feedback on exam scores substantially shifts students' belief distributions regarding their own academic performance. We document relatively larger updates among lower performing students, who display wider gaps between the expected score and their actual performance in the mock exam. While on average, the information intervention does not systematically alter school choices or placement outcomes, it differentially improves the alignment between actual skills and the demand for academically-oriented schools. Better performing (lower performing) students in the treatment group increase (decrease) the share of academic vis-a-vis non-academic options in their school rankings when compared to those in the control group. This choice response translates into differential placement outcomes, which, in turn, systematically alters subsequent educational trajectories. Three years after school assignment, the probability of graduating from high school on time is, on average, 4 percentage points higher among students who received performance feedback ( $p$ -value=0.055). This corresponds to a seven percent increase when compared to the sample average of high school graduation rates for the control group. The effect of the intervention on this relevant downstream educational outcome is more pronounced at the lower-end of the academic achievement distribution. Since lower performing students are less likely to successfully complete their upper secondary education, the information treatment contributes to reducing education inequality in our setting.

We next study the effects of a counterfactual scaled-up version of the experimental intervention that mandates the application of the admission exam and the disclosure of the individual score results to *all* applicants before the submission of their school rankings. We rely on a simple discrete choice model that captures rich heterogeneity across students regarding their valuations over the available schooling alternatives (see, e.g., [Abdulkadiroglu et al. 2017, 2020](#), [Arcidiacono et al. 2016](#)). Taking the mock exam without receiving feedback on their score results makes applicants in the experimental control group comparable to the average applicant in the assignment mechanism, in terms of the information on their own academic skills. This assumption, coupled with the parametric structure of the school choice model, allows us to extrapolate the estimated demand-side parameters for the students in the experimental control group toward the much larger population of applicants in the school assignment mechanism. Analogously, the estimated model parameters for the students in the experimental treatment group (i.e., those who were provided with information on their performance in the mock exam) approximate the counterfactual policy scenario whereby all applicants would be provided with their admission test scores.

We simulate out-of-sample the predicted school valuations for the approximately 280,000 applicants who did not participate in the experiment, and quantify the effect of the information intervention at scale, accounting for equilibrium effects. Supply-side responses are held constant in the

matching equilibrium, since schools simply accept or reject prospective applicants in descending order, based on their priority, until capacity constraints are reached, which are assumed invariant to the intervention. The estimated distribution of school valuations under the status quo scenario tracks the average assignment patterns observed in the data remarkably well, in addition to the equilibrium cutoff scores at the school level. The new matching equilibrium for the counterfactual regime of the information intervention effectively takes into account the sorting and displacement effects resulting from aggregate changes in the demand side, as well as the associated movements in cutoff scores.

On average, the provision of information makes students slightly more likely (by two percentage points) to be assigned to a school of their choice. More importantly, the share of students assigned to their most preferred school option increases by 9 percentage points, from 16 percent to 25 percent. These findings suggest a positive average impact of the information intervention on student welfare. The bulk of the changes in the school choices between the status quo and the policy scenarios are concentrated among applicants who live in relatively wealthy areas. These applicants *increase* their demand for academic schools and symmetrically *decrease* the demand for highly selective and prestigious (elite) schools as a result of the information intervention. Instead, and in line with the experimental evidence, applicants who reside in relatively more disadvantaged neighborhoods are, on average, unresponsive to the information intervention. These sorting patterns spur reallocation effects within the assignment system. The lower demand-side pressure on elite programs leaves open seats available for high-achieving and socio-economically disadvantaged applicants. Indeed, the information intervention effectively doubles the representation of applicants from low-income neighborhoods in elite high school programs, from 10 percent to 20 percent. This result underscores the quantitative importance of equilibrium effects that arise during the large-scale implementation of the information intervention in our setting.

In the last part of the analysis, we leverage sharp discontinuities in admission probabilities generated by the strict-priority feature of the assignment mechanism, in order to gauge the consequences of such an equilibrium effect on socio-economically disadvantaged applicants. Marginal admission to an elite school decreases the rate at which low-income students graduate on time from high school by 11-12 percentage points. These findings are in line with previous evidence from the same setting, which documents a negative effect of marginal elite admission on the educational outcomes of academically weaker students (Dustan et al. 2017, Pariguana & Ortega 2022). The regression-discontinuity estimates are, of course, local in nature, thus, it is difficult to extrapolate these effects to other disadvantaged applicants who have admission scores that are sufficiently far away from the admission cutoff scores of elite schools. However, we show that the sub-set of ap-

plicants who would gain admission into elite schools under the counterfactual policy simulation for the most part coincide with the marginally admitted applicants under the regression discontinuity design. Therefore, the regression-discontinuity evidence is policy relevant for this group.

Taken together, the body of the evidence presented in this paper fully characterizes the impacts of an information intervention implemented in the context of a school assignment mechanism in Mexico City. Experimental results drawn from a relatively small sample of applicants demonstrate that providing students from less advantaged socioeconomic backgrounds with informative feedback about their academic performance triggers a substantial reallocation effect across high school tracks, which allows students to graduate on time from high school at a higher rate. Further insights generated from a counterfactual and yet feasible large-scale implementation, combined with quasi-experimental variations of unpredictable admission cutoffs, suggest that the sorting and displacement effects of the intervention at scale may partly offset the impact on on-time graduation documented in this small-scale evaluation. In fact, on-time graduation rates likely decrease among high-achieving and disadvantaged students who would otherwise be admitted to elite schools under the intervention at scale when compared to those who are admitted to non-elite schools.

**Relationship to the literature** There is an emerging consensus that information interventions in educational settings can shift subjective beliefs and individual choices, although the specific effects depend on context, implementation, and design details (see, e.g., [Haaland et al. 2023](#), [Lavecchia et al. 2016](#)). We build on this line of work by focusing on the role of perceptions about one’s own ability in a context where beliefs are tightly linked to concrete, immediate, and high-stake choices. While other papers have studied the mechanisms through which feedback on students’ academic performance affects educational decisions and outcomes ([Azmat et al. 2019](#), [Bergman 2021](#), [Dizon-Ross 2019](#)), the longitudinal span of our data enables us to assess the medium-run impact on academic trajectories triggered by the short-term changes in placement outcomes.

Evidence regarding the equilibrium effects of large-scale information interventions remains scarce in the literature. In the context of educational policies, [Andrabi et al. \(2017\)](#) evaluate a market-level experiment in Pakistani villages showing that information on school quality and price can lead to changes in aggregate educational outcomes. [Neilson et al. \(2019\)](#) is probably the closest contribution to our study. The authors study a small-scale experiment in Chile that provides personalized information to parents about the characteristics of nearby schools and they approximate the effects of such a program at large. These papers crucially rely on modelling assumptions on the supply-side of the education market, which is the driver behind the equilibrium increase in school quality. One advantage of our study is the presence of a centralized assignment mechanism, which considerably simplifies the simulation of the market equilibrium ([Agarwal & Somaini 2020](#)). This

feature enables us to unpack the equilibrium effects of the information intervention in a transparent way. Another benefit of our analysis is that we are able to track the medium-run consequences of such equilibrium effects.

Methodologically, our approach builds upon recent attempts to combine natural or field experiments with model-based estimation techniques (see, e.g., [Galiani & Pantano 2021](#), [Low & Meghir 2017](#), [Todd & Wolpin 2023](#)). Our experimental design enables us to flexibly incorporate the provision of information into the model of school choice by exploiting the differential valuations over the schooling alternatives across applicants in the treatment and the control groups. In turn, a discrete choice model with rich (observed) heterogeneity along applicants' characteristics serves to extrapolate the demand-side responses outside of the experimental sample. Finally, a credible source of variation from discontinuities that effectively randomize applicants near admission cutoff scores into different high school tracks (elite vs. non-elite) allows us to causally link the model predictions and subsequent schooling trajectories. Thus, we fully leverage the synergies between different empirical approaches in order to generate new insights into the science of scaling ([Al-Ubaydli et al. 2020](#)).

## 2 Context and Data

### 2.1 Centralized School Assignment in Mexico City

Since 1996, a local commission (COMIPEMS, by its Spanish acronym) has centralized public high school admissions in Mexico City's metropolitan area by means of an assignment mechanism. In 2014 (the year of our intervention), over 238,000 students were placed in 628 public high schools, accounting for approximately three-quarters of enrollments in the entire metropolitan area. The remaining portion of high school students sought enrollment in public schools with open admission (10 percent) or private schools (15 percent).

Students apply to the centralized system during the second-to-last term while in ninth grade (i.e., the last year of middle school). Prior to registration, they receive a booklet outlining the timing of the application process and corresponding instructions, as well as a list of available schools, their basic characteristics, and their cutoff scores in the last three rounds. In addition to the registration form, students complete a socio-demographic survey and a ranked list of 20 schools, at most. At the end of the school year, all applicants take a unique standardized achievement test. The submission of school preferences *before* the application of the admission exam is an unusual feature of the COMIPEMS system in relation to other centralized assignment mechanisms that use strict priority rules. The timing of the events in the Mexican case is designed to provide public

officials and sponsoring institutions with a “ballpark estimate” of the number of seats that should be made available within the assignment process.

Based on their scores, students are ranked in descending order, and the matching algorithm goes down the list to sequentially assign applicants to their most preferred schooling option with available seats. Each applicant is matched with one school. Whenever a tie in the score occurs, members of the local commission agree on whether to admit all of the tied students, or none of them. Unplaced applicants can request admission to other schools with available seats after the allocation process is over or to search for a seat in schools with open admissions outside the system. When an applicant is not satisfied with their placement, they can request admission to another school in the same way unplaced applicants do. In this way, the assignment system discourages applicants from remaining unplaced and/or to list schools that they will ultimately not enroll in; specifically, participating in the second round will almost certainly imply being placed in a school that is not included in the student’s original ranking. In practice, the matching algorithm performs well: among all applicants who graduate from middle school and take the admission exam, only 12.8% remain unplaced and 3.2% are admitted through the second round of the matching process.

The Mexican system offers three educational tracks at the upper secondary level: General, Technical, and Vocational Education. Each school within the assignment system offers a unique track. The general track is academically oriented and includes traditional schools that are more focused on preparing students for tertiary education. Technical schools cover most of the curriculum of general education programs, but they also provide additional courses allowing students to become technicians upon completion of high school. The vocational track exclusively trains students to become technically adept.<sup>2</sup>

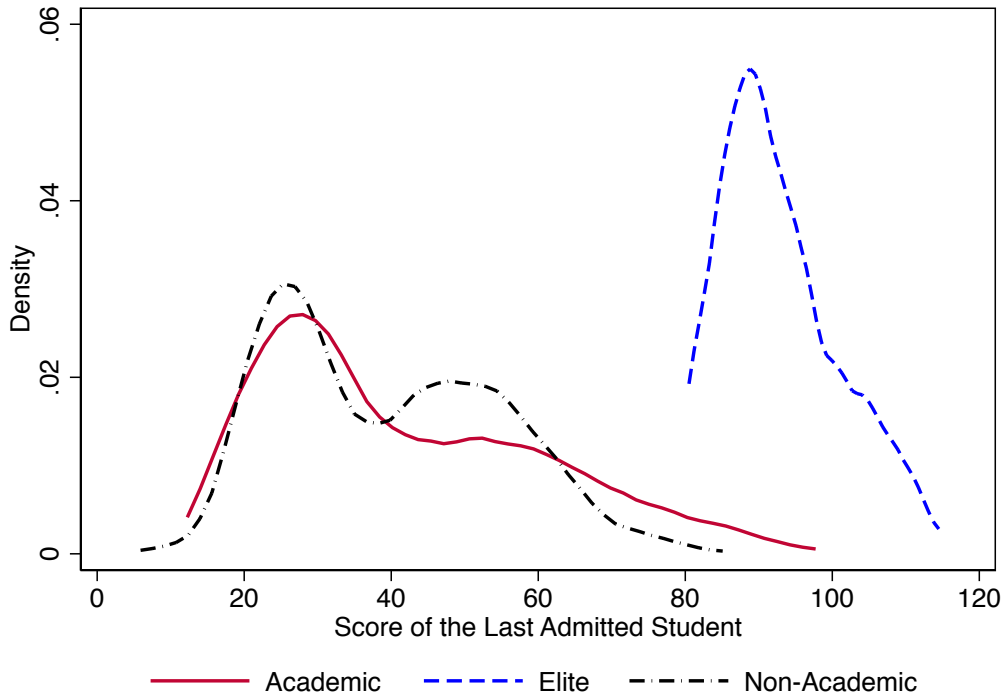
A small sub-set of schools (32 out of 628) within the assignment system in Mexico City is affiliated with two higher education institutions (the National Polytechnic Institute and the National Autonomous University, IPN and UNAM by their Spanish acronyms), which are highly selective and prestigious universities, and as such these high schools are highly demanded. These programs exclude from consideration students with a middle school GPA lower than seven out of a 10-point grade scale. However, most of the applicants meet this requirement (more than 90 percent in every round of the assignment system). In what follows, we define UNAM- and IPN-sponsored high school programs as ‘elite schools’. All the non-elite general track schools are considered ‘academic schools’ while the remaining technical and vocational programs are ‘non-academic schools’. As

---

<sup>2</sup>Data from a nationally representative survey of individuals aged 26-35 in urban Mexico (ENTELEMS, 2008) shows that less than 40 percent of the graduates from technical or vocational high schools successfully obtain a tertiary education degree. Attending the academic track yields a positive premium of 12 percentage points in terms of average hourly wages.



Figure 1: Distribution of Cutoff Scores



Note: Cutoff scores for each high school program refer to the 2014 assignment process. ‘Academic’ schools are defined as the high school programs in the general track, ‘Non-academic’ schools are those in the technical and vocational tracks, and ‘Elite’ schools are affiliated with two higher education institutions (the National Polytechnic Institute and the National Autonomous University, IPN and UNAM by their Spanish acronyms).

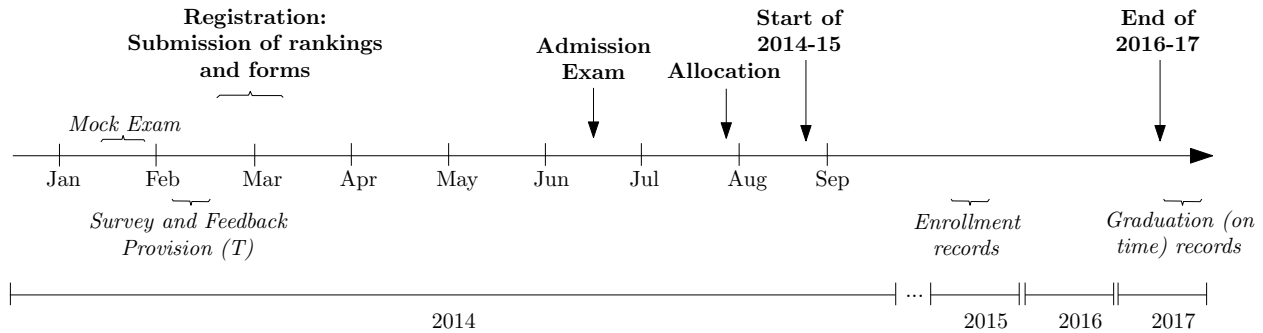
shown in Figure 1, academic schools are, on average, slightly more selective than non-academic schools, but there is a large overlap in the distributions of equilibrium cut-off scores across these two tracks. Elite schools clearly stand out in terms of selectivity and peers’ quality (e.g., median scores are almost perfectly correlated with cutoff scores).

## 2.2 The Information Intervention

Figure 2 depicts the timing of the activities related to the intervention. The admission exam is applied three to four months after students have submitted their rank-ordered lists of preferred high school programs, by mid-June in each academic year. During the second half of the 2013-14 academic year, we administered a mock exam in 90 middle schools as our experimental sample (see Section 2.3). One or two weeks later, and just before the submission of the school rankings, surveyors provided students with individual feedback on their performance in the mock exam. The delivery of the test scores took place in a setting secluded from other students or school staff in



Figure 2: Timeline of Events



Note: This figure depicts the timing of the activities related to the intervention (in italics) and the relevant phases of the assignment process throughout the school calendar year (in bold).

order to avoid reporting biases due to the influence of peers and/or social image concerns (Burks et al. 2013, Ewers & Zimmermann 2015). Surveyors showed each student a personalized graph with two pre-printed bars: the average score in the universe of applicants during the 2013 edition of the school assignment mechanism and the average mock exam score in the student’s class. Surveyors plotted a third bar corresponding to the student’s score in the mock exam. Both pre-printed bars served the purpose of providing the student with additional elements to better frame her own score, which is the main object of interest of the performance feedback.

The mock exam was designed by the same institution responsible for the official admission exam, in order to mirror the latter in terms of structure, content, level of difficulty, and duration (three hours). The test is comprised of 128 multiple-choice questions worth one point each, without negative marking, covering a wide range of subjects that correspond to the public middle school curriculum (Spanish, mathematics, social sciences and natural sciences) as well as mathematical and verbal aptitude sections.<sup>3</sup> We informed students, parents, and school principals about the benefits of additional practice for the admission exam. We also made sure that the school principal was able to assign the person who is usually in charge of the academic discipline and/or a teacher to proctor the exam, alongside the survey enumerators.

In order to support the notion that students took the mock exam seriously, we look at the pattern of skipped questions (Akyol et al. 2018). Without negative marking, the expected value of guessing is always higher than leaving a question blank, which implies that students have no incentive to skip a question. Indeed, the average number of skipped questions in the mock exam was only 1.4

<sup>3</sup>Since the mock exam took place before the end of the school year, 13 questions related to curricular content that was not yet covered were not graded. We normalize the raw scores obtained in the 115 valid questions to the 128-point scale.

out of 128, and more than 80 percent of the students did not leave any question unanswered. Figure B.1 in the Appendix shows that the average patterns of skipping questions are more consistent with binding time constraints, rather than a lack of effort exerted in test taking. Furthermore, we do not find differential skipping patterns according to either the score in the admission exam or individual traits linked to effort and persistence.

We argue that the score in the mock exam was easy to interpret for the applicants in the assignment mechanism while providing additional and relevant information about their academic skills. The linear correlation in our sample between performance in the mock exam and the actual exam is 0.82. This relationship does not vary along the exam score distribution. In turn, the linear correlation between a freely available signal, such as the middle school GPA and the admission exam score, is only 0.48. The mock exam score also predicts later educational outcomes for the applicants in the control group: a one-standard-deviation increase in the mock exam score is associated with a 4.7 percentage-point increase (std.err.=0.019) in the probability of graduating from high school on time.

### **2.3 Sample Selection and Randomization**

To select the experimental sample, we focus on middle schools with a considerable mass of applicants (more than 30) in a previous round of the centralized mechanism and that are located in neighborhoods with high or very high poverty levels (according to the National Population Council in 2010). The latter criterion reflects previous evidence that shows that less privileged students tend to be relatively more misinformed when making educational choices (Avery & Hoxby 2012, Hastings & Weinstein 2008, Jensen 2010).

Table 1 shows that the experimental sample is largely comparable to the general population of applicants in terms of initial credentials, such as GPA or college aspirations. However, the applicants in our sample score 4-points lower (0.2 standard deviations) in the admission exam. Consistent with our focus on relatively disadvantaged neighborhoods, the applicants in the experimental sample are less likely to have parents with tertiary education, they attend middle schools with lower performing students, and reside in neighborhoods with a much higher prevalence of poverty. The average difference in the poverty index computed at the neighborhood-level between the population of applicants and our sample is 1.5 standard deviations.

Schools that comply with our sample selection criteria are grouped into four geographic regions and terciles of school-average math performance amongst ninth graders. Treatment assignment is randomized within strata at the school level. As a result, 44 schools are assigned to a treatment group in which we administer the mock exam and provide face-to-face feedback on performance,

Table 1: Applicants' Characteristics in the Population and in the Sample

	All COMIPEMS	Experiment	Difference [ <i>p</i> -values]
Admission exam score	69.506 (20.705)	65.400 (19.401)	4.107 [0.000]
Grade Point Average in middle school (GPA)	8.058 (0.871)	8.119 (0.846)	-0.061 [0.001]
Has some disabilities (1=yes)	0.118 (0.323)	0.145 (0.352)	-0.027 [0.000]
Scholarship in middle school (1=yes)	0.116 (0.320)	0.110 (0.313)	0.006 [0.401]
Indigenous	0.041 (0.198)	0.093 (0.290)	-0.052 [0.000]
Plans to go to college (1=yes)	0.662 (0.473)	0.670 (0.470)	-0.008 [0.378]
One parent with at least tertiary education (1=yes)	0.236 (0.425)	0.147 (0.354)	0.089 [0.000]
Average math score in middle school (z-score)	0.000 (1.000)	-0.208 (0.712)	0.208 [0.000]
Neighborhood poverty index (z-score)	0.000 (1.000)	1.504 (0.494)	-1.504 [0.000]
Observations	284,412	2,493	

NOTE: The first two columns report means and standard deviation (in parentheses) of individual characteristics between the overall population of applicants and the experimental sample. The third column displays mean differences and the associated *p*-values (in brackets) for the null hypothesis of equal means. The observations in the 'All COMIPEMS' column comprise all the applicants in the year 2014 who were eligible to be assigned through the matching algorithm. The observations in the 'Experiment' column comprise the evaluation sample of the randomized information intervention.

while 46 schools are assigned to the control group in which we only administer the mock exam. Within each school, we randomly select one ninth grade classroom to participate in the experiment.

Since the provision of feedback about test performance took place during the survey (see Section 2.4), it cannot induce differential attrition patterns. The match rate between the survey and the application records is 88 percent (2,828 students) and it is not differential across the treatment and the control groups (see column 1 of Table B.3 in the Appendix). We focus our empirical analysis on the 2,493 participating applicants who were eligible for assignment through the matching algorithm.<sup>4</sup>

Appendix Table B.1 provides basic descriptive statistics and a balancing test of the randomization for various applicants' characteristics. Mean differences are very small in magnitude, with no significant differences detected across the treatment group and the control group. On average, 14 percent of the applicants report previous exposure to a mock version of the admission exam with

<sup>4</sup>There are a few observations with missing values in the survey data (247 observations) and in the scores of the mock exam (146 observations, with 78 missing in both variables), which implies an effective sample size of 2,178 applicants for the analysis presented in Section 3.1.

performance feedback, and this share is balanced between the treatment and the control group.

## 2.4 Data and Measurement

Our study relies on several data sources. First, we have access to administrative data on different cohorts of applicants for several rounds of the assignment mechanisms. These records include socio-demographic variables, such as gender, age, household assets, parental education and occupation, personality traits, and study habits, among others. They also contain information on school preference rankings, admission exam scores, and placement outcomes. We can link these records with middle school-level test scores in mathematics for students in the ninth grade in a national standardized examination (Evaluacion Nacional de Logros Academicos en Centros Escolares, ENLACE).

Second, we collect detailed survey data with information on the subjective distribution of beliefs about performance in the admission exam for the students in the experimental sample. In order to help students understand probabilistic concepts, the survey relies on visual aids (Delavande et al. 2011). We explicitly link the number of beans placed in a cup to a probability measure, where zero beans means that the student assigns zero probability to a given event and 20 beans means that the student believes the event will occur with certainty. Students are provided with a card divided into six discrete intervals of the score. Surveyors then elicit students' subjective expectations about test performance by asking them to allocate the 20 beans across the intervals to represent the chances of scoring in each bin. Appendix A provides more detail on the elicitation of the individual data on beliefs in our setting.

Third, we assemble and harmonize data from each of the nine public institutions in the centralized assignment system on individual trajectories through upper secondary education for the 2014 applicants in the sample. It is not possible to track students who enrol in schools outside the centralized system. The resulting panel dataset allows us to measure enrollment and drop-out during the first year in high school (tenth grade) for all the students in our sample as well as graduation on time from high school (twelfth grade) for more than 90 percent of the students. While we also have access to either individual grades or end-of-year GPAs for a few high school programs in the centralized system, we do not use this information in the empirical analysis, since it is available for less than one-third of the observations in our sample. Furthermore, grades are difficult to compare across high school programs.

Since we have access to the longitudinal schooling trajectories exclusively for the experimental sample, we rely on an alternative strategy in order to construct an indicator variable for on-time graduation from high school for all the other applicants to the centralized system. Previous evi-

dence from Mexico shows that the probability of taking the ENLACE test in the twelfth grade is a good proxy for the probability of graduating from high school (Dustan 2020, Dustan et al. 2017, Estrada & Gignoux 2017). The coverage of the ENLACE evaluation was universal throughout the 2008-2013 school years, except for the UNAM-sponsored high schools (see Section 2.1).<sup>5</sup> We have access to administrative records from the 2005-2009 rounds of the centralized system and we match these at the individual level with the corresponding ENLACE scores collected three years after high school assignment.

### 3 Experimental Evidence

Providing information about individual performance in the mock exam potentially allows students to revise their own beliefs and make more informed choices, which, in turn, may lead to better educational outcomes in the medium-run. In this section, we document the effect of performance feedback in the experimental sample on beliefs, school choices, and placement, as well as downstream outcomes shown by the completion of upper secondary education.

Different treatment effect parameters are estimated using simple OLS models: all specifications include the treatment assignment indicator, as well as a set of dummy variables that correspond to the randomization strata, pre-determined characteristics (gender, characteristics of the school of origin, previous experience with practice exams providing feedback, aspirations to attend college, an index of personality traits, an index of parental characteristics, and a household asset index), as well as a set of indicator variables for whether each of the covariates has missing data (Zhao & Ding 2021). Given the relatively large array of hypotheses considered throughout the analysis, we complement the usual asymptotic inference by computing  $p$ -values that are adjusted for multiple hypothesis testing across different families of outcomes (List et al. 2019).

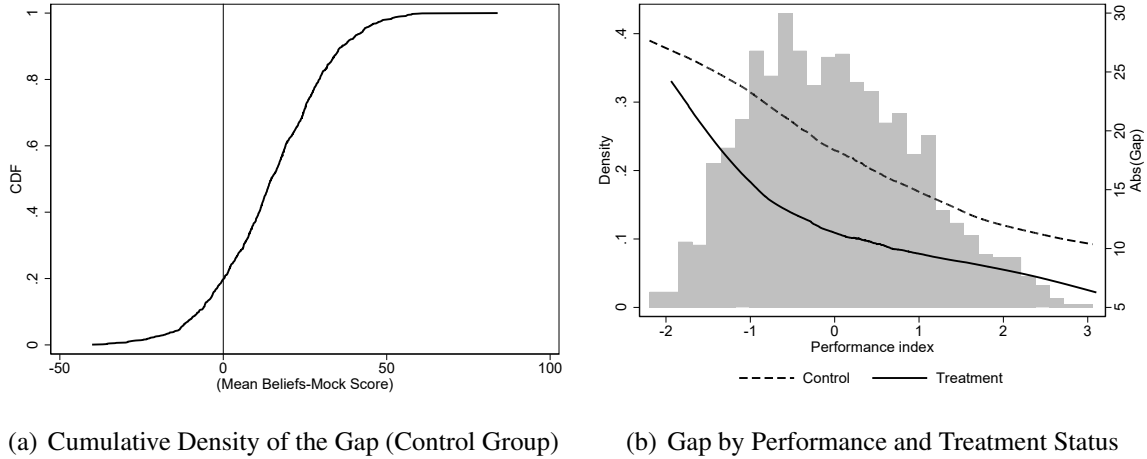
#### 3.1 Subjective Expectations about Test Performance

Panel A in Figure 3 relies on data from the control group and presents the cumulative distribution of the perception gap, defined as the difference between the expected score and the actual performance in the mock exam. Assuming a uniform distribution within each interval of the score, the expected scores are constructed as the summation over intervals of the product of the mid-point of the bin and the probability assigned by the student to that bin. The great majority of applicants, over

---

<sup>5</sup>The ENLACE examination was discontinued after the school-year 2013 to be eventually replaced from 2015 onward with another standardized test that was administered only among a (relatively small) random sample of students within schools; that is, the National Plan for Learning Evaluation (PLANEA).

Figure 3: Gap between Expected Score and Mock Exam Score



NOTE: Using data for the control group, Panel A shows the cumulative density of the difference between the expected scores (as captured in the survey) and scores in the mock version of the admission exam. Panel B depicts the density of the performance index (y-axis on the left), which is a GLS-weighted average of the GPA in middle school, mock exam score, and exam score (Anderson 2008). Overlaid on the density, we show non-parametric locally weighted estimates of the relationship between the absolute value of the gap in beliefs and the performance index separately for applicants in the treatment and control groups (y-axis on the right).

80%, overestimate their performance in the test. Moreover, gaps in performance estimation are substantial for a large share of students. Half of the sample overestimates their performance by more than 14.7 points and 25% of students exhibit absolute gaps greater than 26.7 points out of a 128-point scale.

Panel B in Figure 3 presents evidence on the relationship between the absolute value of the perception gap and academic achievement separately for students in the treatment group and control groups. Achievement is captured through a Generalized-Least-Squares (GLS)-weighted average of the GPA in middle school, the score in the mock exam, and the score in the admission exam (Anderson 2008).<sup>6</sup> Updates on the expected score in response to performance feedback occur along the entire distribution of the performance index, with relatively larger gap reductions among lower performing students, who are also those who display larger prior biases.

Table 2 shows the corresponding OLS estimates of the effect of the information intervention on the absolute value of the difference between the expected score and the score in the mock

<sup>6</sup>The procedure accommodates the construction of the index, even when data on one of the individual performance measures is missing. It does so by setting missing indicator values to zero, which is the mean of the reference group following normalization. The GLS-weighting approach increases efficiency by ensuring that outcomes that are highly correlated with each other receive less weight, while outcomes that are uncorrelated and thus represent new information receive more weight. O'Brien (1984) found this procedure to be more powerful than other popular strategies in the repeated-measures setting.

Table 2: Perception Gap about Test Performance

	Abs(Gap)			
	All Sample	By Performance Index		
		Tercile I	Tercile II	Tercile III
Treatment	-6.809	-8.341	-6.827	-5.170
	[0.000]	[0.000]	[0.000]	[0.000]
	{0.001}	{0.001}	{0.001}	{0.001}
Mean Control	18.8	23.8	18.0	14.6
Number of Observations	2178	683	740	755
Number of Clusters	90	90	90	87
R-squared	0.100	0.131	0.132	0.114

NOTE: All specifications include a set of dummy variables which correspond to the randomization strata, pre-determined characteristics (sex, characteristics of the school of origin, previous experience with practice exams providing feedback, aspirations to attend college, an index of personality traits, an index of parental characteristics, and a household asset index), and indicator variables for whether each of the covariates has missing data. The dependent variable “Abs(Gap)” is the absolute value of the difference between the expected score (as captured in the survey) and the score in the mock exam.  $p$ -values reported in brackets refer to the conventional asymptotic standard errors while those reported in curly brackets are adjusted for testing each null hypothesis across multiple outcomes through the step-wise procedure described in [Romano & Wolf \(2005a,b, 2016\)](#). All inference procedures take into account the clustering of error terms at the middle school level and the block randomization design.

exam. Table B.2 in the Appendix documents additional evidence on the effect of the information intervention for other moments of the individual belief distributions. The delivery of the individual scores in the mock exam shrinks the absolute value of the perception gap by 6.8 points on average, as shown in the first column of Table 2. The magnitude of this effect is quite large as it is over one-third of the mean absolute gap in the control group. The additional columns in Table 2 further document that gap reductions are more pronounced among the lowest performing students, who are also those with greater average perception gaps. For students in the bottom tercile of the performance index distribution, feedback provision reduces the absolute value of the perception gap by 8.3 points on average. While significant gap reductions occur in the second and third terciles, the size of the updating is smaller in absolute terms. In relative terms, the decrease in the perception gap is similar across the distribution of academic achievement and equivalent to about one-third of the perception gap in the control group.

On average, students in our sample seem to hold rather over-optimistic expectations about their performance in the test. Providing information about individual performance in the mock exam allows them to substantially revise their beliefs more toward their performance provided by the



mock exam score. While the information intervention is effective at partly correcting biased self-views about academic performance, an important part of the perception gap (about two-thirds, on average) remains unexplained. These findings underline the informativeness of the score in the mock exam while, at the same time, point toward a high degree of persistence of priors in the distribution of posterior beliefs. We explore these issues in further details in our companion paper (Bobba & Frisancho 2022).

## 3.2 School Choices and Placement Outcomes

In our setting, the high school programs that are available through the centralized assignment system differ in terms of the curricular track that they offer. Academically-oriented schools tend to provide students with general skills and adequate training to pursue a college education. Non-academic schools, that are either technical or vocational, focus more on fostering specific skills that can enable access to the labor market after the completion of secondary schooling. Schools also differ greatly in terms of "peer quality" and the type of college career that their graduates may pursue. For instance, elite high school programs have the highest performing applicants within the centralized admission system (see Figure 1).

As mentioned in Section 2.1, school placement under the assignment mechanism depends exclusively on two student-level observable factors: individual school rankings and the score in the admission exam. Exposure to performance feedback does not systematically affect the fraction of applicants assigned in the first or second round of the assignment process (see Table B.4 in the Appendix), or the scores in the admission exam (see column 2 in Table B.3 in the Appendix).<sup>7</sup> Any treatment-control differences in the final assignment of students are mainly driven by the observed differential changes in their school choices.

Table 3 presents the treatment impacts of feedback provision on school rankings and placement outcomes. The first row in the table shows that, on average, the provision of performance feedback does not affect applicants' demand or placement across high school tracks. However, the estimated coefficients on the interaction term between the performance feedback indicator and the performance index (Treat  $\times$  Performance Index) imply a substantial reallocation effect across tracks of the information intervention. The estimates in the first column indicate that a one-standard-deviation increase in students' scholastic success in the treatment group decreases the share of non-academic schools requested by 3.2 percentage points. This effect is precisely es-

---

<sup>7</sup>This evidence is consistent with recent experimental findings reported in Azmat et al. (2019), whereby the short-term responses to the provision of information on the ordinal ranking of college students in Spain tend to rapidly fade over time. In our setting, the admission exam takes place more than four months after the provision of feedback about performance in the mock exam (see Figure 2).

Table 3: School Choices and Placement Outcomes

	Non-academic Schools		Academic Schools		Elite Schools	
	Share in ROL	Placement	Share in ROL	Placement	Share in ROL	Placement
Treatment	0.002 [0.918] {0.999}	0.047 [0.074] {0.195}	-0.001 [0.923] {0.999}	-0.045 [0.076] {0.195}	-0.000 [0.986] {0.999}	-0.002 [0.845] {0.998}
Performance Index	-0.031 [0.002] {0.007}	-0.079 [0.000] {0.001}	-0.054 [0.000] {0.001}	-0.086 [0.000] {0.001}	0.084 [0.000] {0.001}	0.165 [0.000] {0.001}
Treat × Performance Index	-0.032 [0.010] {0.030}	-0.065 [0.015] {0.041}	0.030 [0.008] {0.023}	0.041 [0.046] {0.107}	0.002 [0.865] {0.999}	0.024 [0.243] {0.549}
Mean Control	0.363	0.401	0.328	0.369	0.309	0.114
Number of Observations	2493	2493	2493	2493	2493	2493
Number of Clusters	90	90	90	90	90	90
R-squared	0.154	0.101	0.129	0.061	0.264	0.335

NOTE: All specifications include a set of dummy variables that correspond to the randomization strata, pre-determined characteristics (sex, characteristics of the school of origin, previous experience with practice exams providing feedback, aspirations to attend college, an index of personality traits, an index of parental characteristics, and a household asset index), and indicator variables for whether each of the covariates has missing data. The dependent variable “Share in ROL” in the odd columns denotes the share of high school programs in the school rankings submitted by each applicant that belong to a given group of schools (i.e., Non-academic, Academic, and Elite schools). The dependent variable “Placement” in even columns denotes an indicator variable that is equal to one if the applicant is assigned to a given group of schools. The performance index is a GLS-weighted average of the GPA in middle school, mock exam score, and exam score.  $p$ -values reported in brackets refer to the conventional asymptotic standard errors, while those reported in curly brackets are adjusted for testing each null hypothesis across multiple outcomes through the step-wise procedure, as described in [Romano & Wolf \(2005a,b, 2016\)](#). All inference procedures take into account the clustering of the error terms at the middle school level and the block randomization design.

timated ( $p$ -values range between 0.01 and 0.03) and sizable, as it corresponds to about 10 percent of the average share of non-academic options in the school rankings of applicants in the control group. This composition change in the demand for non-academic schools significantly alters the assignment patterns realized under the mechanism, as shown in the second column of Table 3. Placement in non-academic schools decreases along the distribution of academic achievement for the students in our sample. A one-standard-deviation increase in the performance index among students in the treatment group is associated with a 6.5 percentage-point lower probability of being placed into a non-academic program (or a 16 percent decrease when compared to the average probability of assignment in the control group).

The third and fourth columns in Table 3 present the estimated effects of the information intervention on school choices and placement into (non-elite) academic programs. A one-standard-deviation increase in achievement increases the share of academic schools in the submitted rankings by 3 percentage points for the applicants who receive the performance feedback. This effect is statistically significant and it is similar in magnitude to that presented in column 1, which suggests that higher performing students who receive the performance feedback substitute non-academic with academic programs in their school rankings almost one-to-one. Column 4 shows the esti-

mated treatment effects on placement into academic programs. The impact is symmetric relative to the estimated effect on placement into non-academic programs, as reported in column 2. A one-standard-deviation increase in academic achievement for the applicants who receive the performance feedback increases placement into such programs by approximately 4 percentage points.

Finally, the estimates reported in the last two columns of Table 3 show that the provision of performance feedback does not systematically alter preferences for or assignment into elite schools. Given the relatively large changes in beliefs documented above (see Section 3.1), the finding that the demand for elite programs is relatively insensitive to information about students' own academic skills is consistent with the fact that the applicants of the centralized mechanism do not factor-in their chances of admission when ranking their preferred schools. Hence, the choice responses observed here should be interpreted as mainly stemming from changes in the individual valuations over the schooling alternatives proposed within the assignment system.

These findings indicate that the provision of performance feedback has real consequences on the sorting patterns across high school tracks in our setting. While the average effect of the information intervention on school choices and placement outcomes are small and statistically insignificant, higher performing students in the treatment group are less likely to be placed in a vocational or technical high school program, and they are more likely to sort into a (non-elite) academic program. The observed composition effect across high school tracks is consistent with the performance feedback, thus enabling greater alignment between applicants' observed individual skills and the careers that they are able to access through the centralized assignment mechanism.

### **3.3 Educational Outcomes**

The centralized assignment mechanism appears to generate school-placement outcomes that are satisfactory for the great majority of the applicants, at least in the short-run. Administrative records from the 2016-17 academic year show that about 80 percent of the students in the control group enroll in the school to which they were assigned through the centralized process. However, among this same group of students, only 52 percent graduate on time from high school. There is very little heterogeneity ascribed by high school track, with timely graduation rates in the academic and non-academic tracks at 50 and 54 percent, respectively. These figures clearly reflect inadequate academic progress through upper secondary education, due to either school dropout or grade retention, which are both strong indicators of a mismatch between schooling careers and students' individual skills. As shown in the previous sub-section, the provision of performance feedback likely improves the alignment between (measured) academic skills and high school track choices. The associated change in the sorting patterns of students across schools may thus result in a better

Table 4: Education Outcomes in Upper Secondary

	Enrollment	Dropout in 1st year	Graduation on Time
Treatment	-0.005 [0.751] {0.697}	0.015 [0.500] {0.614}	0.040 [0.075] {0.055}
Mean Control	0.8	0.2	0.6
Number of Observations	2492	2023	1888
Share of Missing Data	0.000	0.189	0.243
Lee lower bound		0.008	0.035
Lee upper bound		0.011	0.060
R-squared	0.064	0.167	0.127

NOTE: All specifications include a set of dummy variables that correspond to the institution in which the student was placed, pre-determined characteristics (sex, characteristics of the school of origin, previous experience with practice exams providing feedback, aspirations to attend college, an index of personality traits, an index of parental characteristics, and a household asset index), and indicator variables for whether each of the covariates has missing data. The dependent variable “Enrollment” denotes an indicator variable that is equal to one if students enroll in the high school programs they were assigned to, and zero otherwise. The dependent variables “Dropout, 1st year” captures whether the student stopped attending classes or actively dropped out of school, conditional on enrollment. The dependent variable “Graduation on Time” denotes an indicator variable that is equal to one if the student successfully completes the high school programs three years after enrolling in tenth grade, and zero otherwise. Lee bounds (Lee 2009) are reported at the bottom of the table in order to account for potentially non-random sample attrition.  $p$ -values reported in brackets refer to the conventional asymptotic standard errors, while those reported in curly brackets are adjusted for by testing each null hypothesis across multiple outcomes through the step-wise procedure, as described in Romano & Wolf (2005a,b, 2016). All inference procedures take into account clustering of the error terms at the middle school level and the block randomization design.

match that may potentially alter students’ academic trajectories and improve downstream educational outcomes.

The point estimate reported in the first column of Table 4 shows that, on average, there are no discernible differences in the high school enrollment rates between students in the treatment and control groups. Conditional on enrollment, the second column documents that changes in the sorting patterns across tracks induced by feedback provision do not systematically alter dropout rates during the first year of high school. While censoring in the high school sample is unlikely to be differential across the treatment and the comparison group (see estimates in the first column of Table 4), we still report Lee bounds (Lee 2009) at the bottom of the second column in order to account for potentially non-random sample attrition. The bounds are narrow and broadly consistent with the very small point estimate for the effect of the performance feedback on dropout rates during the first academic year of high school.

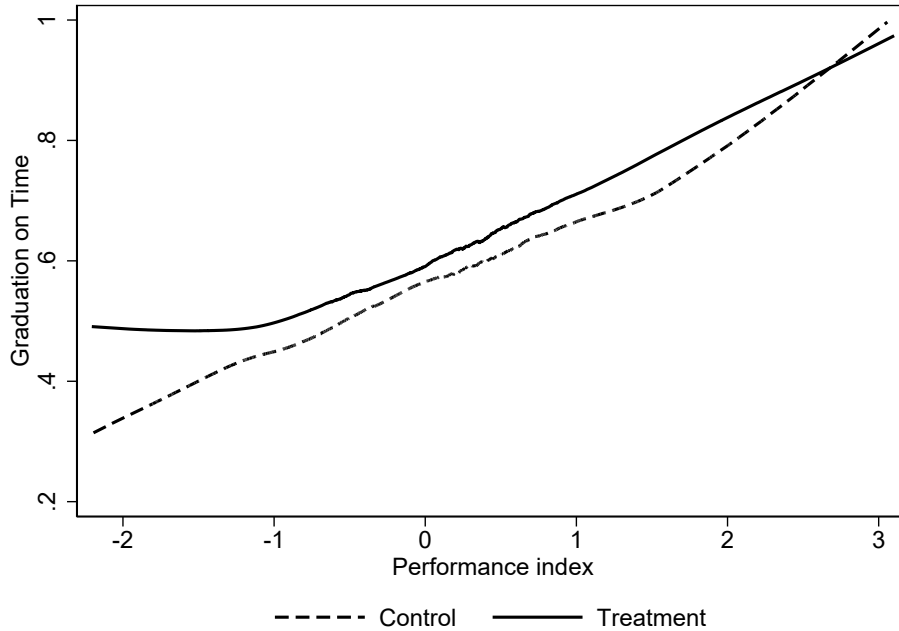
While these findings appear to show only small and insignificant early effects on academic

trajectories, the information intervention does seem to play a role in the medium term. The point estimate reported in the third column of Table 4 documents that the probability of graduation on time is almost 4 percentage points higher for students who receive performance feedback when compared to those who do not. While only marginally significant ( $p$ -values range between 0.055 and 0.075), the magnitude of this average effect is quite remarkable, as it corresponds to a 7 percent increase in high school graduation rates when compared to the sample average in the control group. The effect size also roughly coincides with the magnitude of the impact of a one-standard-deviation increase of the score in the mock exam on the rate of graduation on time for the control group (see Section 2.4).

It is important to note that the sample of students included in the regression analysis displayed in the third column is conditional both on high school enrollment and the availability of high school records (see Section 2.4). The censoring due to high school enrollment is unlikely to generate differential attrition patterns, as shown in Column 1. However, any additional censoring due to missing high school records in our sample (8 percent) may potentially bias the estimated treatment effect on graduation on time. The Lee bounds (Lee 2009) associated with the overall attrition rate (high school enrollment+missing records) in the graduation on time sample range between 3.5 and 6 percentage points. Hence, they are fully informative about the size of the estimated average treatment effect in the presence of the attrition rate in our sample.

The sustained impact of the feedback provision on high school graduation rates is a compounded effect of the school choice and assignment responses to the information provision and subsequent changes in the schooling environment, such as peers and other school inputs. Relative to the control group, the information intervention contributes to successfully placing students in schools that provide a better fit with their skills and allows them to progress through high school at a relatively smoother pace. The average effect displayed in Column 3 of Table 4 may mask potential heterogeneity along the achievement distribution. Figure 4 displays evidence on the relationship between the rates of graduation on time and the performance index separately for the treatment and control groups. The plot clearly shows that the effects of performance feedback are present along the entire distribution of academic achievement, with larger effects around the left tail. Since lower performing students also tend to have lower graduation rates, the information intervention contributes to “levelling the playing field” in our setting.

Figure 4: Graduation on Time and Academic Achievement



NOTE: This plot depicts non-parametric locally weighted estimates of the relationship between the graduation on time and the performance index, which is a GLS-weighted average of middle school GPA, mock exam score, and exam score (Anderson 2008). “Graduation on Time” denotes an indicator variable that is equal to one if the student successfully completes the high school programs three years after enrollment in tenth grade, and zero otherwise.

## 4 Scaling-up the Information Experiment

The experimental evidence reported in the previous section might suggest that the centralized assignment system can improve the allocation of students across high school tracks. This would be achieved by implementing a policy that mandates the application of the admission exam and the disclosure of the individual score results to *all* applicants before the submission of the rank-ordered lists. This conclusion, however, is not warranted.

There are at least two threats to scalability in our setting. First and most importantly, the experimental evaluation comprises only a small portion (approximately one percent) of the applicants in the assignment system. Given the nature of the matching equilibrium (i.e., schools simply accept or reject prospective applicants in descending order based on their exam scores until seat capacities are met), aggregate changes in the demand-side would necessarily trigger feedback effects into the assignment system. Sorting and displacement effects across applicants would, in turn, alter the cutoff scores for the different schooling alternatives. Second, the experiment focuses on applicants from relatively disadvantaged backgrounds (see Section 2.3), thereby making it difficult to extrapolate the experimental treatment effects to a larger and more diverse population of students.

In this section, we embed the experimental data in an empirical model of school choice in order to simultaneously overcome both of these challenges and quantify the effect of the intervention at scale.

## 4.1 School Choice Model

Following the school choice literature (see, e.g., [Abdulkadiroglu et al. 2017, 2020](#), [Arcidiacono et al. 2016](#)) and drawing from the experimental evidence discussed in the previous Section, we rely on a multinomial Logit model that captures rich heterogeneity across applicants in the experimental sample in terms of their preferences for schools within the assignment system. We model the (indirect) utility that student  $i$  gets from attending school  $j$  as:

$$u_{ij} = \alpha_j + \beta'_j \mathbf{x}_i + \gamma' \mathbf{x}_i d_{ij} + \epsilon_{ij}, \quad (1)$$

where the composite term  $\alpha_j + \beta'_j \mathbf{x}_i$  denotes the net returns of attending a particular high school within the centralized system. We allow these net returns to flexibly vary between a school-specific component,  $\alpha_j$ , and a student-school match effect,  $\beta'_j \mathbf{x}_i$ . The vector  $\mathbf{x}_i$  contains a broad array of proxies of academic preparation, such as the individual score in the mock exam, the cumulative GPA in middle school, the poverty index at the neighborhood-level, the middle school-average of ENLACE math test scores, and parental education (see Sections 2.3 and 2.4).

The same vector of individual characteristics is also interacted with the geodesic distance  $d_{ij}$  (in kilometers) between the location of the middle school of each applicant and each high school, available through the centralized assignment system. The  $\gamma$  parameter vector captures the individual cost of attending a particular high school program, while accounting for potential heterogeneity in applicants' willingness to travel to a given school. Tuition fees are negligible in this setting and the small differences in the out-of-pocket expenses across school programs are captured by the  $\alpha_j$  parameters. Conditional on  $\mathbf{x}_i$ ,  $d_{ij}$  is assumed orthogonal to the preference shock ( $\epsilon_{ij}$ ). This assumption is usually invoked in the school choice literature (see, e.g., [Agarwal & Somaini 2020](#)). It is violated if students systematically reside near the schools for which they have idiosyncratic tastes, while it may provide a reasonable approximation in our case as we have rich micro-data on students. Furthermore, priorities in the school assignment mechanism do not depend on student locations, thereby alleviating issues related to strategic residential sorting.

All residual unobserved tastes are captured in the error term,  $\epsilon_{ij}$ , that is assumed to be i.i.d. across  $i$  and  $j$  following a type-I extreme value distribution with normalized scale and location. Since discrete choice models depend on differences in payoffs without loss of generality, we nor-



malize the student’s mean utility of not being assigned to any COMIPEMS school program to zero. This outside option captures the value of other schools that are not part of the centralized assignment mechanism or any other labor market entry opportunity not directly observed in the data.

We embed the information intervention into the model of school choice (1) by allowing all parameters to vary with the treatment assignment indicator ( $\alpha_j^T, \beta_j^T, \gamma^T, T \in \{0, 1\}$ ). This specification captures the fact that feedback provision likely alters the choice environment in which applicants operate. The main advantage of such an approach is that it permits us to incorporate preference heterogeneity over the schooling alternatives in a fully flexible manner that is directly informed by the experimental data on choices. Using this framework, the effect of the intervention at scale on both school choices and assignment outcomes will originate through the estimated differential demand-side responses across treated and untreated applicants in the experiment.

## 4.2 Estimation Procedure and Results

We have access to applicant-level data on rank-ordered lists and realized assignment outcomes. While both sources are potentially valuable in order to infer the distribution of school valuations in our setting, school rankings may in fact deviate from true preference orderings due to the limited uncertainty regarding admission outcomes in our setting. School cutoff scores are highly persistent over time, and information about the previous three rounds of the assignment is contained in the booklet that is provided to students when they apply (see Section 2.1). Furthermore, the performance feedback provided with the experiment may decrease the noise in perceived admission probabilities. The institutional cap of 20 schools in the submitted rank-ordered lists is another potential source of truth-telling violation (Haeringer & Klijn 2009, Calsamiglia et al. 2010). A possibly more robust estimation approach relies on a stability assumption of the realized matching equilibrium, which is likely satisfied in the large-market matching mechanism under study. Under stability, the observed match between an applicant and a given school can be interpreted as the outcome of a discrete choice model with individual-specific choice sets (Fack et al. 2019). In our setting, stability-based choice sets depend solely on students’ scores in the admission exam.

The size of the experimental sample is too small to precisely estimate the 600+ school-specific intercepts ( $\alpha_j$ ) and the associated interaction terms ( $\beta_j$ ). We thus group the various high school programs in the centralized system into 17 college-specific intercepts, or groups of schools, that share the same track and that belong to the same public institution of upper secondary education. By doing so, we substantially reduce the number of parameters that need to be estimated in the school choice model (1). Additional variations in the applicants’ valuations of high school pro-

Table 5: Average Marginal Effects for Selected School-Student Match Parameters

	Control Sample		Treatment Sample	
	Coefficient	Std. Error	Coefficient	Std. Error
Academic $\times$ GPA	0.0005	0.0016	0.0036	0.0020
Academic $\times$ Poverty	-0.0061	0.0024	-0.0074	0.0036
Elite $\times$ GPA	0.0053	0.0046	0.0046	0.0057
Elite $\times$ Poverty	0.0064	0.0081	0.0348	0.0099
Selectivity $\times$ GPA	0.0027	0.0012	0.0039	0.0014
Selectivity $\times$ Poverty	-0.0042	0.0019	-0.0058	0.0026

NOTE: This table depicts the estimated average marginal effects for selected school-student match parameters, which are computed using the model estimates shown in Table B.5. These coefficients show the change in the conditional probability that student  $i$  chooses a school  $j$  with a given characteristic (academic, elite, and the degree of selectivity), resulting from a one-unit increase in the individual covariates (GPA and Poverty). Standard errors are computed using the delta method.

grams within colleges is incorporated in the model in two ways. First, the geodesic distance term and its interactions with the vector  $\mathbf{x}_i$  allows students to differently value schools based on their geographic proximity. Second, we incorporate the admission cutoff score observed by the applicants in the previous round of the assignment mechanism into the model as a proxy for the degree of selectivity of each high school program, along with its interaction terms with the vector  $\mathbf{x}_i$  of applicants’ characteristics. The assignment mechanism described in Section 2.1 generates sorting across schools based on individual performance in the admission exam (see Figure 1). The admission cutoff score is thus a good proxy for the quality of peers and the associated level and pace of instruction within the school.

The parameters of the school choice model (1) are estimated by maximum likelihood and the resulting estimates are reported in Table B.5 in the Appendix. The first four columns show that the estimated school valuations for the control group are similar across alternative specifications. Whether or not we include the score in the mock exam, replacing it with the expected score of the test performance or adding random coefficients to selected colleges, does not systematically alter the magnitudes or signs of the estimated net returns of attending the different colleges for those applicants who did not receive the performance feedback.<sup>8</sup> Table 5 shows the average marginal effects as implied by the model estimates for selected school-student match parameters. The es-

<sup>8</sup>The estimated standard deviations of the two normally distributed random coefficients included in the specification in the third column are not reported in Appendix Table B.5. Their magnitudes are reasonably small, both at approximately one-third of the respective mean coefficient, indicating that the rich sources of observed heterogeneity included in the school choice model (1) may be sufficient to capture substitution patterns across schools in our setting. Indeed, the (negative) value of the log-likelihood at convergence does not decrease much under the specification with random coefficients when compared to the other specifications. Perhaps most importantly, the  $p$ -value of the Likelihood-Ratio test statistic against a nested model with no random coefficients is equal to 0.45, indicating that we cannot reject the null hypothesis that the coefficients on the two colleges are fixed.

estimated match-specific returns for treated applicants differ somewhat from their counterparts for the applicants in the control group, confirming that feedback provision changes the choice environment of the students in our sample. This result is broadly consistent with previous empirical school choice models that allow for imperfect information (Neilson et al. 2019, Wiswall & Zafar 2015).

While the single estimated coefficients are difficult to interpret, in general, we highlight some patterns that may be relevant for the out-of-sample predictions and the associated simulations that are discussed in Sections 4.3 and 4.4. A marginal increase in the GPA in middle school increases the probability of choosing an academic school seven-times more in the treatment group than in the control group. Applicants who receive the performance feedback and are more socio-economically disadvantaged are five-times more likely to choose an elite school when compared to those in the control group. Notice that these marginal effects hold after conditioning on the degree of selectivity/peer quality of the schools, which in fact reveals the opposite sign on the effect of poverty on choice probabilities.

### 4.3 Out-of-Sample Predictions and Model Fit

We extrapolate the estimated school valuations from the experimental sample to the universe of applicants in the centralized assignment system. The out-of-sample prediction uses the parametric linear form of the students' indirect utility (1). The only difference with the experimental set-up is that here we replace the individual scores in the mock exam with the individual scores in the admission exam. The underlying assumption is that the intervention at large would provide timely information about applicants' academic skills that is comparable to the information provided through the delivery of mock exam scores.

The estimated model parameters for the students in the experimental control group (i.e., those who took the mock exam but were not provided with information on their performance) likely capture the status quo scenario for the much larger population of applicants in the school assignment mechanism. As discussed in Section 2.3, applicants may have access to preparatory courses for the admission exam, likely featuring a socioeconomic gradient.<sup>9</sup> We assume that taking the mock exam without receiving personalized feedback puts the relatively disadvantaged applicants of the experimental sample in a position similar to that of the average applicant in the system, in terms of the information set about their own academic skills.

---

<sup>9</sup>Administrative records from the 2012 assignment round indicate that 44 percent of the applicants enrolled in schools from more affluent neighborhoods took preparatory courses before submitting their school rankings, but this figure drops to 12 percent among applicants from schools in high poverty areas. This information was not made available in subsequent rounds of the assignment mechanism (including the year of our intervention, 2014).

Table 6: Model Fit on Average Assignment Outcomes

	Data	Model	Difference
Applied in the system (1=yes)	1.00	0.99	-0.01
Assigned in the system (1=yes)	0.87	0.89	0.02
<u>Assigned in:</u>			
Vocational schools	0.14	0.11	-0.03
Technical schools	0.26	0.26	-0.00
Academic schools	0.38	0.41	0.03
Elite schools	0.23	0.22	-0.00
Selectivity (z-cutoff score)	0.75	0.74	-0.01
Academic (above-median selectivity)	0.50	0.50	-0.01
Academic (below-median selectivity)	0.10	0.13	0.03
Non-academic (above-median selectivity)	0.25	0.24	-0.01
Non-academic (below-median selectivity)	0.15	0.13	-0.02

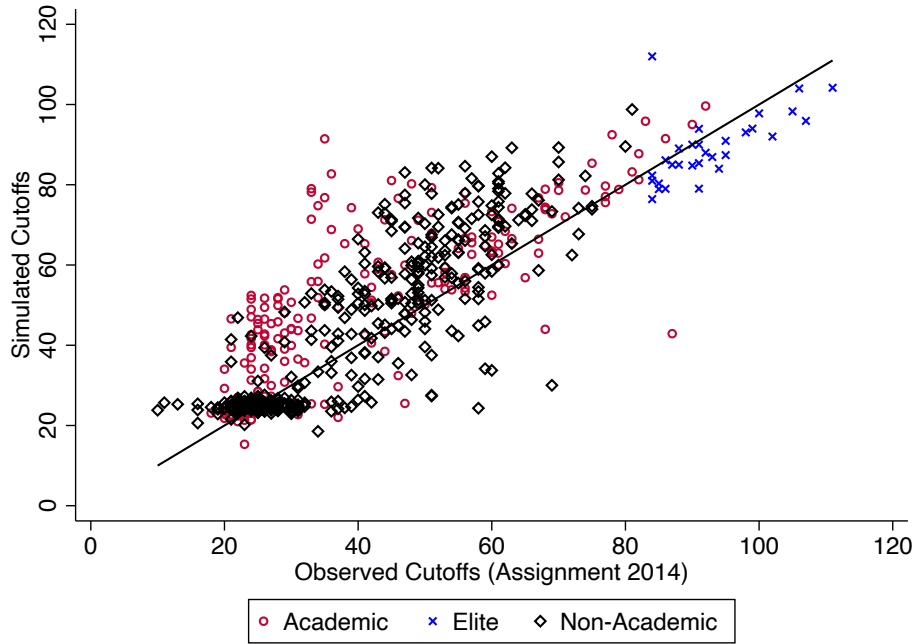
NOTE: The moments displayed in the first column are computed from the data of the assignment mechanism in the year 2014 (see Section 2). The moments displayed in the second column are computed by running the Serial Dictatorship algorithm that is in place for the COMIPEMS system, using the simulated rank-ordered lists from the estimates reported in the fourth column in Appendix Table B.5, the individual scores in the admission exam, and the school capacities as inputs.

We provide two pieces of evidence that lend support to this assumption. First, we compare average moments of assignment patterns from the data with those from the matching equilibrium computed using the simulated rank-ordered lists for all the applicants in the system.<sup>10</sup> Table 6 reports overall means across all applicants for selected assignment outcomes (columns 1 and 2). Mean-differences are very small in size throughout the set of outcomes considered (see column 3 in Table 6), which indicates that the estimated model parameters using the experimental control group appear to adequately approximate the overall assignment patterns observed in the data. Second, the linear correlation between the observed schools' cutoff scores and the model-based simulated cutoff scores (averaged over multiple replications of the matching equilibrium, each time with a different  $\epsilon$  draw of the preference shock) is 0.85. Figure 5 provides more nuanced evidence on the comparison in the cutoff scores between the model and the data. The extrapolated choice model using the experimental control group broadly fits the cutoff distribution in the data throughout the entire spectrum of schools in the system (i.e., non-academic and academic, more or less selective, and elite schools).

The estimated model parameters for the students in the experimental treatment group (i.e., those who were provided with information on their performance in the mock exam) approximate the counterfactual scenario in which *all* applicants would be given their test scores for the actual

<sup>10</sup>Section 4.4 provides a more detailed description on how we compute the matching equilibrium in our simulations.

Figure 5: Model Fit on Schools' Cutoff Scores



NOTE: The observed cutoffs are computed from the data of the assignment mechanism in the year 2014 (see Section 2). The (average) simulated cutoff scores displayed in the scatter plot are computed by running 50 times the Serial Dictatorship algorithm that is in place for the COMIPEMS system using the simulated rank-ordered lists (each time with a different  $\epsilon$  draw of the preference shock), the individual scores in the admission exam, and the school capacities as inputs.

assignment exam. Using the simulated rank-ordered lists, we can compute the new matching equilibrium in the centralized system under the policy counterfactual of interest (see Section 4.4). This approach overcomes the key challenge to scaling-up in our setting to the extent that the simulated assignment patterns effectively embed the equilibrium sorting and displacement effects resulting from shifts in the aggregate demand.<sup>11</sup>

#### 4.4 The Impact of Information Provision at Scale

While we estimate the school choice model (1) by assuming stable matching but not truth-telling, we can allow students to be truthful in order to study matching outcomes. This holds as long as preference estimates are consistent (Artemov et al. 2023). We use the simulated rank-ordered

<sup>11</sup>One potential caveat with our school choice framework is that it precludes applicants from reacting to (expected) policy-induced changes in peer composition across schools. While summary measures of peer quality can be included as school characteristics in the students' indirect utility (1), the endogenous determination of peer quality compromises the interpretation of the estimates for counterfactual or equilibrium calculations. In a setting of decentralized allocation of students across schools, Allende (2019) documents the importance of preferences for peers in the study of school competition, which is beyond the scope of this paper.

Table 7: The Effect of the Information Intervention on Assignment Outcomes

	Status Quo	Information Intervention	Difference
Applied in the system (1=yes)	0.99	0.99	0.00
Assigned in the system (1=yes)	0.89	0.91	0.02
Rank of assigned school	6.41	5.43	-0.98
Assigned in top choice	0.16	0.25	0.09
Assigned in elite schools	0.22	0.22	0.00
Assigned in academic schools	0.41	0.40	-0.01
Assigned in non-academic schools	0.37	0.38	0.01

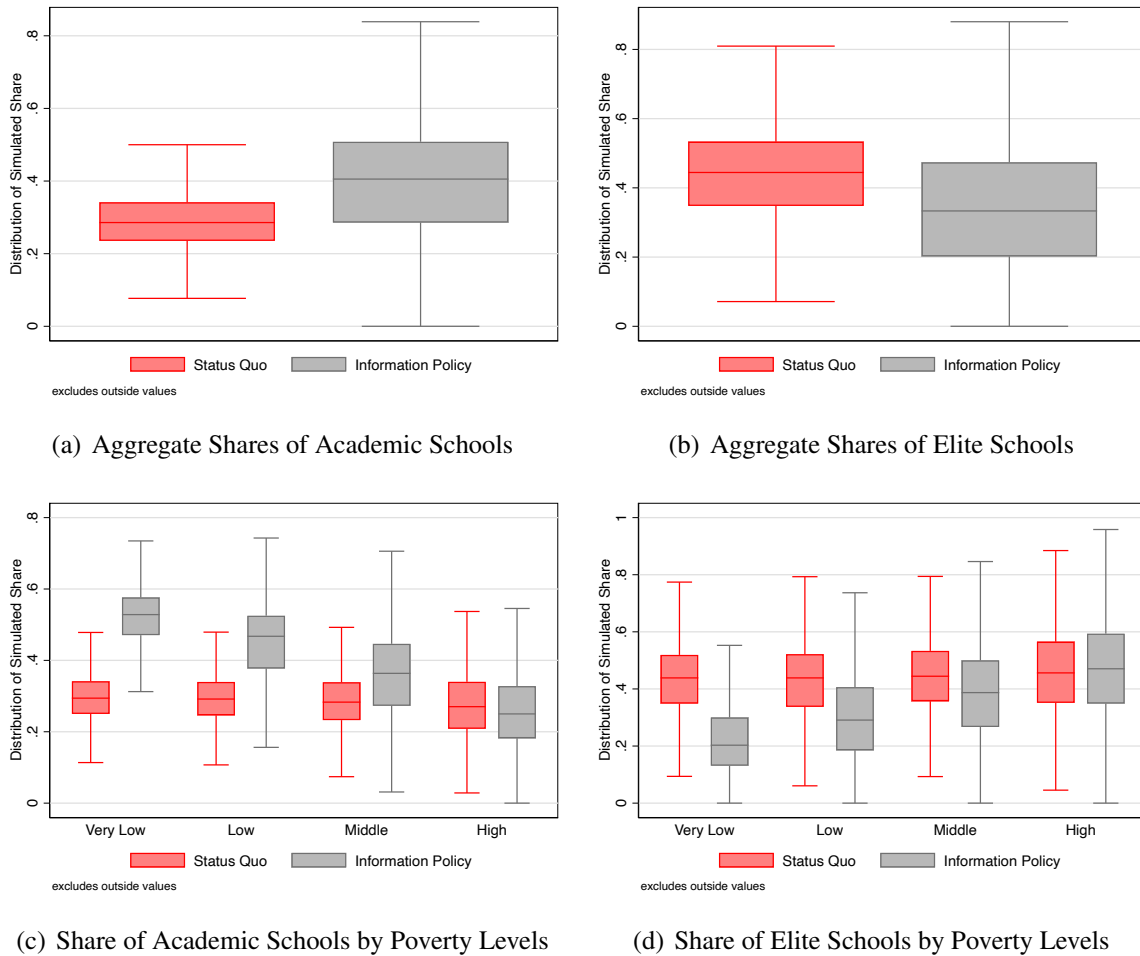
NOTE: The average moments displayed in the first column are obtained by applying the Serial Dictatorship algorithm to the simulated rank-ordered lists based on the model estimates for the control group. The average moments displayed in the second column are computed by running the Serial Dictatorship algorithm using the simulated rank-ordered lists based on the model estimates for the treatment group. The corresponding estimates are reported in the fourth and fifth columns of Table B.5 in the Appendix.

lists, the individual scores in the admission exam, and the school capacities from the assignment mechanism in the 2014 round as inputs to run the same Serial Dictatorship algorithm that is in place in the centralized assignment system.<sup>12</sup> Since school preferences are vertical, this algorithm delivers the unique stable matching equilibrium allocation (Roth & Sotomayor 1992).

Table 7 shows some features of the equilibrium allocation under the information intervention when compared to the status quo allocation. First, there are no changes in the participation rate in the admission process. With the additional information provided, some students may have preferred their outside options and hence opted out of the assignment system. This is not the case in our setting. Second, the share of assigned students through the Serial Dictatorship algorithm increases by 2 percentage points. Albeit marginal, an increase in the number of assigned students is an important result. Unmatched students participate in a second round of assignments where they can only choose between the relatively low-quality schooling alternatives that still have open seats after the first round (see Section 2.1). Third, and perhaps most importantly, students are more likely to get assigned to their preferred options. When moving from the status quo to the counterfactual policy, the average applicant is placed in a school that is one position above her school rankings (5.4, vs. 6.4). Accordingly, the share of students assigned to their most preferred option increases by nine percentage points, from 16 percent to 25 percent. Finally, the average share of applicants assigned across high school tracks does not systematically change with the equilibrium allocation under the information intervention at scale.

<sup>12</sup>The simulated rank-ordered lists included in the algorithm are those that give a higher utility to the applicants when compared to the outside option. In order to mimic the actual assignment mechanism, we truncate the length of the simulated rank-ordered lists to the institutional constraint of 20 and impose the GPA constraint (above 7/10) for the elite school admissions.

Figure 6: The Effect of the Information Intervention on School Choice

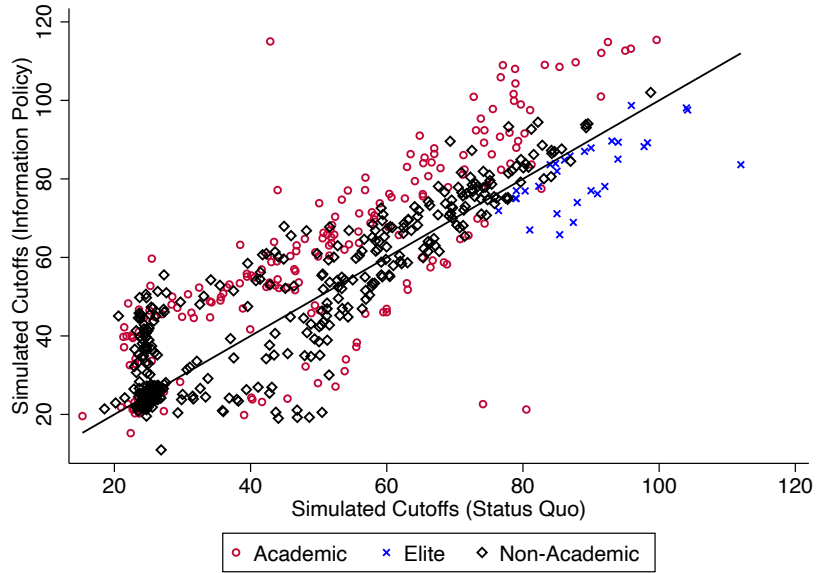


NOTE: This figure displays the empirical distributions of the shares of academic schools and elite schools in the applicants' simulated rank-ordered lists as implied by the model estimates for the control group (Status Quo) and for the treatment group (Information Policy). The corresponding estimates are reported in the fourth and fifth columns of Table B.5 in the Appendix. The central lines within each box denote the sample medians, whereas the upper and lower level contours of the boxes denote the 75th and 25th percentiles, respectively. The whiskers outside of the boxes denote the upper and lower adjacent values, which are values in the data that are furthest away from the median on either side of the box, but are still within a distance of 1.5 times the interquartile range from the nearest end of the box (i.e., the nearer quartile).

These aggregate sorting patterns may mask substantial heterogeneity in the effect of the intervention across different groups of applicants. In fact, it is informative to compare the distributions of school choices as predicted by the estimated model for the control group with the corresponding distributions predicted by the estimated model for the treatment group. Figure 6 displays box-and-whisker plots for the shares of academic and elite schools in the simulated rank-ordered lists. Panels A and B show the distributions across the whole applicants' pool in the Status Quo sce-



Figure 7: The Effect of the Information Intervention on Cutoff Scores



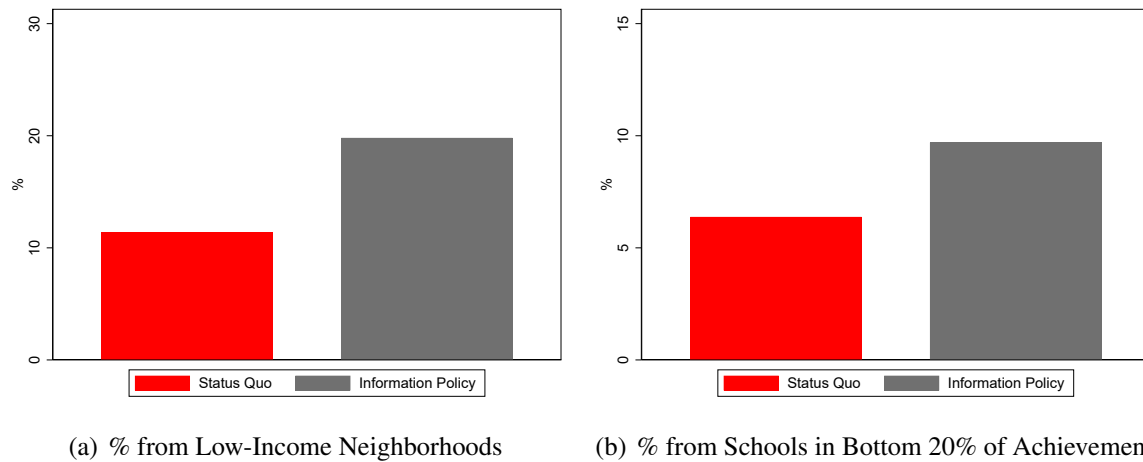
NOTE: The (average) simulated cutoff scores displayed in the x-axis are computed by running 50 times the Serial Dictatorship algorithm that is in place for the COMIPEMS system using the predicted school rankings (each time with a different  $\epsilon$  draw) based on the model estimates for the control group. The (average) simulated cutoff scores displayed in the y-axis are computed by running 50 times the Serial Dictatorship algorithm that is in place for the COMIPEMS system using the predicted school rankings (each time with a different  $\epsilon$  draw) based on the model estimates for the treatment group. The corresponding estimates are reported in the fourth and fifth columns of Table B.5 in the Appendix.

nario (red) versus the Information Policy scenario (gray). Overall, the gray bars show that the information intervention leads to an increase in the demand for academic schools and a symmetric decrease in demand for elite schools. The box plots further reveal that both the location and the scale of the choice distributions for academic schools increase (i.e., the inter-quantile range more than doubles). The same pattern holds for the scale of the choice distribution for elite schools under the information intervention.

Panels C and D display the same effects, but they are broken down by discrete levels of neighborhood poverty. The overall effect of the information intervention observed in the upper-panels can be mostly explained by the associated changes in the simulated rank-ordered lists for relatively better-off applicants (i.e., those who live in neighborhoods with very low, or low poverty levels).<sup>13</sup>

<sup>13</sup>This pattern can be traced back to the estimates of the school choice model, as shown in Table 5. The neighborhood poverty variable generates a large, positive, and statistically different average marginal effect on the choice probabilities of elite schools in the experimental treatment group when compared to the control group. Poverty also has a stronger negative average marginal effect on the choice probabilities of academic schools among applicants in the experimental treatment group, although the difference in the estimated coefficients with the control group is neither large nor statistically significant.

Figure 8: The Effect of the Information Intervention on the Composition of Students at Elite Schools



NOTE: This figure shows the simulated shares of applicants from low-income neighborhoods (Panel A) and from schools in the bottom 20 percent of math achievement in the national evaluation (ENLACE), that would result in being assigned to elite schools under both the Status Quo and the Information Policy scenarios.

The fact that socio-economically disadvantaged applicants are, on average, unresponsive to the intervention is consistent with the experimental evidence presented in Section 3.2.

The decreased demand-side pressure on elite programs is likely to spur some reallocation effects within the system, possibly towards (less selective) academic programs. This can be seen through changes in the simulated schools' cutoff scores, as depicted in Figure 7. Under the Information Policy scenario, the cutoff scores for most of the elite schools slightly decrease when compared to the Status Quo scenario. Instead, the cutoff scores of academic programs increase on average. Non-academic programs feature a more erratic pattern of positive and negative changes in their equilibrium cutoffs. Jointly, the movements in the cutoff scores and the underlying changes in the demand for academic schools and elite schools may explain the muted effect of the intervention on the average sorting patterns across high school tracks (see Table 7).

The composition changes in the demand for elite programs by neighborhood poverty levels may induce some differential sorting and displacement effects of the information intervention at scale. The reduction in the demand for elite programs among the relatively better-off applicants would necessarily leave some open seats in those programs for high-achieving and relatively disadvantaged students. In this scenario, the scaled-up intervention would alter the composition of the student population admitted to elite schools in favor of the relatively disadvantaged students. Figure 8 shows that this is precisely what happens in the (simulated) data. The information intervention doubles the representation of students from low-income neighborhoods in elite schools,

where the admission share starts at 13 percent in the Status Quo scenario to 22 percent in the Information Policy scenario (see Panel A). Similarly, Panel B in the same figure further shows that the share of applicants from the most disadvantaged middle schools (defined as those in the first quintile of the math-score distribution) increases from 7 to 10 percent under the Information Policy scenario.

The equilibrium effect of the information intervention on elite admission for low-income applicants could not possibly be detected in the small-scale evaluation. Indeed, the experimental evidence of Section 3.2 documents no effect of feedback provision on the probability of assignment to elite schools, as shown in the last column of Table 3. Our approach shows that this type of indirect effect that exclusively arises during large-scale policy implementations can be fully characterized, even with partial knowledge of the demand side, in a setting where supply-side responses are muted.

## 5 The Medium-run Consequences of Displacement

The simulation results presented in the previous section show that low-income and high-achieving applicants are disproportionately more likely to be admitted to elite schools under the large-scale implementation of the information intervention. This effect reflects a displacement effect of relatively better-off applicants, who decrease their demand for elite programs under this scenario. In this section, we attempt to assess the impact of such indirect (equilibrium) effect of the intervention on the educational outcomes of low-income applicants.

We consider on-time graduation from high school as our main outcome of interest. As shown in Section 3.3, this variable adequately captures the student-school match effects that are likely at play in our setting. Since the administrative records at our disposal do not cover UNAM-sponsored high schools, we focus on the effects of admission to IPN schools among the elite schools (see Section 2.1). We apply the same definition of socio-economic disadvantage (low-income) that we have used throughout the analysis. Namely, we focus on applicants who reside in neighborhoods with high or very high poverty levels. Differing from the analysis discussed in Section 3.3, we treat admission as equal to enrolment, because enrolment at elite schools is almost universal (97%).

### 5.1 Regression Discontinuity Design

All elite schools are over-subscribed, and admission requires clearing their admission cutoffs (see Figure 1). These cutoffs can be exploited to identify the effect of marginal admission to an elite school on the probability of on-time graduation from upper secondary education. We focus on

approximately 10 percent of the low-income applicants (N=18,011), whose score barely places them in an elite school (instead of a non-elite school), as well as those who just miss elite placement and end up in a non-elite school. Since we consider only two groups of institutions, elite and non-elite, we have students whose first best is an elite school and their second best is a non-elite school in the local institution ranking (Kirkeboen et al. 2016). Table B.6 in the Appendix illustrates the intuition behind our sample selection, which implies that applicants who are at the margin of elite admission have the same ordinal preferences.

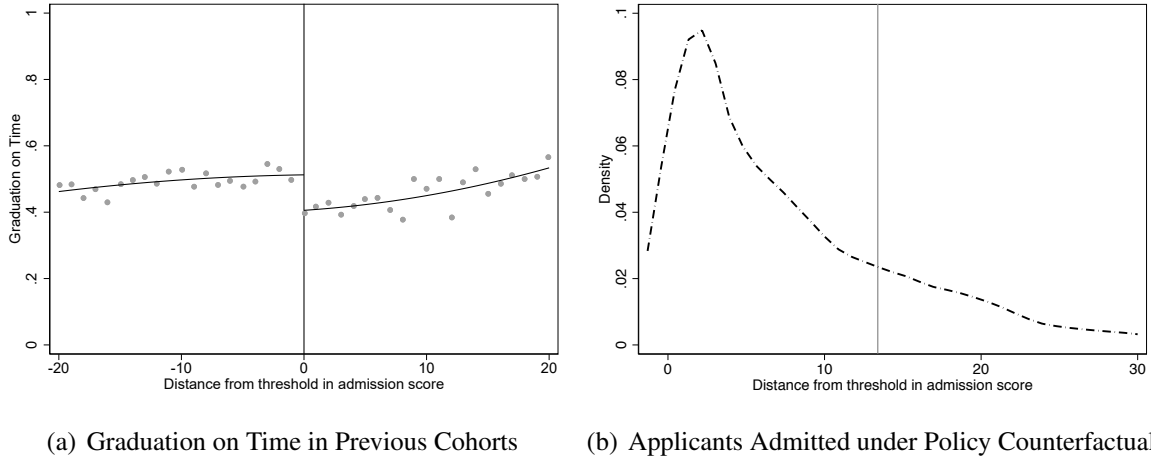
We use the following Regression Discontinuity (RD) regression model for students with a score close to the relevant cutoff scores:

$$y_{ij} = \eta + \theta Elite_i + \delta(s_i - c_{k(i)}) + \tau(s_i - c_{k(i)}) \times Elite_i + \nu_{ij}, \quad (2)$$

where  $y_{ij}$  is an indicator variable that denotes whether student  $i$  in high school  $j$  graduates on time. We center the running variable  $s_i$  by the IPN-specific admission cutoffs,  $c_{k(i)}$ , such that a positive value of  $(s_i - c_{k(i)})$  indicates admission to an elite school. In our estimation sample, each student has a minimum cut-off for elite admission,  $c_{k(i)}$ , that depends on her preferences. The indicator variable  $Elite_i$  takes the value of one when a student is admitted to an elite school, and zero otherwise, and  $\nu_{ij}$  is an error term that is clustered at the school  $\times$  year level. Given the continuity of potential outcomes in terms of the admission cutoffs, the parameter  $\theta$  would capture the effect of marginal admission to an elite school on high school graduation on time. We estimate this non-parametrically by local linear regressions that are defined within the mean squared error (MSE) optimal bandwidths (Calonico et al. 2014).

Manipulation of the running variable in terms of the RD threshold is unlikely in our context, for two reasons. First, admission cutoffs 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 manipulation was to occur, we would expect to observe bunching in the empirical density of the admission score just above the admission cutoffs of elite schools. Figure B.2 in the Appendix confirms that there is no evidence of manipulation in our sample ( $T=-1.0$ ,  $p$ -value=0.31). Accordingly, Appendix Table B.7 documents that the pre-determined individual-level covariates are smooth in relation to the RD threshold, with point estimates that are small relative to the respective means in the local sample and that are not statistically different from zero.

Figure 9: Threshold-crossing and Marginally Admitted Low-income Applicants



NOTE: Panel A of this figure shows the binned means of the outcome variable in terms of the elite admission thresholds, defined by following the IMSE-optimal evenly spaced method by Calonico et al. (2015). Solid lines represent the predictions from local polynomial (quadratic) regressions estimated separately for observations to the left and to the right of the admission threshold. Panel B depicts the empirical density for the sub-set of low-income applicants who would be admitted to elite (IPN) schools under the large-scale implementation of the information intervention (see Panel A of Figure 8). The vertical line in Panel B represents the MSE-optimal bandwidth (see Table 8).

## 5.2 The Causal Effect of Elite School Admission for Low-income Students

We start by showing graphical evidence of marginal admission to an elite school of on-time graduation from high school. Panel A of Figure 9 documents clear evidence that low-income applicants who are marginally admitted to an elite school experience relatively worse rates of on-time high school graduation when compared to those who are marginally rejected. Away from the cutoff, the observed positive correlation between the rate of on-time graduation and the running variable is consistent with the notion that the score in the admission exam captures relevant academic skills.

Panel B of Figure 9 depicts the empirical density of the comparable running variable under the large-scale implementation of the information intervention studied in Section 4 for the sub-set of low-income applicants who would be admitted to elite (IPN) schools. The chart shows that the majority of these applicants would indeed be marginally admitted to elite schools. More than 80 percent of these applicants are within the optimal bandwidth of 13.4 centered-score points, as shown by the vertical bar in Panel B of Figure 9, and the median lies at a 5-point difference between the admission score and the relevant minimum cutoff score for elite admission.

Table 8 reports the corresponding non-parametric estimates of the RD regression model (2). The magnitude and sign of the estimated  $\theta$  parameter are remarkably consistent between the different specifications displayed across the various columns of the table and the graphical evidence

Table 8: Threshold Crossing Effect on Graduation on Time

	Baseline	Cohort-FE	Quadratic	Half BW	Double BW
Elite Admission	-0.112 (0.021)	-0.112 (0.020)	-0.123 (0.030)	-0.117 (0.028)	-0.119 (0.017)
Bias-corrected CI	[-0.157;-0.061]	[-0.156;-0.061]	[-0.196;-0.064]		
Optimal BW	13.368	13.575	12.962		
Ad Hoc BW				6.684	26.737
Mean Dep. Var.	0.463	0.463	0.463	0.466	0.455
Nb. of Obs.	10828	10828	10226	5984	16021

NOTE: This table displays non-parametric estimates of marginal admission to elite schools of on-time graduation for a stacked sample across five consecutive cohorts (2005-2009) of applicants who reside in low-income neighborhoods. Standard errors reported in parentheses are clustered at the school $\times$ year level. The first column displays a baseline specification using local-linear regressions and a triangular kernel. In the second column, we include indicator variables for the different years of the assignment mechanism as additional covariates. In the third column, we consider local polynomial (quadratic) functions of the centered admission score. Finally, the last two columns show local linear regression models that use ad-hoc bandwidth choices, respectively half and double the MSE-optimal bandwidth of the baseline model. For the specifications that are defined within optimal bandwidths, we also report the bias-corrected confidence intervals obtained using the robust estimator proposed in [Calonico et al. \(2014\)](#).

reported in Panel A of Figure 9. Marginal admission to an elite school decreases the rate at which low-income students graduate on time from high school by 11-12 percentage points. The magnitude of this local effect is substantial given that, on average, only 46 percent of the marginally rejected students from elite schools tend to graduate on time from high school.

The evidence presented in this section suggests that it is not clear whether low-income applicants would in fact benefit from the equilibrium effect that arises under the large-scale implementation of the information intervention. Marginally admitted students at elite schools are likely to experience negative consequences on their subsequent academic trajectories, possibly through grade retention, switching across high schools, or dropping-out from upper secondary education. While there may be other channels through which low-income applicants may take advantage of elite admission, such as peer effects or social networks that may determine future educational or employment opportunities, comparable evidence from Chile shows that these gains are muted for students outside of historically advantaged groups ([Zimmerman 2019](#)).

The sub-sample of low-income applicants who would gain admission into elite schools under the counterfactual policy simulation is largely comparable to the marginally admitted applicants of the RD analysis presented in this section. Taken together, these findings suggest that the sorting and displacement effects of the intervention are likely to offset, at least in part, the positive average impact of the small-scale intervention on on-time graduation, as discussed in Section 3.3.

## 6 Conclusion

The aim of this paper is to characterize the effect of an intervention in education, while taking into account the equilibrium effects that often arise during a large-scale implementation (see, e.g., [Agostinelli, Luflade & Martellini 2021](#), [Heckman et al. 1998a,b](#), [Khanna 2023](#)). We study a randomized experiment designed and implemented within a centralized school assignment mechanism in Mexico City. The treatment consists of providing ninth-graders with timely performance feedback regarding their academic skills through the application of a mock version of the admission exam used to determine priorities within the assignment mechanism. The experimental evidence documents that, relative to a control group, applicants in the treatment group are placed in better school-student matches, which yields higher rates of on-time graduation by the end of the twelfth grade.

While these findings point toward a positive impact of the information intervention in our setting, it is not clear whether the results are also informative when the same intervention is implemented at scale. The key challenge to scaling-up in our setting is because providing information about individual scores in the admission exam to all applicants would necessarily trigger aggregate sorting and displacement effects within the centralized mechanism, and this would then potentially alter individual placement outcomes. We embed an empirical model of school choice in a matching equilibrium that is consistent with the centralized assignment mechanism in order to overcome this challenge and quantify the effect of the intervention at scale. The modelling framework leverages the variation induced by the experiment in terms of the differential valuations over the schooling alternatives between treated and untreated applicants. We do so in order to extrapolate the counterfactual effects of the information intervention toward a larger and more diverse population of applicants.

Comparing simulated assignment outcomes between the status quo and the policy counterfactual reveals a positive average impact of the information intervention on student welfare. We also document substantial heterogeneity in the school choice responses across the diverse spectrum of applicants, which unlocks equilibrium effects within the centralized system that ultimately hinder the educational trajectories of socio-economically disadvantaged and high-achieving applicants. We conclude that the information intervention has quantitatively sizable equilibrium effects that only occur at scale, and that may partly offset its effectiveness for the sub-population of applicants targeted by the small-scale randomized evaluation.

This paper contributes to the current debate about the challenges to scale-up experimental evaluations. Some authors argue that one way in which experiments can be made more representative and hence have a greater impact on policy is by means of at-scale implementations ([Muralidharan](#)



& Niehaus 2017). Others propose an iterative process of sequential experimentation so as to gradually build external validity from small-scale trials to larger-scale evaluations (Banerjee et al. 2017). We argue that solely relying on randomized control trials is unlikely to generate real progress in the science of scaling for at least three reasons. First, it takes time and significant resources to implement at-scale evaluations in many applications, which then makes it difficult to provide quantitative policy advice in a timely way. Second, shifting the behavior of a large number of individuals often spurs nontrivial equilibrium/spillover effects. To the extent that these market-level changes systematically influence individual outcomes, it is difficult to think about experimental designs that are robust to possible violations of the stable unit treatment value assumption, or SUTVA (Imbens & Rubin 2015).<sup>14</sup> Lastly, as we show in our analysis, these indirect effects can have negative consequences on beneficiaries, and so quantifying them through ex-ante counterfactual model simulations seems a sensible (and ethical) alternative to ex-post program evaluation.

Another related strand of the literature attempts to bridge small-scale experimental work and structural macro analysis in the study of the aggregate and distributional effect of development policies (Bergquist et al. 2022, Buera et al. 2021, 2023). The authors advocate for a blend of models and data as a “middle ground”, in order to make progress on the complex issues that arise in scaling-up such policies. The mixed empirical approach pursued in this paper fits well into this characterization, which we hope will become a blueprint for future work that intends to study the impact of large-scale interventions.

---

<sup>14</sup>In contexts, different from ours, whereby equilibrium effects can be assumed spatially concentrated, or “local” in nature, an obvious solution is to enlarge the unit of randomization in order to account for those in the experimental analysis (Egger et al. 2022, Muralidharan et al. 2023).

## References

- Abdulkadiroglu, A., Agarwal, N. & Pathak, P. A. (2017), 'The Welfare Effects of Coordinated Assignment: Evidence from the New York City High School Match', *American Economic Review* **107**(12), 3635–3689.
- Abdulkadiroglu, A., Pathak, P. A., Schellenberg, J. & Walters, C. R. (2020), 'Do Parents Value School Effectiveness?', *American Economic Review* **110**(5), 1502–1539.
- Agarwal, N. & Somaini, P. (2020), 'Revealed preference analysis of school choice models', *Annual Review of Economics* **12**(1), 471–501.
- Agostinelli, F., Avitable, C. & Bobba, M. (2021), Enhancing Human Capital at Scale, Working Papers 2021-012, Human Capital and Economic Opportunity Working Group.
- Agostinelli, F., Luflade, M. & Martellini, P. (2021), On the Spatial Determinants of Educational Access, Working Papers 2021-042, Human Capital and Economic Opportunity Working Group.
- Akyol, P., Krishna, K. & Wang, J. (2018), Taking PISA seriously: How Accurate are Low-Stakes Exams? NBER Working paper No. 24930.
- Al-Ubaydli, O., List, J. A. & Suskind, D. (2020), '2017 klein lecture: The science of using science: Toward an understanding of the threats to scalability', *International Economic Review* **61**(4), 1387–1409.
- Allcott, H. (2015), 'Site Selection Bias in Program Evaluation', *The Quarterly Journal of Economics* **130**(3), 1117–1165.
- Allende, C. (2019), Competition under social interactions and the design of education policies, Technical report, Job Market Paper.
- Anderson, M. L. (2008), 'Multiple inference and gender differences in the effects of early intervention: A reevaluation of the abecedarian, perry preschool, and early training projects', *Journal of the American Statistical Association* **103**(484), 1481–1495.
- Andrabi, T., Das, J. & Khwaja, I. (2017), 'Report Cards: The Impact of Providing School and Child Test Scores on Educational Markets', *American Economic Review* **107**(6), 1535–1563.
- Arcidiacono, P., Aucejo, E. M. & Hotz, V. J. (2016), 'University Differences in the Graduation of Minorities in STEM Fields: Evidence from California', *American Economic Review* **106**(3), 525–562.
- Artemov, G., Che, Y.-K. & He, Y. (2023), 'Stable matching with mistaken agents', *Journal of Political Economy Microeconomics* (Forthcoming).
- Avery, C. & Hoxby, C. (2012), The Missing "One-Offs": The Hidden Supply of High Achieving, Low Income Students. NBER Working paper No. 18586.

- Azmat, G., Bagues, M., Cabrales, A. & Iriberry, N. (2019), ‘What You Don’t Know Can’t Hurt You? A Natural Field Experiment on Relative Performance Feedback in Higher Education’, *Management Science* **65**(8), 3714–3736.
- Banerjee, A. V., Banerji, R., Berry, J., Duflo, E., Kannan, H., Mukerji, S., Shotland, M. & Walton, M. (2017), ‘From proof of concept to scalable policies: Challenges and solutions, with an application’, *Journal of Economic Perspectives* **31**(4), 73–102.
- Bergman, P. (2021), ‘Parent-child information frictions and human capital investment: Evidence from a field experiment’, *Journal of Political Economy* **129**(1), 286–322.
- Bergquist, L. F., Faber, B., Fally, T., Hoelzlein, M., Miguel, E. & Rodriguez-Clare, A. (2022), Scaling agricultural policy interventions, Working Paper 30704, National Bureau of Economic Research.
- Bobba, M. & Frisancho, V. (2016), Perceived Ability and School Choices, TSE Working Papers 16-660, Toulouse School of Economics (TSE).
- Bobba, M. & Frisancho, V. (2022), ‘Self-perceptions about academic achievement: Evidence from Mexico City’, *Journal of Econometrics* **231**(1), 58–73.
- Buera, F. J., Kaboski, J. P. & Shin, Y. (2021), ‘The Macroeconomics of Microfinance’, *Review of Economic Studies* **88**(1), 126–161.
- Buera, F. J., Kaboski, J. P. & Townsend, R. M. (2023), ‘From Micro to Macro Development’, *Journal of Economic Literature* (forthcoming).
- Burks, S. V., Carpenter, J. P., Goette, L. & Rustichini, A. (2013), ‘Overconfidence and social signalling’, *The Review of Economic Studies* **80**(3), 949–983.
- Calonico, S., Cattaneo, M. D. & Titiunik, R. (2014), ‘Robust nonparametric confidence intervals for regression discontinuity designs’, *Econometrica* **82**, 2295–2326.
- Calonico, S., Cattaneo, M. D. & Titiunik, R. (2015), ‘Optimal data-driven regression discontinuity plots’, *Journal of the American Statistical Association* **110**(512), 1753–1769.
- Calsamiglia, C., Haeringer, G. & Klijn, F. (2010), ‘Constrained School Choice: An Experimental Study’, *American Economic Review* **100**(4), 1860–1874.
- Caron, E., Bernard, K. & Metz, A. (2021), Fidelity and properties of the situation, challenges and recommendations, in ‘The Scale-up Effect in Early Childhood and Public Policy. Edited by John A. List, Dana Suskind, and Lauren H. Supplee’, Routledge.
- Cattaneo, M. D., Jansson, M. & Ma, X. (2020), ‘Simple local polynomial density estimators’, *Journal of the American Statistical Association* **115**(531), 1449–1455.
- Davis, J., Guryan, J., Hallberg, K. & Ludwig, J. (2021), Studying properties of the population: Designing studies that mirror real world scenarios, in ‘The Scale-up Effect in Early Childhood and Public Policy. Edited by John A. List, Dana Suskind, and Lauren H. Supplee’, Routledge.

- Delavande, A., Giné, X. & McKenzie, D. (2011), ‘Eliciting probabilistic expectations with visual aids in developing countries: how sensitive are answers to variations in elicitation design?’, *Journal of Applied Econometrics* **26**(3), 479–497.
- Dizon-Ross, R. (2019), ‘Parents beliefs about their children academic ability: Implications for educational investments’, *American Economic Review* **109**(8), 2728–65.
- Duflo, E. (2017), ‘The economist as plumber’, *American Economic Review* **107**(5), 1–26.
- Duflo, E. (2020), ‘Field experiments and the practice of policy’, *American Economic Review* **110**(7), 1952–73.
- Dustan, A. (2020), ‘Can large, untargeted conditional cash transfers increase urban high school graduation rates? Evidence from Mexico City’s Prepa SÃ’, *Journal of Development Economics* **143**(C).
- Dustan, A., de Janvry, A. & 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.
- Egger, D., Haushofer, J., Miguel, E., Niehaus, P. & Walker, M. W. (2022), ‘General equilibrium effects of cash transfers: Experimental evidence from kenya’, *Econometrica* **90**(6), 2603–2643.
- Estrada, R. & Gignoux, J. (2017), ‘Benefits to elite schools and the expected returns to education: Evidence from mexico city’, *European Economic Review* **95**, 168–194.
- Ewers, M. & Zimmermann, F. (2015), ‘Image And Misreporting’, *Journal of the European Economic Association* **13**(2), 363–380.
- Fack, G., Grenet, J. & He, Y. (2019), ‘Beyond Truth-Telling: Preference Estimation with Centralized School Choice and College Admissions’, *American Economic Review* **109**(4), 1486–1529.
- Galiani, S. & Pantano, J. (2021), Structural Models: Inception and Frontier, NBER Working Papers 28698, National Bureau of Economic Research, Inc.
- Haaland, I. K., Roth, C. & Wohlfart, J. (2023), ‘Designing Information Provision Experiments’, *Journal of Economic Literature* (forthcoming).
- Haeringer, G. & Klijn, F. (2009), ‘Constrained School Choice’, *Journal of Economic Theory* **144**(5), 1921–1947.
- Hastings, J. S. & Weinstein, J. M. (2008), ‘Information, School Choice, and Academic Achievement: Evidence from Two Experiments’, *The Quarterly Journal of Economics* **123**(4), 1373–1414.
- Heckman, J. J., Lochner, L. & Taber, C. (1998a), ‘General-Equilibrium Treatment Effects: A Study of Tuition Policy’, *American Economic Review* **88**(2), 381–386.
- Heckman, J., Lochner, L. & Taber, C. (1998b), ‘Explaining Rising Wage Inequality: Explanations With A Dynamic General Equilibrium Model of Labor Earnings With Heterogeneous Agents’, *Review of Economic Dynamics* **1**(1), 1–58.

- Imbens, G. W. & Rubin, D. B. (2015), *Causal Inference for Statistics, Social, and Biomedical Sciences*, number 9780521885881 in 'Cambridge Books', Cambridge University Press.
- Jensen, R. (2010), 'The (Perceived) Returns to Education and The Demand for Schooling', *Quarterly Journal of Economics* **125**(2), 515–548.
- Khanna, G. (2023), 'Large-Scale Education Reform in General Equilibrium: Regression Discontinuity Evidence from India', *Journal of Political Economy* **131**(2), 549–591.
- Kirkeboen, L. J., Leuven, E. & Mogstad, M. (2016), 'Editor's Choice Field of Study, Earnings, and Self-Selection', *The Quarterly Journal of Economics* **131**(3), 1057–1111.
- Lavecchia, A., Liu, H. & Oreopoulos, P. (2016), Chapter 1 - behavioral economics of education: Progress and possibilities, Vol. 5 of *Handbook of the Economics of Education*, Elsevier, pp. 1–74.
- Lee, D. S. (2009), 'Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects', *The Review of Economic Studies* **76**(3), 1071–1102.
- List, J. A. (2022), *The Voltage Effect: How to Make Good Ideas Great and Great Ideas Scale*, Penguin Books.
- List, J. A., Shaikh, A. M. & Xu, Y. (2019), 'Multiple Hypothesis Testing in Experimental Economics', *Experimental Economics* **22**(4), 773–793.
- Low, H. & Meghir, C. (2017), 'The use of structural models in econometrics', *Journal of Economic Perspectives* **31**(2), 33–58.
- Muralidharan, K. & Niehaus, P. (2017), 'Experimentation at Scale', *Journal of Economic Perspectives* **31**(4), 103–124.
- Muralidharan, K., Niehaus, P. & Sukhtankar, S. (2023), 'General equilibrium effects of (improving) public employment programs', *Econometrica* (Forthcoming).
- Neilson, C., Allende, C. & Gallego, F. (2019), Approximating the Equilibrium Effects of Informed School Choice, Working Papers 628, Princeton University, Department of Economics, Industrial Relations Section.
- O'Brien, P. C. (1984), 'Procedures for comparing samples with multiple endpoints', *Biometrics* **40**(4), 1079–1087.
- Pariguana, M. & Ortega, M. E. (2022), School Admissions, Mismatch, and Graduation, PhD thesis, University of Western Ontario.
- Romano, J. P. & Wolf, M. (2005a), 'Exact and Approximate Stepdown Methods for Multiple Hypothesis Testing', *Journal of the American Statistical Association* **100**, 94–108.
- Romano, J. P. & Wolf, M. (2005b), 'Stepwise Multiple Testing as Formalized Data Snooping', *Econometrica* **73**(4), 1237–1282.

- Romano, J. P. & Wolf, M. (2016), 'Efficient Computation of Adjusted p-Values for Resampling-Based Stepdown Multiple Testing', *Statistics & Probability Letters* **113**(C), 38–40.
- Roth, A. E. & Sotomayor, M. (1992), Two-sided matching, in R. Aumann & S. Hart, eds, 'Handbook of Game Theory with Economic Applications', Vol. 1 of *Handbook of Game Theory with Economic Applications*, Elsevier, chapter 16, pp. 485–541.
- Shapiro, S. S. & Wilk, M. B. (1965), 'An analysis of variance test for normality (complete samples)', *Biometrika* **52**(3-4), 591–611.
- Todd, P. & Wolpin, K. (2023), 'The best of both worlds: Combining rcts with structural modeling', *Journal of Economic Literature* (forthcoming).
- Wiswall, M. & Zafar, B. (2015), 'Determinants of College Major Choice: Identification using an Information Experiment', *Review of Economic Studies* **82**(2), 791–824.
- Zhao, A. & Ding, P. (2021), 'To adjust or not to adjust? estimating the average treatment effect in randomized experiments with missing covariates'.
- Zimmerman, S. D. (2019), 'Elite Colleges and Upward Mobility to Top Jobs and Top Incomes', *American Economic Review* **109**(1), 1–47.

# Appendices

## A Beliefs Data

We collect rich survey data with detailed information on the subjective distribution of beliefs about performance in the admission exam. In order to help students understand probabilistic concepts, we explicitly linked the number of beans placed in a cup to a probability measure, where zero beans means that the student assigns zero probability to a given event and 20 beans means that the student believes the event will occur with certainty. Students were provided with a card divided into six discrete intervals of the score. Surveyors then elicited students' expected performance in the test by asking them to allocate the 20 beans across the intervals so as to represent the chances of scoring in each bin.

We include a set of practice questions before eliciting beliefs:

1. How sure are you that you are going to see one or more movies tomorrow?
2. How sure are you that you are going to see one or more movies in the next two weeks?
3. How sure are you that you are going to travel to Africa next month?
4. How sure are you that you are going to eat at least one *tortilla* next week?

If respondents grasp the intuition behind our approach, they should provide an answer for question 2 that is larger than or equal to the answer in question 1, since the latter event is nested in the former. Similarly, respondents should report fewer beans in question 3 (close to zero probability event) than in question 4 (close to one probability event). Whenever students made mistakes, the surveyor repeated the explanation as many times as necessary before moving forward. We are confident that the elicitation of beliefs has worked well since only 11 students (0.3%) ended up making mistakes in these practice questions. The survey question eliciting beliefs reads as follows (authors' translation from Spanish):

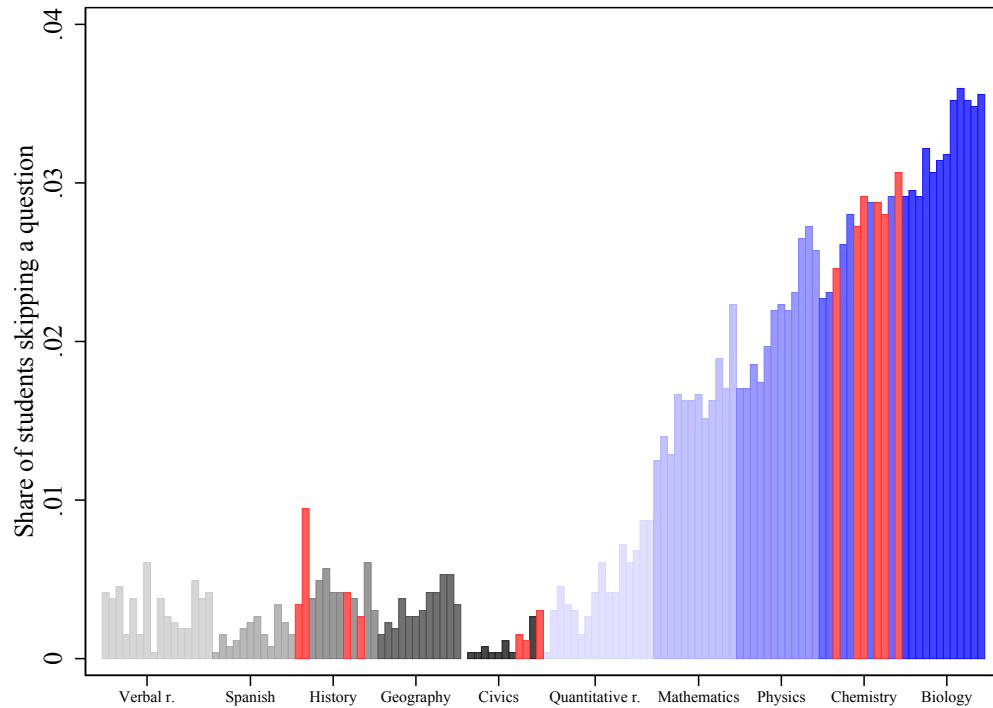
“Suppose that you were to take the COMIPEMS exam today, which has a maximum possible score of 128 and a minimum possible score of zero. How sure are you that your score would be between ... and ...”

During the pilot activities, we tested different versions with less bins and/or fewer beans to evaluate the trade-off between coarseness of the grid and students' ability to distribute beans across all intervals. We settled for six intervals with 20 beans as students were at ease with that format. Only 6% of the respondents concentrate all beans in one interval, which suggests that the grid was too coarse only for a few applicants. The resulting individual ability distributions seem well-behaved: using the 20 observations (i.e., beans) per student, we run a normality test ([Shapiro & Wilk 1965](#)) and reject it for only 11.4% of the respondents.



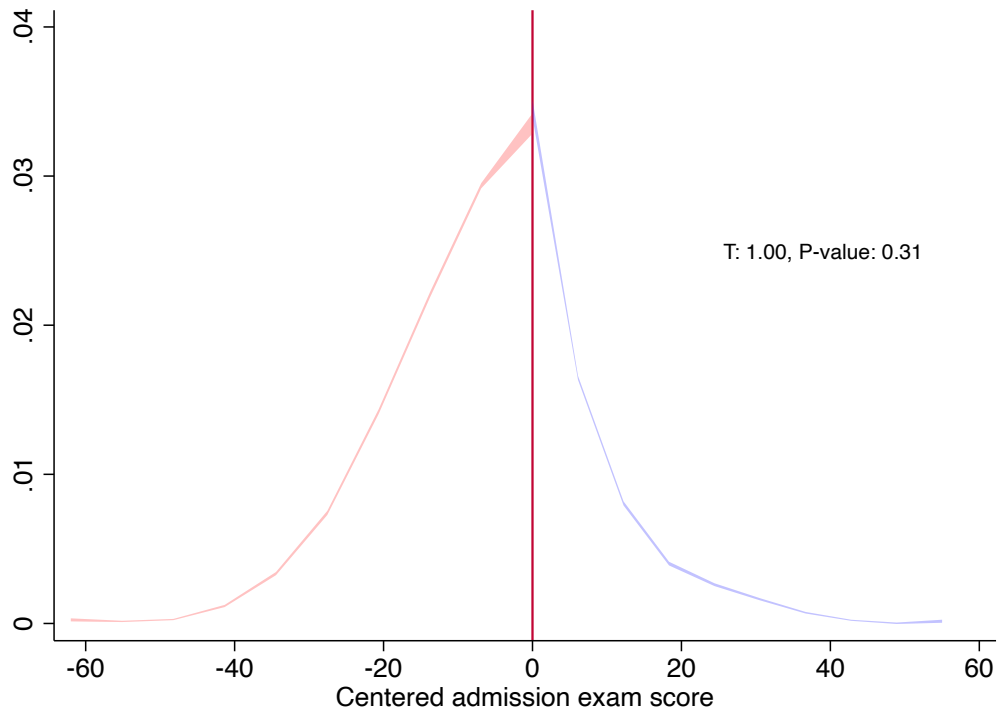
## B Additional Figures and Tables

Figure B.1: Average Skipping Patterns in the Mock Exam



Note: The x-axis orders the 128 questions of the exam in order of appearance. Different colors are used to group together questions from the same section in the exam. Questions in red are the ones excluded from grading since the school curriculum did not cover those subjects by the time of the application of the mock exam.

Figure B.2: Manipulation of the Elite Admission Cutoff



Note: The figure displays the empirical densities with the corresponding confidence intervals for the admission score running variable. The density is computed separately before and after the admission cutoffs of elite schools using the local-polynomial estimator proposed in [Cattaneo et al. \(2020\)](#).

Table B.1: Summary Statistics and Randomization Check

	Control Group	Treatment Group	Treatment-Control
Mock exam score	60.540 (15.416)	62.366 (16.290)	1.496 [1.065]
Exam score	65.541 (19.516)	65.248 (19.284)	-0.169 [1.248]
GPA (middle school)	8.116 (0.846)	8.122 (0.846)	-0.013 [0.047]
Scholarship in MS	0.106 (0.308)	0.115 (0.319)	0.007 [0.015]
Grade retention in MS	0.125 (0.331)	0.131 (0.338)	0.010 [0.019]
Does not skip classes	0.971 (0.169)	0.971 (0.169)	-0.001 [0.010]
Plans to go to college	0.670 (0.470)	0.671 (0.470)	-0.003 [0.022]
Male	0.444 (0.497)	0.461 (0.499)	0.016 [0.020]
Disabled student	0.142 (0.349)	0.148 (0.355)	0.006 [0.015]
Indigenous student	0.085 (0.278)	0.101 (0.302)	0.017 [0.014]
Does not give up	0.878 (0.327)	0.889 (0.315)	0.015 [0.014]
Tries his best	0.735 (0.442)	0.722 (0.448)	-0.016 [0.021]
Finishes what he starts	0.720 (0.449)	0.712 (0.453)	-0.015 [0.020]
Works hard	0.725 (0.447)	0.739 (0.439)	0.010 [0.022]
Experienced bullying	0.142 (0.349)	0.152 (0.359)	0.010 [0.013]
Parental background and supervision	0.032 (0.786)	0.058 (0.760)	0.011 [0.035]
High SES (asset index)	0.463 (0.499)	0.485 (0.500)	0.019 [0.025]
Took prep courses	0.488 (0.500)	0.467 (0.499)	-0.026 [0.026]
Exam Preparation	0.421 (0.494)	0.443 (0.497)	0.027 [0.033]
Previous mock exam	0.269 (0.444)	0.290 (0.454)	0.017 [0.037]
Previous mock exam with feedback	0.133 (0.340)	0.166 (0.372)	0.028 [0.033]
N. Obs.	1290	1203	2493

NOTE: The first two columns report means and standard deviations (in parenthesis). The last column displays the OLS coefficients of the treatment dummy along with the standard errors clustered at the middle school level (in brackets) for the null hypothesis of zero effect.

Table B.2: Belief Updating

	Mean	Median	IQR
Treatment	-6.519	-8.449	-2.898
	[0.000]	[0.000]	[0.000]
	{0.001}	{0.001}	{0.001}
Mean Control	75.6	78.8	24.2
Number of Observations	2246	2246	2246
Number of Clusters	90	90	90
R-squared	0.156	0.156	0.044

NOTE: The dependent variable “Mean Beliefs” is constructed as the summation of the mid-values in each discrete interval of the score multiplied by the associated probability assigned by the student. The dependent variable “Median Beliefs” is defined as the midpoint of the interval in which the cumulative density of beans first surpasses 0.5 (11 beans or more). The dependent variable “Inter-Quantile Range (IQR) of Beliefs” is defined as the difference between the midpoints of the intervals that accumulate 75 percent and 25 percent of the probability mass. All specifications include a set of dummy variables which corresponds to the randomization strata, pre-determined characteristics (sex, characteristics of the school of origin, previous experience with practice exams providing feedback, aspirations to attend college, an index of personality traits, an index of parental characteristics, and a household asset index), and indicator variables for whether each of the covariates has missing data.  $p$ -values reported in brackets refer to the conventional asymptotic standard errors while those reported in curly brackets are adjusted for testing each null hypothesis across multiple outcomes through the step-wise procedure described in [Romano & Wolf \(2005a,b, 2016\)](#). All inference procedures take into account clustering of the error terms at the middle school level and the block randomization design.

Table B.3: Treatment Effects on Application Outcomes

	Participates COMIPEMS	Exam Score	Length of ROL	Max cutoff in ROL	Min cutoff in ROL
Treatment	0.000 [0.987] {0.983}	-0.669 [0.348] {0.908}	0.126 [0.564] {0.982}	1.641 [0.247] {0.777}	-0.366 [0.637] {0.982}
Performance index	0.023 [0.000] {0.001}	16.147 [0.000] {0.001}	0.079 [0.475] {0.978}	4.019 [0.000] {0.001}	4.368 [0.000] {0.001}
Treatment X Performance index	-0.002 [0.777] {0.982}	0.223 [0.582] {0.982}	-0.108 [0.489] {0.978}	0.262 [0.757] {0.982}	0.483 [0.510] {0.978}
Mean Control	0.881	65.541	9.465	90.491	35.022
Number of Observations	3160	2493	2493	2493	2493
Number of Clusters	90	90	90	90	90
R-squared	0.609	0.735	0.032	0.266	0.243

NOTE: Standard errors clustered at the middle school level. All specifications include a set of dummy variables which corresponds to the randomization strata, pre-determined characteristics (sex, characteristics of the school of origin, previous experience with practice exams providing feedback, aspirations to attend college, an index of personality traits, an index of parental characteristics, and a household asset index), and indicator variables for whether each of the covariates has missing data. Sample in column 1 includes all students in the survey records. Sample in columns 2-5 consists of placed applicants.  $p$ -values reported in brackets refer to the conventional asymptotic standard errors while those reported in curly brackets are adjusted for testing each null hypothesis across multiple outcomes through the step-wise procedure described in [Romano & Wolf \(2005a,b, 2016\)](#). All inference procedures take into account clustering of the error terms at the middle school level and the block randomization design.

Table B.4: Average Treatment Effects on Admission Outcomes

	Placed in 1st Round	Placed Any	Ranking of placement school
Treatment	-0.004 [0.796] {0.963}	-0.006 [0.719] {0.961}	0.141 [0.411] {0.804}
Performance index	0.068 [0.000] {0.001}	0.064 [0.000] {0.001}	-0.690 [0.000] {0.001}
Treatment X Performance index	-0.007 [0.647] {0.934}	-0.005 [0.751] {0.963}	-0.000 [1.000] {1.000}
Mean Control	0.857	0.884	3.692
Number of Observations	2824	2824	2493
Number of Clusters	90	90	90
R-squared	0.068	0.080	0.085

NOTE: Standard errors clustered at the middle school level. All specifications include a set of dummy variables which corresponds to the randomization strata, pre-determined characteristics (sex, characteristics of the school of origin, previous experience with practice exams providing feedback, aspirations to attend college, an index of personality traits, an index of parental characteristics, and a household asset index), and indicator variables for whether each of the covariates has missing data. Sample in columns 2-3 include all students who are matched in the administrative records of the COMIPEMS exam. Sample in column 3 consists of placed applicants.  $p$ -values reported in brackets refer to the conventional asymptotic standard errors while those reported in curly brackets are adjusted for testing each null hypothesis across multiple outcomes through the step-wise procedure described in [Romano & Wolf \(2005a,b, 2016\)](#). All inference procedures take into account clustering of the error terms at the middle school level and the block randomization design.

Table B.5: Model Estimates

	Control Sample				Treatment Sample
	Baseline	Beliefs	Random Coeff.	Mock Score	Mock Score
zsel	1.1220 (0.1865)	1.1137 (0.1874)	1.1215 (0.1866)	1.1002 (0.1882)	1.5163 (0.2135)
zselXzgp	0.1708 (0.0614)	0.1677 (0.0626)	0.1709 (0.0614)	0.1353 (0.0657)	0.1561 (0.0677)
zselXzmarg	-0.1006 (0.1040)	-0.0959 (0.1042)	-0.1004 (0.1040)	-0.0961 (0.1050)	-0.2217 (0.1239)
zselXzcal_mat	0.1280 (0.0841)	0.1191 (0.0856)	0.1275 (0.0841)	0.1092 (0.0854)	0.2556 (0.0837)
zselXedu	0.5391 (0.1718)	0.5342 (0.1739)	0.5391 (0.1719)	0.5126 (0.1740)	-0.0602 (0.1726)
zselXzmu1		0.0049 (0.0067)			
zselXzscore				0.1033 (0.0718)	0.1670 (0.0723)
geodist	-0.2755 (0.0197)	-0.2751 (0.0197)	-0.2768 (0.0197)	-0.2757 (0.0197)	-0.2005 (0.0239)
geodistXzgp	0.0094 (0.0064)	0.0106 (0.0066)	0.0096 (0.0064)	0.0030 (0.0070)	0.0076 (0.0076)
geodistXzmarg	0.0021 (0.0116)	0.0021 (0.0116)	0.0027 (0.0116)	0.0019 (0.0116)	-0.0450 (0.0149)
geodistXzcal_mat	0.0215 (0.0083)	0.0239 (0.0085)	0.0218 (0.0083)	0.0173 (0.0086)	0.0212 (0.0098)
geodistXedu	0.0302 (0.0160)	0.0324 (0.0163)	0.0305 (0.0161)	0.0252 (0.0163)	0.0042 (0.0176)
geodistXzmu1		-0.0009 (0.0007)			
geodistXzscore				0.0151 (0.0072)	0.0170 (0.0074)
school_d2Xzgp	-0.4900	-0.4854	-0.4921	-0.4536	-0.3656

Continued on next page



Table B.5 Model Estimates – Continued from Previous Page

	Control Sample				Treatment Sample
	Baseline	Beliefs	Random Coeff.	Mock Score	Mock Score
	(0.1680)	(0.1707)	(0.1680)	(0.1786)	(0.1895)
school_d2Xzmarg	0.0459 (0.3011)	0.0535 (0.3015)	0.0399 (0.3011)	0.0239 (0.3030)	1.1455 (0.3797)
school_d2Xzcal_mat	0.1057 (0.2153)	0.0689 (0.2179)	0.1027 (0.2153)	0.1550 (0.2198)	-0.0927 (0.2555)
school_d2Xedu	-0.5865 (0.4330)	-0.5755 (0.4363)	-0.5906 (0.4330)	-0.5215 (0.4390)	0.2154 (0.5107)
school_d2	2.1514 (0.5141)	2.0977 (0.5158)	2.1650 (0.5143)	2.3253 (0.5276)	-0.1198 (0.6189)
school_d2Xzmu1		0.0018 (0.0186)			
school_d2Xzscore				-0.1272 (0.2122)	-0.4005 (0.2157)
school_d3Xzgpa	-0.3409 (0.1882)	-0.3590 (0.1916)	-0.3422 (0.1882)	-0.3890 (0.2014)	-0.5132 (0.1995)
school_d3Xzmarg	0.4562 (0.3427)	0.4650 (0.3432)	0.4508 (0.3427)	0.3995 (0.3485)	1.1633 (0.3941)
school_d3Xzcal_mat	0.4423 (0.2318)	0.3767 (0.2339)	0.4384 (0.2318)	0.4091 (0.2358)	-0.3214 (0.2483)
school_d3Xedu	-0.4164 (0.4882)	-0.4436 (0.4916)	-0.4187 (0.4881)	-0.4239 (0.4945)	-0.0141 (0.5370)
school_d3	-0.6555 (0.5814)	-0.7139 (0.5822)	-0.6442 (0.5815)	-0.4159 (0.6048)	-1.9262 (0.6309)
school_d3Xzmu1		0.0147 (0.0205)			
school_d3Xzscore				0.2645 (0.2337)	0.0489 (0.2149)
school_d4Xzgpa	0.0030 (0.5471)	-0.0004 (0.5478)	0.0010 (0.5471)	0.0189 (0.5606)	-0.2476 (0.8303)
school_d4Xzmarg	-0.3096	-0.2683	-0.3161	-0.3621	2.0701

Continued on next page

Table B.5 Model Estimates – Continued from Previous Page

	Control Sample				Treatment Sample
	Baseline	Beliefs	Random Coeff.	Mock Score	Mock Score
	(1.4066)	(1.4046)	(1.4060)	(1.3662)	(1.8677)
school_d4Xzcal_mat	0.4560 (0.6966)	0.3904 (0.7100)	0.4510 (0.6962)	0.4019 (0.7206)	-0.3692 (1.2363)
school_d4Xedu	-1.0015 (1.2642)	-1.0365 (1.2748)	-1.0043 (1.2640)	-0.9636 (1.2587)	-7.3388 (76.2827)
school_d4	2.1428 (2.0790)	1.9942 (2.0944)	2.1575 (2.0780)	2.1849 (2.0571)	-2.6485 (2.7534)
school_d4Xzmu1		0.0159 (0.0615)			
school_d4Xzscore				0.3300 (0.7165)	-0.2664 (1.1829)
school_d5Xzgpa	-0.4680 (0.3380)	-0.4464 (0.3441)	-0.4692 (0.3381)	-0.5051 (0.3680)	0.0477 (44.0106)
school_d5Xzmarg	0.6763 (0.5203)	0.6712 (0.5173)	0.6708 (0.5204)	0.6364 (0.5193)	0.1239 (111.0151)
school_d5Xzcal_mat	0.2031 (0.5232)	0.2014 (0.5276)	0.1983 (0.5233)	0.1772 (0.5329)	0.3491 (74.2039)
school_d5Xedu	-0.5802 (1.1287)	-0.5535 (1.1304)	-0.5807 (1.1284)	-0.6255 (1.1306)	0.6071 (87.1108)
school_d5	1.3488 (1.1428)	1.3315 (1.1383)	1.3599 (1.1430)	1.5062 (1.1599)	-6.8738 (186.2420)
school_d5Xzmu1		-0.0160 (0.0358)			
school_d5Xzscore				0.1638 (0.4216)	0.5097 (40.7020)
school_d6Xzgpa	-0.5034 (0.2606)	-0.6263 (0.2709)	-0.5053 (0.2606)	-0.6576 (0.2800)	-0.1901 (0.2474)
school_d6Xzmarg	-0.4732 (0.3856)	-0.4307 (0.3896)	-0.4791 (0.3856)	-0.5044 (0.3887)	0.1023 (0.4312)
school_d6Xzcal_mat	-0.3591	-0.3816	-0.3609	-0.4022	0.5233

Continued on next page

Table B.5 Model Estimates – Continued from Previous Page

	Control Sample				Treatment Sample
	Baseline	Beliefs	Random Coeff.	Mock Score	Mock Score
	(0.3955)	(0.3883)	(0.3953)	(0.3984)	(0.3026)
school_d6Xedu	-1.2103 (0.8684)	-1.3210 (0.8714)	-1.2116 (0.8685)	-1.4034 (0.8961)	0.3299 (0.6197)
school_d6	2.0773 (0.6938)	1.9030 (0.7097)	2.0908 (0.6941)	2.1907 (0.7063)	0.9946 (0.7585)
school_d6Xzmu1		0.0441 (0.0288)			
school_d6Xzscore				0.5917 (0.3262)	-0.4600 (0.2707)
school_d7Xzgpa	-0.4686 (0.1595)	-0.4647 (0.1622)	-0.4705 (0.1596)	-0.5108 (0.1710)	-0.7586 (0.1816)
school_d7Xzmarg	0.2240 (0.2751)	0.2314 (0.2754)	0.2173 (0.2751)	0.1635 (0.2790)	0.9612 (0.3458)
school_d7Xzcal_mat	0.1418 (0.2068)	0.1060 (0.2084)	0.1384 (0.2068)	0.1153 (0.2108)	0.1620 (0.2295)
school_d7Xedu	-0.8298 (0.4437)	-0.8226 (0.4464)	-0.8327 (0.4437)	-0.8419 (0.4511)	0.0578 (0.4746)
school_d7	1.3646 (0.4717)	1.3242 (0.4721)	1.3789 (0.4718)	1.6000 (0.4900)	-0.3695 (0.5717)
school_d7Xzmu1		0.0011 (0.0174)			
school_d7Xzscore				0.2551 (0.2034)	-0.0882 (0.2021)
school_d8Xzgpa	-0.4209 (0.2753)	-0.4408 (0.2792)	-0.4109 (0.2749)	-0.3449 (0.2859)	-0.5504 (0.3131)
school_d8Xzmarg	0.5170 (0.4912)	0.4736 (0.4927)	0.5296 (0.4909)	0.5121 (0.4903)	2.0198 (0.5712)
school_d8Xzcal_mat	-0.2240 (0.3503)	-0.3229 (0.3537)	-0.2221 (0.3502)	-0.1521 (0.3594)	-0.8218 (0.4050)
school_d8Xedu	-1.5097	-1.5752	-1.5016	-1.4205	0.7425

Continued on next page

Table B.5 Model Estimates – Continued from Previous Page

	Control Sample				Treatment Sample
	Baseline	Beliefs	Random Coeff.	Mock Score	Mock Score
	(0.6523)	(0.6575)	(0.6522)	(0.6551)	(0.7090)
school_d8	2.8707 (0.8756)	2.7543 (0.8764)	2.8343 (0.8743)	3.1237 (0.9074)	0.0656 (1.0629)
school_d8Xzmu1		0.0382 (0.0299)			
school_d8Xzscore				-0.2418 (0.3727)	-0.8604 (0.3814)
school_d9Xzgpa	-0.8674 (0.6517)	-1.0825 (0.7667)	-1.5338 (1.2238)	-0.4433 (0.7054)	-0.8429 (0.7255)
school_d9Xzmarg	-3.9261 (4.0107)	-4.0808 (4.2425)	-5.8883 (6.4934)	-4.4321 (4.3531)	1.7951 (1.4758)
school_d9Xzcal_mat	0.3388 (1.1657)	0.3707 (1.2460)	-0.4466 (1.9763)	0.5158 (1.2804)	1.1686 (1.0350)
school_d9Xedu	-6.6213 (18.5332)	-7.0666 (21.3624)	-18.0524 (3.2e+03)	-8.5800 (35.3352)	-7.7138 (42.5010)
school_d9	8.6761 (4.9403)	7.8062 (5.3762)	8.4710 (7.2752)	10.6771 (5.3848)	-0.2287 (2.9337)
school_d9Xzmu1		0.1297 (0.1364)			
school_d9Xzscore				-1.7571 (0.9231)	-0.9392 (0.9438)
school_d10Xzgpa	-1.7266 (0.5639)	-1.8671 (0.5980)	-1.7272 (0.5640)	-1.9192 (0.5713)	-0.8270 (0.9407)
school_d10Xzmarg	2.2972 (0.8357)	2.7499 (1.0006)	2.2940 (0.8360)	2.5223 (0.9385)	3.8977 (2.6261)
school_d10Xzcal_mat	0.6398 (1.1149)	0.5967 (1.1127)	0.6372 (1.1153)	0.6962 (1.1210)	-0.2802 (1.0640)
school_d10Xedu	0.3447 (1.0174)	0.4360 (1.0343)	0.3434 (1.0174)	0.2131 (1.0135)	-7.0258 (73.9448)
school_d10	-5.5771	-7.0059	-5.5702	-6.3276	-9.0535

Continued on next page

Table B.5 Model Estimates – Continued from Previous Page

	Control Sample				Treatment Sample
	Baseline	Beliefs	Random Coeff.	Mock Score	Mock Score
	(2.1025)	(2.6035)	(2.1034)	(2.4699)	(5.5151)
school_d10Xzmu1		0.0873 (0.0557)			
school_d10Xzscore				1.1159 (0.5822)	-0.6322 (1.0539)
school_d11Xzgpa	-0.0696 (0.1709)	-0.0567 (0.1735)	-0.0717 (0.1709)	-0.1399 (0.1834)	-0.1929 (0.2037)
school_d11Xzmarg	-0.3847 (0.2955)	-0.3727 (0.2960)	-0.3888 (0.2954)	-0.4173 (0.2981)	-0.2980 (0.3966)
school_d11Xzcal_mat	0.1265 (0.2233)	0.1059 (0.2237)	0.1234 (0.2233)	0.0825 (0.2291)	0.4417 (0.2759)
school_d11Xedu	-0.9143 (0.5300)	-0.8896 (0.5321)	-0.9150 (0.5300)	-0.9300 (0.5377)	-0.6561 (0.6377)
school_d11	0.1453 (0.4892)	0.0974 (0.4901)	0.1539 (0.4893)	0.3222 (0.5039)	-0.4086 (0.6524)
school_d11Xzmu1		-0.0016 (0.0183)			
school_d11Xzscore				0.2641 (0.2103)	-0.7440 (0.2367)
school_d12Xzgpa	-0.5456 (0.2072)	-0.4731 (0.2092)	-0.5480 (0.2071)	-0.5830 (0.2243)	-0.1712 (0.2145)
school_d12Xzmarg	-0.2134 (0.3599)	-0.1925 (0.3589)	-0.2201 (0.3599)	-0.2652 (0.3628)	0.7170 (0.4180)
school_d12Xzcal_mat	0.8073 (0.2923)	0.9102 (0.3032)	0.8028 (0.2921)	0.7823 (0.2961)	0.2892 (0.2733)
school_d12Xedu	-1.6815 (0.7086)	-1.4674 (0.7122)	-1.6829 (0.7084)	-1.6647 (0.7152)	-0.1300 (0.5924)
school_d12	0.3454 (0.6347)	0.2105 (0.6379)	0.3597 (0.6349)	0.5492 (0.6478)	-1.3730 (0.7469)
school_d12Xzmu1		-0.0542			

Continued on next page

Table B.5 Model Estimates – Continued from Previous Page

	Control Sample			Treatment Sample	
	Baseline	Beliefs (0.0222)	Random Coeff.	Mock Score	Mock Score
school_d12Xzscore				0.1626 (0.2543)	-0.6003 (0.2323)
school_d13Xzgpa	-0.2689 (0.1793)	-0.2504 (0.1830)	-0.2697 (0.1793)	-0.3052 (0.1910)	-0.0992 (0.2065)
school_d13Xzmarg	-0.0140 (0.3112)	-0.0080 (0.3111)	-0.0192 (0.3113)	-0.0515 (0.3137)	0.9202 (0.4200)
school_d13Xzcal_mat	0.2582 (0.2383)	0.2417 (0.2382)	0.2554 (0.2383)	0.2353 (0.2417)	-0.2489 (0.3014)
school_d13Xedu	-0.4517 (0.5552)	-0.4199 (0.5592)	-0.4519 (0.5552)	-0.4605 (0.5634)	-0.2792 (0.6476)
school_d13	1.1265 (0.5466)	1.0897 (0.5469)	1.1364 (0.5468)	1.2961 (0.5669)	-0.7107 (0.7262)
school_d13Xzmu1		-0.0097 (0.0194)			
school_d13Xzscore				0.1650 (0.2318)	-0.3089 (0.2328)
school_d14Xzgpa	-0.4737 (0.1648)	-0.4427 (0.1678)	-0.4745 (0.1647)	-0.5208 (0.1748)	-0.3661 (0.1816)
school_d14Xzmarg	0.2396 (0.2767)	0.2491 (0.2769)	0.2353 (0.2767)	0.2097 (0.2803)	0.2678 (0.3556)
school_d14Xzcal_mat	0.1451 (0.2324)	0.1340 (0.2330)	0.1420 (0.2324)	0.1185 (0.2356)	-0.1823 (0.2639)
school_d14Xedu	-0.9139 (0.5757)	-0.8581 (0.5785)	-0.9143 (0.5757)	-0.9059 (0.5817)	0.0955 (0.4909)
school_d14	-0.3591 (0.4722)	-0.4182 (0.4743)	-0.3511 (0.4722)	-0.2067 (0.4970)	-0.6650 (0.5929)
school_d14Xzmu1		-0.0183 (0.0177)			
school_d14Xzscore				0.1913	-0.3536

Continued on next page

Table B.5 Model Estimates – Continued from Previous Page

	Control Sample			Treatment Sample	
	Baseline	Beliefs	Random Coeff.	Mock Score (0.2131)	Mock Score (0.2009)
school_d15Xzgpa	-0.4223 (0.1433)	-0.3837 (0.1459)	-0.4237 (0.1432)	-0.4413 (0.1563)	-0.1745 (0.1630)
school_d15Xzmarg	-0.1938 (0.2424)	-0.1955 (0.2426)	-0.1982 (0.2424)	-0.2282 (0.2453)	-0.2634 (0.3076)
school_d15Xzcal_mat	0.0406 (0.1831)	0.0482 (0.1833)	0.0371 (0.1831)	0.0219 (0.1868)	0.5696 (0.2143)
school_d15Xedu	-0.5312 (0.4082)	-0.4541 (0.4117)	-0.5317 (0.4082)	-0.5103 (0.4153)	0.1797 (0.4259)
school_d15	0.6226 (0.4040)	0.5969 (0.4041)	0.6314 (0.4040)	0.7729 (0.4223)	0.4014 (0.5192)
school_d15Xzmu1		-0.0205 (0.0153)			
school_d15Xzscore				0.1187 (0.1818)	-0.4891 (0.1772)
school_d16Xzgpa	-0.0527 (29.9480)	-0.1297 (40.6054)	-0.0542 (1.0e+04)	-0.3235 (33.9656)	0.3034 (1.5405)
school_d16Xzmarg	0.1215 (50.3901)	0.0704 (61.0544)	0.1196 (1.7e+04)	0.2465 (66.1300)	0.4916 (3.0450)
school_d16Xzcal_mat	0.5501 (37.0224)	0.4532 (46.0118)	0.5413 (1.2e+04)	0.4017 (50.0650)	1.8845 (1.9489)
school_d16Xedu	-0.4988 (43.3660)	-0.5767 (52.7771)	-0.4859 (1.5e+04)	-0.6337 (59.4293)	-7.5862 (87.9174)
school_d16	-4.2685 (104.3225)	-4.7037 (127.4938)	-15.9019 (3.5e+04)	-5.8524 (176.0093)	-2.0386 (6.7791)
school_d16Xzmu1		0.0362 (3.4851)			
school_d16Xzscore				0.7880 (50.9079)	0.7052 (2.2986)
school_d17Xzgpa	-0.2929	-0.3021	-0.2902	-0.2693	0.0685

Continued on next page



Table B.5 Model Estimates – Continued from Previous Page

	Control Sample				Treatment Sample
	Baseline	Beliefs	Random Coeff.	Mock Score	Mock Score
	(0.3025)	(0.3073)	(0.3023)	(0.3160)	(0.3356)
school_d17Xzmarg	0.1280 (0.5937)	0.1103 (0.5925)	0.1267 (0.5946)	0.0713 (0.5961)	2.3590 (0.5964)
school_d17Xzcal_mat	0.8140 (0.3915)	0.7299 (0.3941)	0.8086 (0.3916)	0.7971 (0.3987)	-0.3870 (0.3985)
school_d17Xedu	-2.8503 (0.7347)	-2.8997 (0.7395)	-2.8418 (0.7346)	-2.7196 (0.7361)	1.0096 (0.7239)
school_d17	4.2555 (1.0180)	4.1349 (1.0202)	4.2529 (1.0183)	4.3824 (1.0367)	-0.2632 (1.1303)
school_d17Xzmu l		0.0314 (0.0320)			
school_d17Xzscore				-0.0428 (0.4250)	-0.6045 (0.3984)
school_d18Xzgpa	-0.2760 (0.3790)	-0.2781 (0.3906)	-0.2581 (0.3892)	-0.1992 (0.3891)	0.4907 (0.4923)
school_d18Xzmarg	0.2244 (0.6832)	0.1621 (0.6886)	0.2318 (0.6928)	0.2275 (0.6865)	2.4687 (0.7835)
school_d18Xzcal_mat	0.3536 (0.4610)	0.2837 (0.4768)	0.3601 (0.4670)	0.4095 (0.4677)	-0.1632 (0.5454)
school_d18Xedu	-1.8263 (0.8288)	-1.8711 (0.8369)	-1.7942 (0.8431)	-1.6847 (0.8299)	1.2642 (0.9170)
school_d18	3.4045 (1.2318)	3.4123 (1.2695)	3.3414 (1.2631)	3.5000 (1.2871)	-2.1179 (1.5437)
school_d18Xzmu l		0.0206 (0.0416)			
school_d18Xzscore				-0.1878 (0.4973)	-0.5527 (0.5583)
N. of Obs.	637901	637901	637901	637901	590526
Log lik	-4414	-4399	-4413	-4401	-4189

NOTE: This table displays the full set of maximum-likelihood estimates and standard errors (in parenthesis) of the parameters of the school choice model (1).

Table B.6: Stylized Example of Two Applicants at the Margin (RD Sample)

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

NOTE: This table provides an example where two applicants are on the margin of receiving an offer for an elite school and a non-elite school. One applicant has a application score of 79 and receives an offer from an IPN school, whereas the other receives an offer from another non-IPN school because she has a slightly lower application score of 77. By comparing the outcomes of these applicants we can estimate the effect of getting an offer of elite schools, while ruling out that differences in their outcomes are driven by unobserved heterogeneity.

Table B.7: Covariate Smoothness Around the Admission Threshold

	Female	Number of Siblings	Mother with High School	Father with High School	Age
IPN Admission	0.032 (0.025)	0.072 (0.073)	-0.021 (0.016)	-0.018 (0.022)	0.029 (0.041)
Bias-corrected CI	[-.017 ; .091]	[-.106 ; .23]	[-.063 ; .011]	[-.074 ; .029]	[-.077 ; .118]
Optimal BW	13.701	12.488	12.246	10.699	9.479
Mean Dep. Var.	0.432	2.005	0.186	0.319	15.555
Nb. of Obs.	10,828	9,077	8,854	7,452	8,186

NOTE: This table displays non-parametric estimates of marginal admission in elite schools on individual-level characteristics of the applicants. Standard errors reported in parenthesis are clustered at the school×year level. All the specifications are defined within mean-square error optimal bandwidths (BW). We also report the bias-corrected confidence intervals obtained using the robust estimator proposed in [Calonico et al. \(2014\)](#).