WORKING PAPERS

N° 660

July 2024

“Perceived Ability and School Choices: Experimental Evidence and Scale-up Effects”

Matteo Bobba, Veronica Frisancho and Marco Pariguana

Toulouse School of Economics

Perceived Ability and School Choices: Experimental Evidence and Scale-up Effects[*]

Matteo Bobba[†] Veronica Frisancho[‡] Marco Pariguana[§]

July 15, 2024

Abstract

This paper explores an information intervention designed and implemented within a school assignment mechanism in Mexico City. Through a randomized experiment, we show that providing a subset of applicants with feedback about their academic performance can enhance sorting by skill across high school tracks. This reallocation effect results in higher completion rates three years post-assignment. We further integrate the experimental evaluation into an empirical model of school choice and educational outcomes to assess the impact of the intervention for the overall population of applicants. Information provision is shown to increase the ex-ante efficiency of the student-school allocation, while congestion externalities are detrimental for the equity of education outcomes.

[*]This paper is dedicated to the memory of YingHua He, whose work and generous feedback have inspired parts of the analysis. The current draft supersedes our earlier working papers "Perceived Ability and School Choices" and "Learning About Oneself: The Effect of Performance Feedback on School Choices" (Bobba and 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, Thierry Magnac, Christopher Neilson, Parag Pathak, Imran Rasul, Michela Tincani, 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.

[†]Toulouse School of Economics, University of Toulouse Capitole & CEPR & IZA. E-mail: matteo.bobba@tse-fr.eu.

[‡]CAF, Development Bank of Latin America and the Caribbean. E-mail: vfrisancho@caf.com.

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

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, 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). Recent 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.[1]

In this paper we are, instead, concerned with situations where it is often feasible to follow the original blueprint of the program. This is the case of information interventions in the context of human capital investments. Education choices are made under uncertainty and rely on subjective expectations about present and future returns. Information provision may partly resolve this specific source of subjective uncertainty. Yet, it is unclear to what extent students internalize signals that may be more or less informative for subsequent outcomes. Perhaps even more importantly, informing a large number of agents in education markets is prone to generating a variety of spillover and equilibrium effects that may fundamentally alter the inference drawn from small-scale studies.

This study represents an attempt to generalize the results from a randomized evaluation of an information intervention that provides students with individualized feedback about their academic skills. Experimental evidence drawn from a socio-economically disadvantaged sample demonstrates that the performance feedback contributes to better aligning skills with high school tracks, and partly enhances the persistence in upper secondary education. A model-based implementation of the same intervention at full scale shows that information provision enhances the ex-ante efficiency of the student-school allocation. However, congestion externalities across high school programs would largely offset the positive impact on education outcomes for the sub-group of students targeted by the original evalu-

[1] See, e.g., Banerjee et al. (2017); Muralidharan and Niehaus (2017); Al-Ubaydli et al. (2020) for recent reviews of the growing literature on scaling up randomized experiments and August et al. (2006); Caron et al. (2021); Agostinelli et al. (2023) on the issues of accurately replicating the exact program designs from a controlled research setting to real-world implementation by the government. Eligible participants in experiments may also be positively selected and not representative of the general population (Heckman, 1992; Allcott, 2015; Davis et al., 2021).

ation. These findings offer novel insights on the channels through which cost-effective and ex-ante scalable policy solutions may fail to deliver the expected results when implemented at scale.

The setting of our analysis is the secondary education market of the metropolitan area of Mexico City, in which a centralized clearinghouse coordinates admission to the quasi-universe of public high schools in the region. Close to 300,000 students apply every year to the system by submitting rank-ordered lists of their preferred high school programs during the second-to-last term while in the ninth grade (i.e., the last year of middle school). At the end of the school year, all applicants take a unique standardized admission test that determines priority in the assignment system and assesses curricular knowledge as well as verbal and analytical aptitude. The timing of the events, which is common across school/college assignment mechanisms in other countries, implies that high stake decisions regarding school choice may not incorporate relevant information about an applicant's academic skills. Indeed, detailed survey data that elicit information on the subjective distribution of beliefs indicate that over 80% of the students overestimate their performance in the test by the time they apply to the system.

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 in Mexico City, and communicate individual score results to a randomly chosen subset of students before the school rankings are submitted. In contrast to other interventions such as tutoring or career counseling (see, e.g., Goux et al., 2017; Carlana et al., 2022), the score in the mock exam provides students with a "light-touch" signal about their own academic skills that is both easy to interpret and to implement to the broader population of the applicants. The linear correlation in our sample between performance in the mock exam and the actual admission exam is 0.82. In turn, the linear correlation between a freely available signal of academic readiness, such as the middle school GPA and the mock exam score, is only 0.48. The mock exam score also predicts later educational outcomes for the students in the control group (i.e. those who took the test but did not receive any performance feedback).

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.

The school choice response translates into differential placement outcomes, which, in turn, alter subsequent educational trajectories. Three years after school assignment, the probability of graduating from high school on time is, on average, 5.4 percentage points higher among under achieving students who received performance feedback. Although noisily estimated, this effect on education outcomes is sizable as it corresponds to a 13 percent increase when compared to the sample average in the control group. Importantly, the observed gains in persistence throughout secondary education do not seem to be explained by the fact that lower-performing students tend to sort into easier-to-graduate schools as a result of the information intervention

We next study the effect of a scaled-up version of the randomized experiment that mandates the universal implementation of a mock exam or, alternatively, discloses admission exam scores to the applicants before the submission of the rank-ordered lists. Under the assumption that the realized matching equilibrium is stable, we can express the student-school allocation as the outcome of a discrete choice problem with student-specific choice sets (Fack et al., 2019). We specify a school choice model that captures rich heterogeneity across students regarding their valuations over the available schooling alternatives (see, e.g., Arcidiacono et al., 2016; Abdulkadiroglu et al., 2020). Using the estimated preference parameters for the evaluation sample along with student-level data for the universe of applicants, we predict the sorting patterns across schools under the status quo as well as the counterfactual regime of the information intervention.

The estimated preference distribution over schools for the students in the control group replicates the main features of the data for the broader population of the applicants. Namely, it tracks the school placement outcomes, even for those students that are far beyond the support of the experimental data, as well as the equilibrium cutoff scores at the school level. The new matching equilibrium of the information intervention based on the estimated parameters of the students in the treatment group effectively takes into account the sorting and congestion effects resulting from aggregate changes in the demand side. Supply-side responses are straightforward to quantify in our setting, since schools simply accept or reject prospective applicants in descending order, based on their priority, until capacity constraints are reached.

We link the equilibrium predictions based on the choice model with a school value added framework featuring substantial heterogeneity across students (Walters, 2018; Abdulkadiroglu et al., 2020). The model further allows for equilibrium changes in school-level peer composition to affect education outcomes. We leverage the key features of the assignment mechanism in order to obtain consistent estimates of our parameters of interest. Notably, we condition on skill measures, demographics, and ranked-order lists as a way to minimize the bias arising from the non-random assignment of students across schools. The estimated value added model allows us to map the individual sorting patterns into subsequent education outcomes under both the status quo and the counterfactual regime of the information intervention.

Overall, the provision of information enhances the ex-ante efficiency of the matching equilibrium. While there are no changes at the extensive margin of the admission process, the share of students assigned to their most preferred option increases by nine percentage points, from 16 percent to 25 percent, under the scenario with performance feedback. These aggregate patterns mask substantial heterogeneity in the demand-side responses across applicants' socio-economic status (SES). The bulk of the changes in the school choices between the status quo and the information intervention are concentrated among applicants who are socio-economically better off (high-SES). These students increase their demand for academic schools and symmetrically decrease the demand for selective and prestigious (elite) schools as a result of the information intervention. Instead, and in line with the randomized evaluation, low-SES applicants are, on average, unresponsive to the information intervention. The lower demand-side pressure on elite programs crowds in high-achieving and socio-economically disadvantaged applicants. Overall, the information intervention increases the representation of low-SES applicants at elite schools by more than 20 percentage points.

We finally use the predictions of the value added model in order to assess the impact of the equilibrium allocation associated to the information intervention, when compared to the status-quo allocation, on education outcomes. The model predicts a large and negative effect of attending elite schools on the probability of graduation. Accordingly, low-SES applicants would be worse off in terms of high-school graduation. This is particularly the case for those with a relatively high admission score, who would be 4-7 percentage points less likely to complete upper secondary education on time. Conversely, by opting out from elite schools towards other academic schools, high-SES applicants would increase school completion rates by 2 to 9 percentage points depending on their admission score.

The implication of these results contrasts with the conclusion drawn from the experimental evaluation. This discrepancy can be explained by a congestion effect across high-school tracks among applicants, which is absent in the small-scale implementation. Low-SES students who are more likely to attend elite schools are found to experience negative consequences on their subsequent academic trajectories, possibly by dropping-out from upper secondary education. While there may be other channels through which these students may take advantage of elite admission in the medium term, such as social networks that may positively influence later employment opportunities, comparable evidence from Chile shows that these gains are muted for students outside of historically advantaged groups (Zimmerman, 2019). In this sense, successfully scaling up the intervention may require providing additional signals targeted at low-SES applicants that are informative about their probability of graduation when attending elite schools.

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., Lavecchia et al., 2016; Haaland et al., 2023). 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, both in partial and in general equilibrium, the medium-run impact on academic trajectories triggered by the 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. This feature greatly simplifies the simulation of the matching equilibrium (Agarwal and Somaini, 2020), and it further allows us to unpack the congestion externalities of the information intervention at scale.

2 Context, Experimental Design, and Data

In this section, we first describe the relevant features of the study setting. We next provide a few details on the design and implementation of the information intervention. We finally discuss the rich combination of administrative and survey datasets that we use throughout the empirical analysis.

2.1 Centralized School Assignment in Mexico City

Since 1996, a local commission (COMIPEMS, by its Spanish acronym) of 16 upper secondary public institutions, or colleges, has centralized 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 and their basic characteristics. In addition to the registration form, students complete a socio-demographic survey and submit a ranked list of 20 schools, at most. At the end of the school year, applicants take a standardized achievement test. Priority in the system is determined based on total scores in that test. The matching algorithm goes down the roster of priority-ranked applicants to sequentially assign them to their most preferred schooling option with available seats. The submission of school preferences before the application of the admission exam is not unique of the COMIPEMS system. Other school assignment mechanisms that use strict priority rules have this institutional feature, such as in Ghana (Ajayi, 2022), Kenya (Lucas and Mbiti, 2014), Barbados (Beuermann and Jackson, 2020), Trinidad and Tobago (Beuermann et al., 2022), and some Chinese provinces (Chen and Kesten, 2017). The timing of the events in the Mexican case is supposed to provide public officials with a "ballpark estimate" of the number of seats that should be made available by the sponsoring colleges in each yearly round of the assignment process.

Each applicant is matched with one school. Whenever a tie in the score occurs for the last available spot in a given school, 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 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 general, 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.

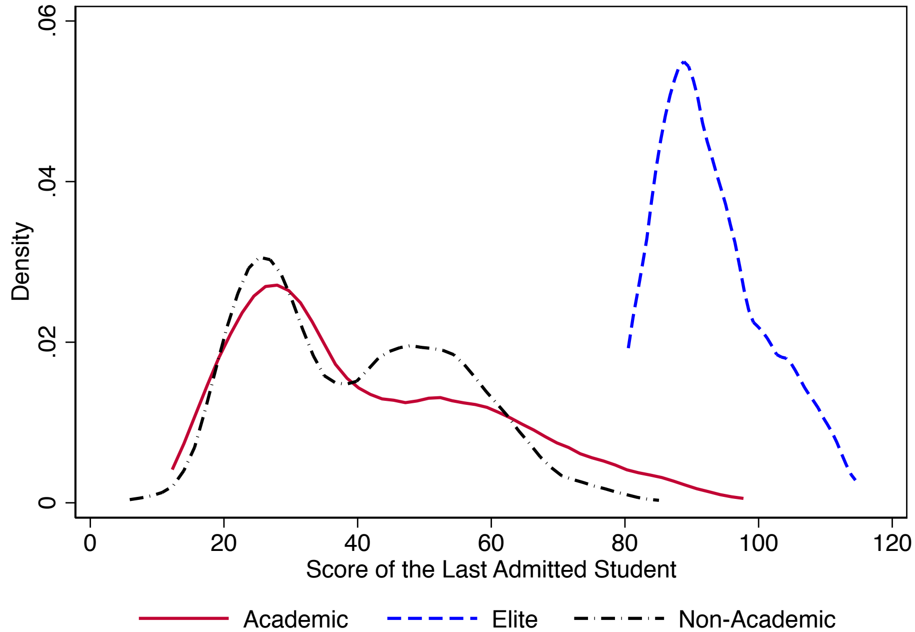
The Mexican system offers three educational tracks at the upper secondary level: General, Technical, and Vocational Education. Each college within the assignment system offers a unique track. The general track is academically oriented and includes traditional schools that are more focused on preparing 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 high school completion. The vocational track exclusively trains students to become technically adept. A small sub-set of schools (32 out of 628) within the assignment system in Mexico City are 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 the associated colleges are highly demanded. 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'.

Figure 1 depicts the distribution of cutoff scores across the three main types of high-schools, or tracks. 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. Some non-academic schools have gained popularity in the system due to their reputation in placing graduates in vocationally related occupations, which explains their relatively high cutoff scores. Elite schools clearly stand out in terms of selectivity.

2.2 The Information Intervention

Figure 2 depicts the timing of the activities related to the intervention. During the second half of the 2013-14 academic year, we implemented a mock exam in 90 middle schools (see

Figure 1: Distribution of Cutoff Scores

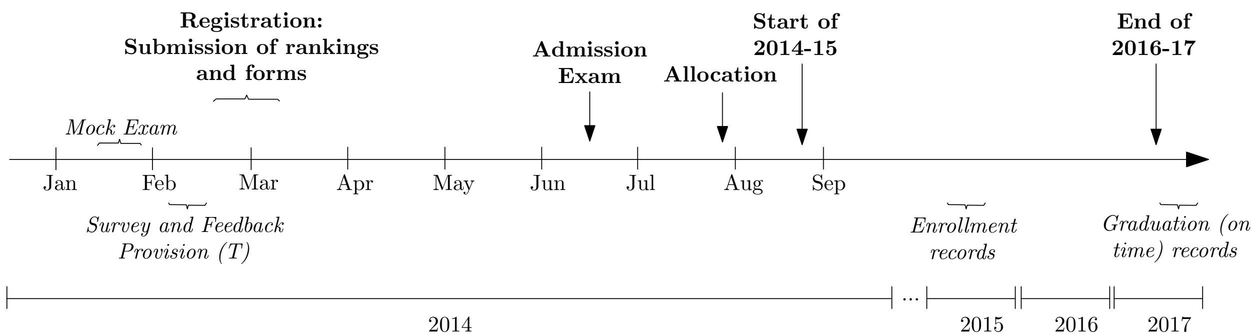


Note: The cutoff score for each high school program refers to the lowest score in the admission exam of the students accepted there in the 2014 assignment process. 'Academic' schools are defined as the high school programs in the general track, 'Non-academic' schools are those in the technical and vocational 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).

Section 2.3). One or two weeks later, and just before the submission of the school rankings, we implemented a survey in those schools during which enumerators 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 order to avoid reporting biases due to the influence of peers and/or social image concerns (Burks et al., 2013; Ewers and 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 (see Figure A.1 in the Appendix). Both pre-printed bars served the purpose of providing the student with additional elements to better frame her own score.

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

Figure 2: Timeline of Events



sciences) as well as mathematical and verbal aptitude sections.[2] 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., 2021). 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 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 own academic skills. The linear correlation in our sample between performance in the mock exam and the actual exam is 0.82. In turn, the linear correlation between a freely available proxy of academic readiness, such as the middle school GPA, and the mock exam score is only 0.48 (see also Figure B.2 in the Appendix). Both the scores in the admission exam and in the mock exam are strong predictors of later academic success. As shown in Table B.1 of

[2]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.

the Appendix, a one-standard-deviation increase in the mock exam score is associated with a 7.2 percentage-point increase (p-value=0.001) in the probability of graduating from high school on time.

2.3 Sample Selection and Randomization

To select the evaluation sample, we focus on middle schools with (i) a considerable mass of applicants, more than 30, in the 2012-2013 round of the centralized mechanism and (ii) that are located in neighborhoods with high or very high poverty levels (CONEVAL, 2018). The latter criterion was largely influenced by the literature showing that less privileged students tend to be relatively more misinformed when making educational choices (Avery and Hoxby, 2012; Hastings and Weinstein, 2008; Jensen, 2010). In our context, 44 percent of the applicants enrolled in schools from more affluent neighborhoods took preparatory courses for the admission exam before submitting their school rankings. This figure drops to 12 percent among applicants from schools in high poverty areas.

Schools that comply with our sample selection criteria are stratified by region and performance terciles. We group them into four geographic regions and terciles of school-average math test scores amongst ninth graders (see Section 2.4). 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, 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, it cannot induce differential attrition patterns.

The match rate between the survey and the application records is 88 percent (2,828 students). As shown in Appendix Table B.2, the participation in the assignment system is balanced between the treatment and the control group. The evaluation sample comprises the 2,493 applicants who were eligible for assignment through the matching algorithm. Appendix Table B.3 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.

Table 1 shows that the evaluation sample is largely comparable to the general population of applicants in terms of initial credentials, such as GPA or college aspirations. However, the average applicant in our sample scores 4-points less in the admission exam than the average applicant in the universe (0.2 standard deviations). Consistent with our focus on

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

	All Applicants Mean (Std. Dev)	Experiment Mean (Std. Dev)	All-Experiment Mean Difference [p-value]
Grade Point Average in middle school (GPA)	8.058	8.119	-0.061
	(0.871)	(0.846)	[0.001]
Has some disabilities (1=yes)	0.118	0.145	-0.027
	(0.323)	(0.352)	[0.000]
Scholarship in middle school (1=yes)	0.116	0.110	0.006
	(0.320)	(0.313)	[0.401]
Indigenous	0.041	0.093	-0.052
	(0.198)	(0.290)	[0.000]
Plans to attend higher education (1=yes)	0.662	0.670	-0.008
	(0.473)	(0.470)	[0.378]
Admission exam score	69.506	65.400	4.107
	(20.705)	(19.401)	[0.000]
One parent with at least tertiary education (1=yes)	0.236	0.147	0.089
	(0.425)	(0.354)	[0.000]
Average math score in middle school (z-score)	0.000	-0.208	0.208
	(1.000)	(0.712)	[0.000]
Neighborhood SES index (z-score)	0.000	-1.504	1.504
	(1.000)	(0.494)	[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 evaluation sample. The third columns displays mean differences and the associated p-values (in brackets) for the null hypothesis of equal means. The observations in the first column comprise all the applicants in the year 2014 who were eligible to be assigned through the matching algorithm. The observations in the second column comprise the evaluation sample of the randomized information intervention.

relatively disadvantaged students, the applicants in the evaluation sample are less likely to have parents with tertiary education, they attend middle schools with lower performing students, and reside in poorer neighborhoods.

2.4 Data and Measurement

Our analysis draws on several data sources. First, we have access to administrative data on different cohorts of applicants for several rounds of the assignment mechanism. These records include socio-demographic variables, such as gender, age, and parental education, among others. They also contain information on school preference rankings, admission exam scores, and placement outcomes. We link this dataset with the school-average math test scores in the national standardized examination (Evaluacion Nacional de Logros Academicos en Centros Escolares, ENLACE) applied in ninth grade.

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 evaluation 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 details on the elicitation of the individual data on beliefs in our setting. There are a few students with missing values in the beliefs data (247 observations, or 10% of the sample), which implies an effective sample size of 2,246 applicants for the analysis presented in Section 3.2. The incidence of missing values is balanced between the treatment and the control group (coeff=0.006, p-value=0.367).

Third, we assemble and harmonize longitudinal data on the schooling trajectories through upper secondary education for the students in the experimental sample. The resulting dataset allows us to measure high school enrollment, drop-out during the tenth grade, and graduation on time from high school (twelfth grade). It is not possible to track those applicants who end up enrolling in schools outside the centralized system. About 80 percent of the applicants in the control group enroll by the next academic year in the high school program in which they were assigned through the centralized process. However, only 45 percent successfully graduate from high school after three years. These figures clearly reflect inadequate academic progress through upper secondary education, due to either school dropout or grade retention, both strong indicators of a mismatch between schooling careers and students' individual skills.

Fourth, we match the individual identifiers for all the applicants who participated in the centralized assignment system with their ENLACE scores in twelfth grade. This is a good proxy for the probability of graduating from almost any high school in the country, including private schools (Dustan et al., 2017; Estrada and Gignoux, 2017; Dustan, 2020), except for the UNAM-sponsored high-schools (16 school programs out of 628 participating schools in the centralized assignment system).[3] The ENLACE test was discontinued in 2014, and so

[3]The UNAM is a higher education institution that is officially autonomous of the government. It's representatives opted for not administering the ENLACE exam to its students.

we collect the schooling trajectories for the ninth graders in the 2010 cohort of applicants.

Conditional on assignment in the centralized system, only 41 percent of all high-school students in Mexico City graduate on time. As Appendix Table B.4 shows, more than three quarters of the students who complete secondary education graduate from high-school in the statutory three-year period. This share is pretty much stable across high-school tracks. This evidence strongly suggests that the majority of the students who don't graduate on-time are likely to drop out from the education system, rather than being held back or transferring to another school.

3 Experimental Evidence

Providing information about individual performance in the mock exam potentially allows students to revise their own beliefs and thereby make high-school track choices that are better aligned with their own academic potential, which, in turn, may lead to better educational outcomes. In this section, we document the effect of the performance feedback on subjective expectations about academic performance, as well as individual outcomes regarding school choices, placement, and subsequent schooling trajectories.

3.1 Empirical Model

We estimate linear regression models of the following form:

$$Y_i = \alpha_0 + \alpha_1 T_{j(i)} + \alpha_2 A_i + \alpha_3 A_i T_{j(i)} + \boldsymbol{\delta' X_i} + \epsilon_i, \tag{1}$$

where Y_i is an individual-level choice or outcome (expected or realized) for student i in one of the 90 middle schools j where we administer the mock exam. The indicator variable $T_{j(i)}$ takes a value of one if the school is in the treatment group and hence its students receive performance feedback in the mock exam, and zero otherwise. The A_i variable is a standardized index of academic achievement, which is obtained as the weighted average of the GPA in middle school, the score in the mock exam, and the score in the admission exam.[4]

The vector $\boldsymbol{X_i}$ contains a set of dummy variables that correspond to the randomization strata (location $\times$ school-average test score indicators), pre-determined characteristics

[4] A GLS-weighting approach (Anderson, 2008) 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.

(gender, type and day-shift of the school of origin, previous experience with practice exams providing feedback, aspirations to attend higher education, 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 and Ding, 2024).

The parameter α_1 measures the average treatment effect of receiving the performance feedback on the outcome Y_i, while α_3 captures how students differentially respond to the feedback in terms of the achievement index, A_i. This specification captures the fact that inaccurate beliefs about academic proficiency can shape the perceived value of attending a given high-school program. Hence, providing performance feedback can potentially alter those beliefs as well as the slope of students' choices and outcomes with respect to their actual academic readiness.

We estimate the parameters of equation (1) by OLS. 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).[5]

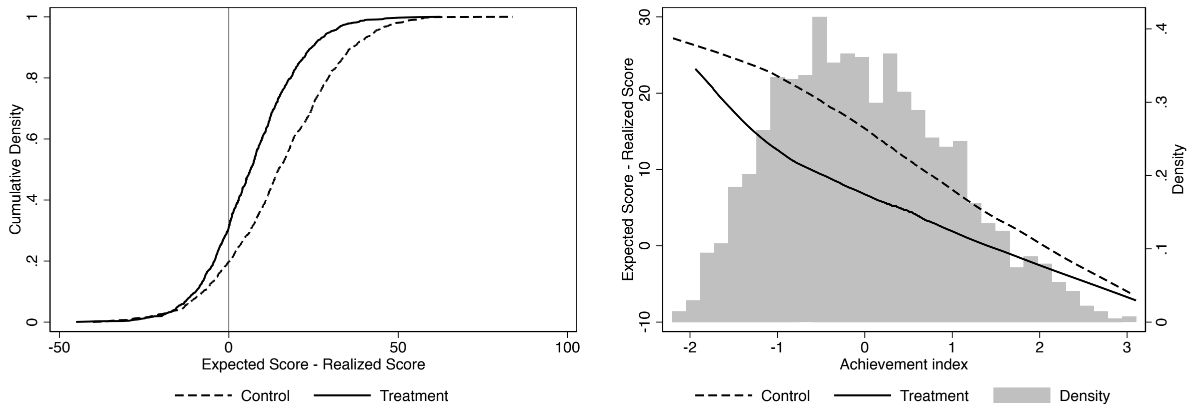
3.2 Subjective Expectations about Test Performance

Panel A in Figure 3 displays the cumulative distributions of the perception gap, defined as the difference between the expected score and the realized performance in the mock exam, for students in the treatment and control groups. In the control group, over 80% of the applicants overestimate their performance in the test. The performance feedback substantially shifts to the left the distribution of the perception gap, with an average gap of 6.5 points for treated applicants and 14.7 points for control applicants (out of a 128-point scale). Panel B in Figure 3 presents evidence on the relationship between the perception gap and academic achievement (A_i) for students in the treatment and control groups. Updates on the expected score in response to performance feedback occur along the entire distribution of the achievement index, with relatively larger gap reductions among lower performing students, who are those who display larger prior biases.

Table 2 shows the OLS estimates of the effect of the information intervention, as depicted by equation (1), on different moments of the individual distribution of beliefs about test

[5]The Romano-Wolf correction (Romano and Wolf, 2005a,b, 2016) asymptotically controls the family-wise error rate, that is, the probability of rejecting at least one true null hypothesis among a family of hypotheses under test. This correction is considerably more powerful than earlier multiple-testing procedures, given that it takes into account the dependence structure of the test statistics by re-sampling from the original data.

Figure 3: Gap between Expected Scores and Realized Scores in the Mock Exam



(a) Cumulative Density of the Gap

(b) Gap and Academic Achievement

NOTE: Assuming a uniform distribution within each interval of the score, the expected scores are constructed as the summation of the mid-values in each discrete interval of the support multiplied by the associated probability assigned by the student. Panel A shows the cumulative density of the difference between the expected scores and the realized scores in the mock version of the admission exam. Panel B shows non-parametric locally weighted estimates of the relationship between the perception gap and the achievement index. For more details on the elicitation of beliefs in the survey data, refer to Appendix A.

performance. The first column documents that providing feedback about test performance decreases the mean of the belief distributions by 6.9 points out of a sample average of 75.6 in the control group (p-value = 0.001). We find a similar effect when we alternatively consider the median of the individual belief distributions in the second column, with an 11% drop relative to the corresponding sample average in the control group (p-value = 0.001). The negative treatment effect on expected performance in the mock test is partly attenuated by the positive updating patterns among the highest performing students. An increase of one standard deviation in the achievement index corresponds to a right-shift in the location of the belief distribution by 2.9-3.1 points in the treatment group relative to the control.

Finally, the estimates reported in the last two columns of Table 2 document that the performance feedback meaningfully decreases the uncertainty of students' predictions, with average reductions in the standard deviation/inter-quartile range of the belief distributions of 2.8 points (p-value = 0.001), or 11-16% of the average in the control group. We find limited evidence of heterogeneous updating on the second moment of the belief distributions by the value of the achievement index.

These findings establish that providing information about individual performance in the mock exam allows applicants to substantially revise their expectations about academic readiness. The evidence underscores the informativeness of the performance feedback for all students in our sample. However, the magnitude and direction of the adjustment is strongly

Table 2: Performance Feedback and Beliefs about Test Performance

	Mean	Median	Std. Dev.	IQR
Treatment	-6.935	-8.892	-2.773	-2.789
	[0.000]	[0.000]	[0.000]	[0.000]
	{0.001}	{0.001}	{0.001}	{0.001}
Achievement index	4.550	4.839	-0.621	-1.234
	[0.000]	[0.000]	[0.021]	[0.011]
	{0.001}	{0.001}	{0.010}	{0.007}
Treatment × Achievement index	2.908	3.109	-0.441	-0.875
	[0.000]	[0.000]	[0.232]	[0.193]
	{0.001}	{0.001}	{0.106}	{0.106}
Mean Control	75.6	78.8	17.4	24.2
Number of Observations	2246	2246	2246	2246
Number of Clusters	90	90	90	90
R-squared	0.289	0.282	0.082	0.057

NOTE: The dependent variable "Mean" is constructed as the summation of the mid-values in each discrete interval of the support multiplied by the associated probability assigned by the student. The dependent variable "Median" is defined as the midpoint of the interval in which the cumulative density first surpasses 0.5 (11/20 beans or more). The dependent variable "Std. Dev." is constructed as the square root of the summation of the mid-values in each discrete interval of the support multiplied by the square of the associated probability assigned by the student minus the square of the constructed mean. The dependent variable "Inter-Quantile Range (IQR)" is defined as the difference between the midpoints of the intervals that accumulate 75 percent and 25 percent of the probability mass. For more details on the elicitation of beliefs in the survey data, refer to Appendix A. The achievement index is a GLS-weighted average (Anderson, 2008) 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 in curly brackets are adjusted for testing each null hypothesis across multiple outcomes through the step-wise procedure described in Romano and Wolf (2005a,b, 2016). Both inference procedures take into account the clustered structure of the individual error terms at the middle school level.

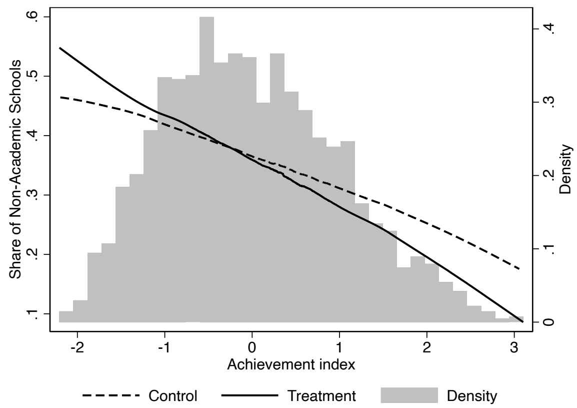
associated with their level of academic achievement.[6]

3.3 School Choices and Placement Outcomes

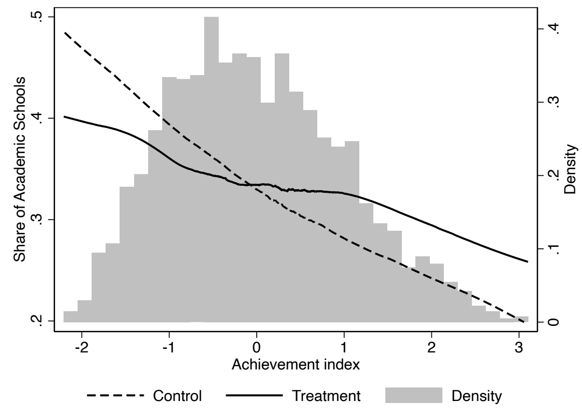
Figure 4 presents evidence on the relationship between the track-composition of the school rankings and academic achievement. We use the share of academic or non-academic schools in the applicants' ranked-order lists as a proxy of their valuations for a high-school track. While, on average, the provision of performance feedback does not affect applicants' choices across high school tracks, it clearly affects the slope of the relationship between track choices

[6] In our companion paper (Bobba and Frisancho, 2022), we explore in further details the process of belief updating spurred by the performance feedback.

Figure 4: Performance Feedback and High-School Track Choices



(a) Share of Non-Academic Schools in Ranking



(b) Share of Academic Schools in Ranking

NOTE: This figure depicts the density of the achievement index (y-axis on the right), which is a GLS-weighted average (Anderson, 2008) of the GPA in middle school, mock exam score, and exam score. Overlaid on the density, we show non-parametric locally weighted estimates of the relationship between the share of non-academic (left panel) and academic/non-elite (right-panel) schools in the applicant's ranking and the performance index separately for applicants in the treatment and control groups (y-axis on the left).

and academic achievement.[7]

This composition effect in the valuations for non-academic and academic schools can potentially alter the sorting patterns realized under the centralized school mechanism. Placement in non-academic schools decreases along the distribution of academic achievement for the students in our sample. As shown in the first column of Table 3, 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 second column in Table 3 presents the estimated effects of the information intervention on placement into (non-elite) academic programs. Consistently with the school choice responses depicted in Figure 4, 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. This effect corresponds to a 10 percent increase when compared to the average probability of assignment in academic programs for the students in the control group. On average, placement in non-academic programs increases

[7] As further shown in Table B.5 in the Appendix, a one-standard-deviation increase in academic achievement for the students in the treatment group decreases (increases) the share of non-academic (academic) schools requested by approximately 3 percentage points. These effects are precisely estimated (p-values$\leq$0.01) and sizable, as they correspond to about 10 percent of the corresponding share of same-track options in the school rankings of applicants in the control group.

Table 3: Performance Feedback and Placement Outcomes

	Non-Academic	Academic	Elite
Treatment	0.046	-0.044	-0.002
	[0.077]	[0.078]	[0.861]
	{0.110}	{0.110}	{0.842}
Achievement index	-0.079	-0.086	0.165
	[0.000]	[0.000]	[0.000]
	{0.002}	{0.001}	{0.001}
Treatment × Achievement index	-0.065	0.041	0.024
	[0.015]	[0.045]	[0.247]
	{0.020}	{0.065}	{0.253}
Mean Control	0.453	0.418	0.129
Number of Observations	2493	2493	2493
Number of Clusters	90	90	90
R-squared	0.100	0.061	0.336

NOTE: The dependent variable is an indicator variable that is equal to one if the applicant is assigned to a given group of schools (i.e., Non-Academic, Academic, and Elite schools). The achievement index is a GLS-weighted average (Anderson, 2008) 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 and Wolf (2005a,b, 2016). Both inference procedures take into account the clustered structure of the error terms at the middle school level.

by 4.6 percentage-points (p-value=0.11), while it decreases in academic programs by roughly the same amount.

The estimates reported in the last column of Table 3 show that the provision of performance feedback does not systematically alter assignment into elite schools. This is consistent with the fact that the valuation for elite programs is relatively inelastic to information about students' own academic skills (see the third column of Table B.5 in the Appendix). Given the relatively large changes in beliefs established above, the evidence on elite schools 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.

Finally, exposure to performance feedback does not systematically affect the scores in the admission exam (see column 2 in Table B.2 in the Appendix), indicating that most of the treatment effect on school placement is likely driven by the observed differential changes in the school choices. The performance feedback does not systematically alter the fraction of applicants assigned in the first or second round of the assignment process (see Table B.6 in the Appendix).

Taken together, 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 placement outcomes are small and statistically insignificant, higher performing students in the treatment group are less likely to be placed in non-academic (vocational or technical) high-school programs, and they are more likely to sort into (non-elite) academic programs.

3.4 Educational Trajectories

As shown in the previous sub-section, the provision of performance feedback likely improved the alignment between (measured) academic skills and high school track choices. The associated change in the sorting patterns of students across schools may therefore potentially alter students' academic trajectories and improve downstream educational outcomes. Given the relatively high returns of a secondary schooling degree in the labor market and other life-cycle outcomes (Heckman et al., 2018; Psacharopoulos and Patrinos, 2018), we consider high-school graduation as our main educational outcome.

The point estimates reported in the first two columns of Table 4 document that there are no discernible differences in the high school enrollment rates or the dropout rates during the first year of high school between students in the treatment and control groups. These findings indicate limited short-run consequences of the reallocation effect across high-school document above on educational trajectories.

However, the information intervention does seem to play a role in the medium term. The results reported in the third column of Table 4 show that the probability of graduating on time (conditional on placement though not on enrollment in tenth grade) is 5.4 percentage points higher for students who receive performance feedback and who score one-standard deviation below the mean (=2.2p.p+3.2p.p) when compared to equally achieving students who do not receive any feedback. While statistically imprecise, the magnitude of this effect is economically significant, as it corresponds to a 13 percent increase in high-school graduation rates when compared to the sample mean in the control group of 0.45. The effect size on the rate of high-school graduation roughly coincides with the magnitude of the impact of a one-deviation increase of the score in the mock exam (see Section 2.4).[8]

Figure 5 visually displays the relationship between the rates of graduation on time from

[8]We were unable to obtain the high-school graduation records for approximately 5% of the students in our sample, which explains the discrepancy in the number of observations between column 1 and column 3 of Table 4. The associated Lee bounds (Lee, 2009) are narrow and broadly consistent with the point estimates reported in the main text (see Appendix Table B.7).

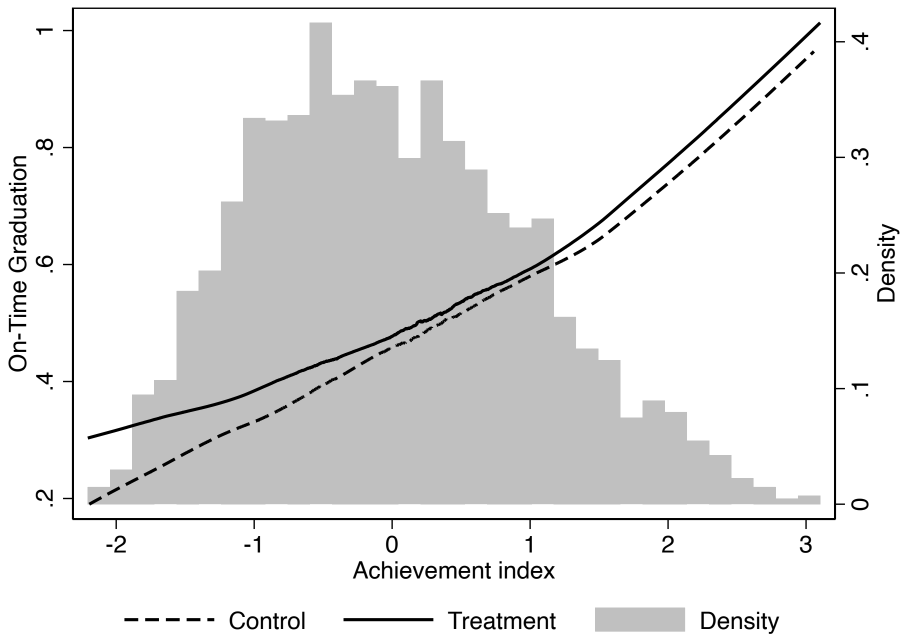
Table 4: High School Outcomes

	Enrollment	Dropout 1st year	Graduation on Time
Treatment	-0.003	0.012	0.022
	[0.789]	[0.668]	[0.252]
	{0.936}	{0.920}	{0.497}
Achievement index	0.068	-0.095	0.138
	[0.000]	[0.001]	[0.000]
	{0.001}	{0.003}	{0.001}
Treatment × Achievement index	-0.021	-0.006	-0.032
	[0.352]	[0.807]	[0.088]
	{0.590}	{0.936}	{0.198}
Mean Control	0.813	0.248	0.447
Number of Observations	2493	2024	2358
R-squared	0.045	0.076	0.090

NOTE: 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 placement in tenth grade and zero otherwise. The achievement index is a GLS-weighted average (Anderson, 2008) 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 and Wolf (2005a,b, 2016). Both inference procedures take into account the clustered structure of the error terms at the high school level.

secondary education and the achievement index separately for the treatment and control groups. While there is a small effect of performance feedback along the entire distribution of academic achievement, its impact on schooling trajectories becomes more clearly visible around the left tail. Since under-achieving students also tend to have lower graduation rates, the information intervention effectively contributes to "levelling the playing field" in our setting. Importantly, the observed gains in persistence throughout secondary education do not seem to be explained by the fact that lower-performing students tend to sort into easier-to-graduate schools as a result of the information intervention (see Table B.8 in the Appendix).

Figure 5: On-time Graduation and Academic Performance



NOTE: This plot depicts non-parametric locally weighted estimates of the relationship between the graduation on time and the achievement index, which is a GLS-weighted average (Anderson, 2008) of middle school GPA, mock exam score, and exam score. "On-Time Graduation" denotes an indicator variable that is equal to one if the student successfully completes the assigned high school program in three years after placement in the centralized system, and zero otherwise.

4 Scaling-up the Information Experiment

The evidence reported in the previous section suggests that it may be possible to improve the allocation of skills across high-school tracks by providing students with information about their academic proficiency during the transition from lower to upper secondary education. In this section, we embed the randomized evaluation in an empirical model of student sorting across schools and educational trajectories in order to characterize the effect of scaling up the information intervention.

4.1 Preferences Over School Characteristics

We rely on a Conditional Logit model with parameters that vary according to observed student characteristics. We model the indirect utility that student i gets from attending school j as:

$$u_{ij} = \alpha_{s(j)} + \boldsymbol{\beta}'_{\boldsymbol{s(j)}}\boldsymbol{x_i} + \boldsymbol{\gamma}'\boldsymbol{x_i}d_{ij} + \boldsymbol{\rho}'\boldsymbol{x_i}c_j + \epsilon_{ij}, \tag{2}$$

where the composite term $\alpha_{s(j)} + \boldsymbol{\beta}'_{\boldsymbol{s(j)}}\boldsymbol{x_i}$ denotes the net returns of attending a particular "college" s, or groups of high-school programs j that share the same track (non-academic,

academic, or elite) and that belong to the same public institution of upper secondary education. The size of the experimental sample is too small to precisely estimate the 600+ school-specific intercepts and the associated interaction terms. We thus group the high-school programs in the centralized system into 16 college-specific intercepts, $\alpha_{s(j)}$, and the associated student-college match effects, $\beta_{s(j)}$. By doing so, we substantially reduce the number of parameters that need to be estimated in equation (2).

Vector $\boldsymbol{x_i}$ contains an array of standardized individual characteristics, which broadly capture skill measures and demographics (observed or unobserved to the applicant), such as the score in the mock exam, the cumulative GPA in middle school, an index of socio-economic conditions in the neighborhood of residence of the applicants, the average ENLACE math score in the students' middle school of origin, and parental education. The same vector $\boldsymbol{x_i}$ of individual characteristics is also interacted with two school-level characteristics— the geodesic distance d_{ij} (in kilometers) between the location of the middle school of applicant i and high-school program j as well as with the degree of selectivity of each high-school program c_j, which we measure through the admission equilibrium cutoff score in the previous round of the assignment mechanism.

The vector of parameters $\boldsymbol{\gamma}$ captures the average commuting cost of attending a particular high school program while accounting for potential heterogeneity across applicants. Analogously, the parameter vector $\boldsymbol{\rho}$ embeds any utility cost or benefit associated to being assigned to schools with a given level of peers' quality and/or academic requirements. Tuition fees are negligible in this setting and they do not vary between schools in the same college s, so that the small differences in the out-of-pocket expenses across high-school programs are captured by the $\alpha_{s(j)}$ parameters.

Conditional on $\boldsymbol{x_i}$, d_{ij} is assumed orthogonal to the preference shock ϵ_{ij}. The preference shock is assumed to be i.i.d. across i and j, following a type-I extreme value distribution with normalized scale and location. This assumption is usually invoked in the school choice literature (see, e.g., Agarwal and Somaini, 2020). It is violated if students systematically reside near the schools for which they have idiosyncratic tastes. Conditional on the rich set of controls that we introduce, this assumption becomes plausible 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.

Since discrete choice models depend on differences in payoffs, we normalize students' mean utility of not being assigned to any school program within the assignment system to zero. This outside option further captures the value of not attending high-school, or the value of any other labor market entry opportunity not directly observed in the data.

4.2 Estimating Preferences

We have access to individual-level data on rank-ordered lists and placement outcomes. Both sources of information are potentially valuable for estimating preference parameters. However, school rankings may deviate from true preference orderings due to the cap of 20 schools in the submitted rank-ordered lists (Haeringer and Klijn, 2009; Calsamiglia et al., 2010) or possible strategic mistakes in applications (Hassidim et al., 2017; Artemov et al., 2023). A more robust estimation approach relies on the assumption that the realized matching equilibrium is stable, which is likely satisfied in the large-market matching mechanism that we 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). These choice sets solely depend on the scores in the admission exam for most programs.[9]

The parameters of the indirect utility function in (2) are estimated by maximum likelihood separately for the treatment and the control samples. This strategy flexibly allows for the possibility that feedback provision can alter the choice environment in which applicants operate. Table 5 shows the estimates of student-school match effects for high-school tracks and programs' selectivity that are re-scaled by the disutility of the distance to the various high-school options (or willingness to travel). The full set of estimated parameters in utility values for the more flexible specification at the college-level is reported in Table B.9 in the Appendix.

The estimates of preferences over school characteristics for the applicants who received the performance feedback differ somewhat from those for the applicants in the control group, as shown in the third column of Table 5. Relative to the control group, higher-SES students (i.e. those who reside in more socio-economically advantaged neighborhoods and/or who come from higher quality middle schools) who receive the performance feedback attach a more negative value to an elite school. The size of the estimated effect for a one-standard-deviation increase in the neighborhood SES index is equivalent to commuting to a school that is 8.5 km away when compared to the applicants at the mean of the SES distribution. The corresponding commuting cost for the applicants in the control group is 1.5 km, which is

[9]Elite schools further impose a GPA requirement of at least 7 out of 10 points. However, most of the applicants to those high school programs meet this requirement (more than 90 percent in every round of the assignment system).

Table 5: Willingness to Travel for School Characteristics

	Control Sample WTT Est (Std. Err.)	Treated Sample WTT Est (Std. Err.)	Control-Treated T-test [p-value]
$[\text{Elite}]_j \times [\text{Mock Score}]_i$	-0.2144	-2.3344	1.0516
	(1.3191)	(1.5246)	[0.2931]
$[\text{Elite}]_j \times [\text{SES Index}]_i$	-1.5265	-8.4959	2.0903
	(1.8019)	(2.8053)	[0.0367]
$[\text{Elite}]_j \times [\text{Middle-School Math Score}]_i$	0.7386	-2.9515	1.7464
	(1.2932)	(1.6710)	[0.0809]
$[\text{Academic}]_j \times [\text{Mock Score}]_i$	0.6327	-1.4862	2.0833
	(0.6604)	(0.7735)	[0.0373]
$[\text{Academic}]_j \times [\text{SES Index}]_i$	-0.8522	0.2950	-0.6901
	(0.9252)	(1.3809)	[0.4902]
$[\text{Academic}]_j \times [\text{Middle-School Math Score}]_i$	-0.0962	0.2064	-0.2664
	(0.6732)	(0.9144)	[0.7899]
$[\text{Non-Academic}]_j \times [\text{Mock Score}]_i$	1.1555	-1.1870	2.3631
	(0.6517)	(0.7470)	[0.0182]
$[\text{Non-Academic}]_j \times [\text{SES Index}]_i$	-0.2780	-1.9743	-1.0007
	(0.9432)	(1.4085)	[0.3171]
$[\text{Non-Academic}]_j \times [\text{Middle-School Score}]_i$	0.2538	-0.4473	0.6388
	(0.6673)	(0.8713)	[0.5230]
$[\text{Cutoff Score}]_j \times [\text{Mock Score}]_i$	0.2707	0.6569	-1.0162
	(0.2378)	(0.2965)	[0.3097]
$[\text{Cutoff Score}]_j \times [\text{SES Index}]_i$	0.7705	1.0730	-0.4850
	(0.3479)	(0.5175)	[0.6277]
$[\text{Cutoff Score}]_j \times [\text{Middle-School Math Score}]_i$	0.1628	1.2866	-2.3816
	(0.2887)	(0.3733)	[0.0173]

NOTE: This table displays maximum-likelihood estimates that are normalized by the distance coefficient for selected match coefficients of equation (2). Standard errors reported in parenthesis are computed using the delta method. The third column displays the t-statistics and the associated p-values (in brackets) for the null hypotheses of equal coefficients between the control and the treated samples. The full set of model estimates at the college-level is reported in Table B.9.

not statistically different from zero. The willingness to travel of attending a non-elite school (academic or non-academic) in the treatment group also features a negative gradient with respect to students' academic achievement, as measured by the score in the mock test. The magnitude of these estimates is equivalent to a commuting cost of 1-1.5 km, although it is not always statistically different from the corresponding estimates for the applicants in the control group.

As shown by the results reported at the bottom of Table 5, the willingness to travel attached to the degree of selectivity of the schools—as measured by the cutoff scores from the previous year—shows a positive gradient with respect to both academic achievement and SES, which is steeper for the applicants in the treatment group when compared to those in the control group. This last finding may help to reconcile the experimental evidence discussed in Section 3.3 with the estimates of the school choice model. For instance, the positive and negative gradients tend to offset each others for treated students' valuations of elite schools thereby explaining the lack of both average and heterogeneous effects of the intervention on the share of those schools in the school rankings (see Figure B.3).

4.3 Out-of-Sample Sorting Predictions

We use the parametrization of the indirect utility function (2) in order to extrapolate the sorting patterns across schools in the evaluation sample to the universe of applicants in the centralized assignment system. We have access to student-level data on assignment outcomes and the individual characteristics contained in the vector $\boldsymbol{x_i}$ for all the applicants. Since the applicants outside the evaluation sample do not take our mock exam, we replace it in the expression for the indirect utility with the admission exam score. We run the Serial Dictatorship algorithm that is in place in the assignment system relying on the priority criteria and school capacities. Since school preferences are vertical (i.e., school programs simply accept or reject prospective applicants in descending order based on their exam scores until seat capacities are met), this algorithm delivers the unique stable matching equilibrium allocation (Roth and Sotomayor, 1992).[10]

In Table 6 we compare the average outcomes of school assignment from the data with those based on the matching equilibrium computed using the estimated preferences of the applicants in the control group. Mean-differences are very small for the outcomes considered across the entire SES distribution. This result was not guaranteed *a priori*, given the fact that the experiment is targeted toward applicants from relatively disadvantaged backgrounds (see Section 2.3). Another way to assess the validity of the extrapolation is by looking at the equilibrium cutoff scores. The linear correlation between the observed cutoff scores and the model-based cutoff scores is 0.88. Figure 6 provides a scatter plot of the relationship between the cutoff scores in the model and in the data for the schools in the assignment mechanism. As expected, the fit of the model improves for more selective options with high cutoff scores—mostly elite schools, but also for a few academic and non-academic options. These schools are more likely to be oversubscribed, which implies that the associated cutoff

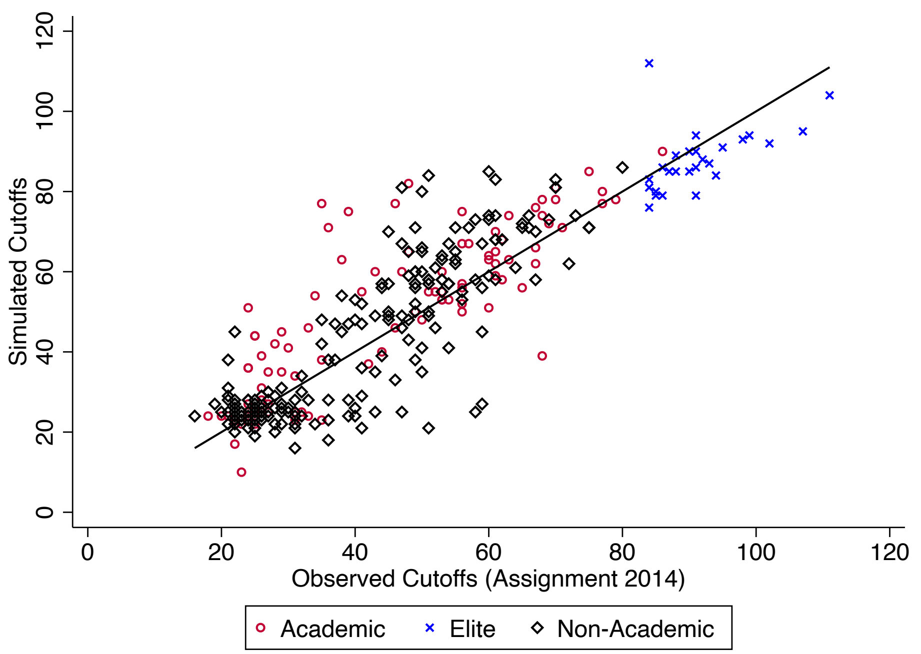
[10] While we estimate the school choice model by assuming stable matching but not truth-telling (see Section 4.2), we can allow students to be truthful when studying matching outcomes. This holds as long as preference estimates are consistent (Artemov et al., 2023).

Table 6: Model Fit on Average Assignment Outcomes by SES Categories

	Very Low SES		Low SES		Middle SES		High SES	
	Data	Model	Data	Model	Data	Model	Data	Model
Applied in the system (1=yes)	1.00	0.97	1.00	0.99	1.00	0.99	1.00	1.00
Assigned in the system (1=yes)	0.91	0.91	0.88	0.92	0.86	0.93	0.84	0.94
Non-Academic schools, vocational track	0.16	0.18	0.14	0.13	0.13	0.10	0.10	0.08
Non-Academic schools, technical track	0.30	0.27	0.27	0.27	0.25	0.25	0.22	0.23
Academic, above-median selectivity	0.23	0.20	0.30	0.28	0.30	0.31	0.32	0.33
Academic, below-median selectivity	0.17	0.21	0.09	0.13	0.04	0.08	0.03	0.04
Elite schools	0.13	0.14	0.20	0.18	0.28	0.26	0.34	0.32
Selectivity (z-cutoff score)	0.32	0.24	0.65	0.56	0.96	0.90	1.21	1.15

NOTE: The averages displayed in the first column are computed from the data of the assignment mechanism in the year 2014 (see Section 2). The averages displayed in the second column are computed by running the Serial Dictatorship algorithm that is in place for the COMIPEMS system, using the estimated preferences across schools of the control group, the individual scores in the admission exam, and the school capacities as inputs.

Figure 6: Model Fit on Cutoff Scores



NOTE: In this figure, we report the cutoff scores for the 429 school programs (68% of the total participating programs) that are contained in the choice sets of the applicants of the evaluation sample. For an analogous chart with the cutoffs of all the 628 school programs, refer to Figure B.4 in the Appendix. The observed cutoffs are computed from the data of the assignment mechanism in the year 2014 (see Section 2). The simulated cutoff scores displayed in the scatter plot are computed by running the Serial Dictatorship algorithm that is in place for the COMIPEMS system using the estimated school preferences of the control group, the individual scores in the admission exam, and the school capacities as inputs.

scores are well-defined equilibrium objects under stable matching (Azevedo and Leshno, 2016; Fack et al., 2019).

Taken together, this evidence broadly supports the validity of the out-of-sample prediction based on the estimated preferences of the experimental control group and the parametric linear form of the indirect utility function (2). Therefore, we postulate that the corresponding predictions based on the estimated preferences of the applicants in the treatment group likely approximate a counterfactual scenario in which the broader population of applicants would be given additional information about their academic skills. This is akin to implementing a policy that mandates the universal implementation of a mock exam or, alternatively, disclosing admission exam scores to the applicants before the submission of the rank-ordered lists (see Figure 2).

4.4 Linking Sorting with Education Outcomes

We consider a potential outcomes framework that maps any student-school match into educational outcomes. In particular, we posit that the potential outcome of student i if she is matched to school j can be written as:

$$Y_{ij} = \delta_{s(j)} + \boldsymbol{\gamma'_{s(j)}} \boldsymbol{x_i} + \nu_{ij}, \tag{3}$$

where, as before, $s(j)$ denotes a particular college (i.e., group of high-school programs). The vector $\boldsymbol{x_i}$ contains the same standardized individual characteristics of equation (2) except for the mock score (i.e. the cumulative GPA in middle school, an index of socio-economic conditions in the neighborhood of residence of the applicants, the middle school-average of ENLACE math test scores, and parental education). In this framework, $\delta_{s(j)}$ measures the average effect of college s, $\boldsymbol{\gamma_{s(j)}}$ corresponds to the vector of match effects for students with observed type $\boldsymbol{x_i}$ in college s, and ν_{ij} denotes any unobserved factor influencing education outcomes.

Students are not randomly assigned to schools or colleges. However, school placement under the assignment mechanism depends exclusively on two student-level observable factors: ranked-order lists (ROL) and the scores in the admission exam. We leverage this feature in order to obtain consistent estimates of our parameters of interest (Angrist and Rokkanen, 2015; Abdulkadiroglu et al., 2020). In particular, we add ROL fixed effects in equation (3), $r_k = \mathbb{1}[R_i = k]$ with R_i denoting student i's ROL. Under the serial dictatorship, the only part of each student's ROL that ultimately plays a role in the allocation is the subset of schools ranked in cut-off descending order, $\tilde{R}_i \subseteq R_i$. Given the relatively long size of the submitted ROLs (10 schools on average), we use $\tilde{R}_i$ instead of R_i to construct the ROL fixed effect r_k in order to increase the number of students who share the same school rankings.

Conditioning on the vector of skill measures and demographics $\boldsymbol{x_i}$, and the school rankings

$\tilde{R}_i$, we assume that the remaining variation in school placement is due to idiosyncratic differences in the score of the admission test (e.g., a good or a bad exam day) that are assumed to be uncorrelated with potential outcomes:

$$\mathbb{E}[Y_{ij} \mid \boldsymbol{x_i}, \tilde{R}_i] = \delta_{s(j)} + \boldsymbol{\gamma'_{s(j)}} \boldsymbol{x_i} + r_k. \tag{4}$$

A new matching equilibrium can potentially affect the characteristics of the peers to whom a student is exposed. We allow for equilibrium changes in peer composition to affect education outcomes by taking advantage of the fact that the parameters of the value added model (3) vary at the college-level. Therefore, we can include average peer characteristics at a more granular level. Under this specification, the matching equilibrium associated to the information intervention will also affect outcomes through changes in the vector $\boldsymbol{\bar{x}_j}$ of average skills and demographics at the school level.

Consistent with the experimental evidence discussed in Section 3.4, we consider as our education outcome the on-time graduation from the assigned high school program in the centralized system. We use as a proxy for graduation the attendance to the ENLACE twelfth-grade test. The test was discontinued in 2014, and so we estimate the parameters of the value added model (3) for an earlier cohort of assigned applicants (see Section 2.4). Based on these estimates, we predict the education outcomes for the cohort of the experiment. The predictions are based on the characteristics of the applicants and their sorting patterns across schools, both under the status-quo and the new matching equilibrium under the performance feedback.

Table 7 displays the OLS estimates for selected coefficients of the value added model aggregated at the track-level. The first column shows the estimates of a specification without ROL fixed effects, while the second column reports the results of the specification shown in equation (4). Table B.10 in the Appendix displays the full set of OLS coefficients at the college-level for our preferred specification with ROL fixed effects. On average, attending an elite high-school program decreases on-time graduation rates by 18 percentage points when compared to both an academic or a non-academic program out of a basis of 41 percent graduation rate after high-school assignment. Some of the match effects based on students' skills and demographics are statistically significant and their sign makes intuitive sense, although they are much less important than the corresponding average effect of attending a given track in explaining education trajectories throughout high school.

Table 7: Estimates of the Value Added Model (On-time Graduation)

	OLS	OLS with ROL fixed effects
$[\text{Elite}]_j$	-0.192	-0.175
	(0.006)	(0.024)
$[\text{Elite}]_j \times [\text{GPA}]_i$	0.081	0.069
	(0.004)	(0.008)
$[\text{Elite}]_j \times [\text{SES index}]_i$	0.027	0.014
	(0.004)	(0.009)
$[\text{Elite}]_j \times [\text{Parent Education}]_i$	0.005	0.002
	(0.003)	(0.006)
$[\text{Academic}]_j$	0.027	-0.002
	(0.003)	(0.013)
$[\text{Academic}]_j \times [\text{GPA}]_i$	0.020	0.011
	(0.002)	(0.005)
$[\text{Academic}]_j \times [\text{SES index}]_i$	0.001	0.009
	(0.002)	(0.006)
$[\text{Academic}]_j \times [\text{Parent Education}]_i$	-0.002	-0.006
	(0.003)	(0.006)
Number of Observations	182,824	182,824

NOTE: This table displays OLS estimates and asymptotic standard errors (in parenthesis) for selected coefficients of equation (3). The full set of estimates is reported in Table B.10 in the Appendix.

5 Feedback Provision at Scale

In this section, we leverage the empirical framework outlined in Section 4 in order to quantify the impact of a counterfactual at-scale implementation of the information intervention. We first uncover the effect of feedback provision on school admission taking into account the equilibrium effect that naturally arises within the centralized assignment system. We then study the consequences on educational outcomes of the equilibrium allocation under the information intervention when compared to the status quo.

5.1 School Choices and Placement Outcomes

We compare the predictions of the choice model (2) across all the applicants using the estimated parameters of the control group with the corresponding predictions based on the estimates of the treatment group. We start by discussing a few features of the equilibrium allocations under the two scenarios, which are reported in Table 8. First, there are no changes at the extensive margin of the admission process. With the additional information

Table 8: The Effect of the Information Intervention on Aggregate Outcomes

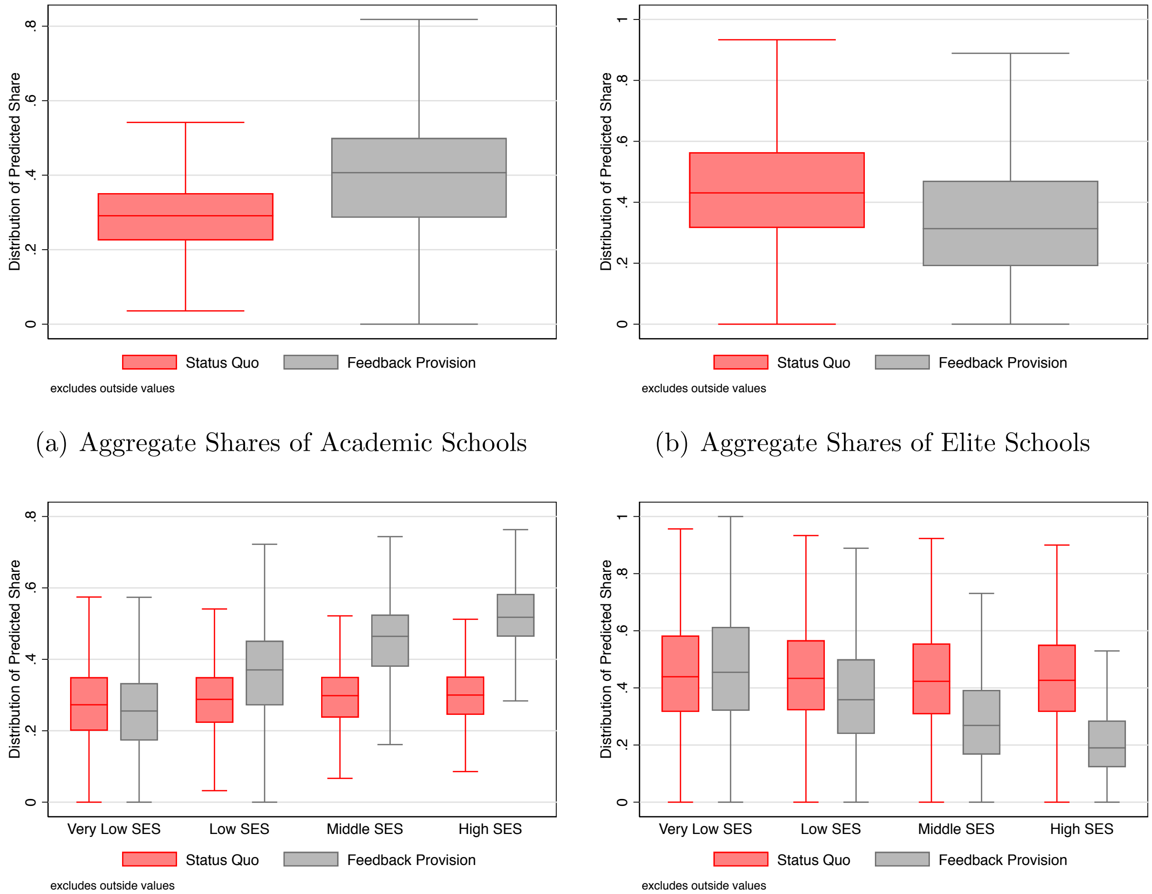
	Status Quo	Information Intervention	Difference
Applied in the system (1=yes)	0.99	0.99	0.00
Assigned in the system (1=yes)	0.89	0.91	0.02
Rank of assigned school	6.41	5.43	-0.98
Assigned in top choice	0.16	0.25	0.09
Assigned in elite schools	0.22	0.22	0.00
Assigned in academic schools	0.41	0.40	-0.01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

NOTE: The average outcomes displayed in the first column are obtained by running the Serial Dictatorship algorithm on the estimated school valuations of the experimental control group, the individual scores in the admission exam, and the school capacities. The average outcomes displayed in the second column are computed by running the Serial Dictatorship algorithm on the estimated school valuations of the experimental treatment group, the individual scores in the admission exam, and the school capacities as inputs.

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 that directly maps to the efficiency of the new matching equilibrium. Third, and perhaps more importantly for efficiency, students are more likely to get assigned to their preferred options. When moving from the status quo to the counterfactual information policy, the average applicant is placed in a school that is one position above her school ranking (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.

Figure 7 displays the distribution of the extrapolated preference over high-school tracks. Under the two scenarios with and without feedback provision, we compute the predicted shares of academic and elite schools among the school programs that give each applicant higher utility than the outside option. Panels A and B show that the information intervention leads to an overall increase in the demand for academic schools and a symmetric decrease in the demand for elite schools. The lower demand-side pressure on elite programs is likely to spur some reallocation effects within the system, possibly towards (less selective) academic programs. Those can be seen through the changes in the equilibrium cutoff scores, as depicted in Figure B.5 in the Appendix. Under the new matching equilibrium associated to the information policy, the cutoff scores for most of the elite schools slightly decrease when compared to the status quo. 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 cutoff scores, which is consistent with the limited average impact on pref-

Figure 7: The Effect of Providing Performance Feedback on Track Choices



(a) Aggregate Shares of Academic Schools

(b) Aggregate Shares of Elite Schools

(c) Shares of Academic Schools by SES

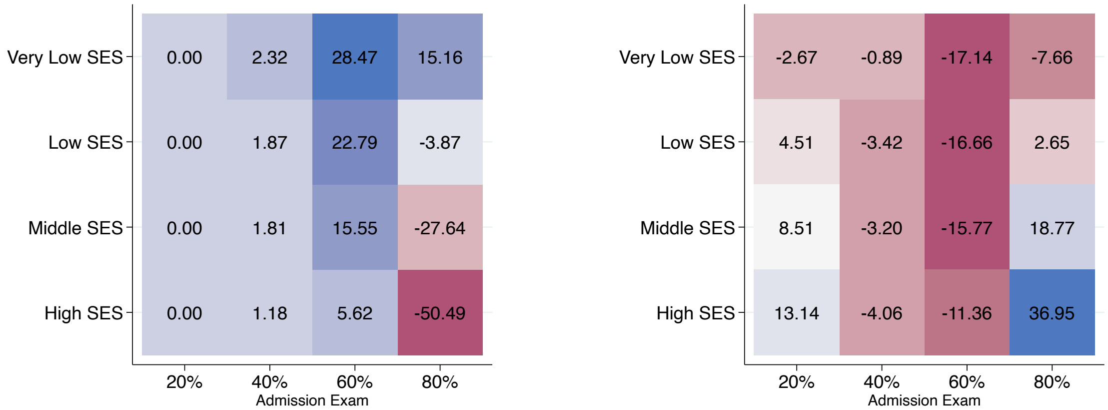
(d) Shares of Elite Schools by SES

NOTE: This figure displays box- and-whisker plots for the shares of academic schools and elite schools as implied by the model estimates for the control group (red bars) and for the treatment group (grey bars). 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).

erences toward this high school track. The aggregate changes in the demand for academic schools and elite schools, together with the associated movements in the cutoff scores, can explain the muted effect of the intervention on the average sorting patterns across high school tracks (see the last three rows of Table 8).

Panels C and D in Figure 7 display the same effects on the predicted shares of academic and elite schools under the two scenarios, but they are broken down by discrete categories of socio-economic conditions in the neighborhood of residence of the applicants (our SES index). The overall effect of the information intervention observed in the upper-panels can be mostly explained by the associated changes in the distribution of school preferences for relatively

Figure 8: The Effect of Providing Performance Feedback on High-School Admission (Percentage Points)



(a) Elite Schools (b) Academic Schools

NOTE: This figure shows the percentage changes between the Information Policy and the Status Quo scenarios in the shares of applicants that are assigned to academic schools (Panel A) and elite schools (Panel B) by discrete categories of socio-economic status (Y-axis) and the score in the admission exam (x-axis).

better-off applicants—i.e., those who live in neighborhoods with low poverty levels. This pattern can be traced back to the estimates of the school choice model, as shown in Table 5, which featured a marked negative gradient with respect to the socio-economic status for the applicants in the treatment group in terms of the value of attending an elite school. Note that our evaluation sample focused on disadvantaged students. Thus, the fact that low-SES applicants are, on average, unresponsive to the intervention in terms of their high-school track choices is consistent with the evidence discussed in Section 3.3.

To sum up, scaling up the information intervention would alter the socio-economic composition of the student population admitted to elite schools. The reduction in the demand for elite programs among the relatively better-off applicants documented above would necessarily leave some open seats in those programs for disadvantaged students with high admission scores. Figure 8 shows that this is precisely what happens in the new matching equilibrium. The information intervention increases the representation of high-achieving students from low-SES neighborhoods in elite schools by more than 20 percentage points (see Panel A). Symmetrically, Panel B in the same figure shows that high-SES students in the top quintile of the score distribution are 37 percentage-points more likely to get placed in a academic (non-elite) high school programs. These students are less likely to get admitted in elite programs by 50 percentage-points (see Panel A).

The equilibrium effect of the information intervention on elite admission for low-SES applicants could not possibly be detected in our small-scale evaluation. Indeed, the experimental evidence documents no effect of feedback provision on the probability of assignment to elite schools, as shown in the last column of Table 3.

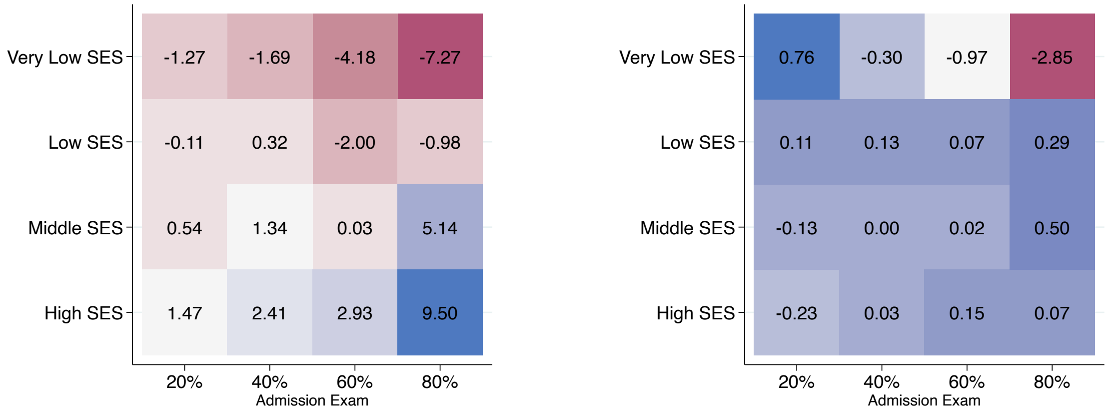
5.2 High-School Graduation

The simulation results presented in the previous sub-section establish that low-SES (high-SES) and high-achieving students are more likely to be assigned to elite (academic) schools under the large-scale implementation of the information intervention. A crowd-in effect for low-SES applicants arises in the new matching equilibrium due to the associated change in sorting across high-school tracks by relatively better-off applicants, who decrease their demand for elite programs and thereby switch to other academic (non-elite) high school programs. In this section, we use the predictions of the value added model (3) in order to assess the impact of the equilibrium allocation associated to the information intervention, when compared to the status-quo allocation, on education outcomes.

Panel A in Figure 9 documents that, under the full-scale implementation of the information intervention, low-SES applicants would be likely worse off in terms of high-school graduation. This is particularly the case for those with a relatively high admission score, who would be 4-7% less likely to graduate. The estimates of the value added model shown in Table 7 indicate that attending elite schools substantially decreases the probability of graduation, and low-SES applicants are more likely to be admitted in elite programs under the information intervention (see Panel A in Figure 8). Conversely, high-SES applicants would likely benefit from the information intervention in terms of on-time completion rates in upper secondary education. The presence of a positive gradient with respect to the admission score for these applicants mirrors the changes in assignment patterns in high school induced by the provision of the performance feedback (see Panel B in Figure 8).

As a way to mimic the experimental setting, we simulate the assignment outcomes in a counterfactual scenario where we provide the performance feedback only to the very-low SES applicants, who comprise 18% of the total number of participating students in 2014. We then use the estimated parameters of the value added model in order to predict the education outcomes associated to the matching equilibrium for such a targeted variant of the information intervention. Panel B in Figure 9 displays the impact of the policy on graduation rates resulting from this alternative simulation. As expected, given the relatively small-scale nature of the intervention, the effect on graduation rates is concentrated among the low-SES beneficiaries of the performance feedback, with a negligible spillover effect to

Figure 9: The Effect of Providing Performance Feedback on High-School Graduation (Percentage Points)



(a) Full Scale Implementation

(b) Feedback Targeted to Very Low SES

NOTE: This figure shows the percentage changes, by discrete categories of socio-economic status (Y-axis) and the score in the admission exam (x-axis), between the Information Policy and the Status Quo scenarios in the shares of applicants who complete the high-school program of their assignment in the centralized system.

the non-targeted applicants. Consistently with the experimental evidence shown in Table 4 and Figure 5, there is a small and positive effect of the performance feedback on graduation rates that accrues to low-achieving students.

These findings demonstrate that the sorting and crowd-in effects induced by the large-scale implementation of the information intervention largely offset the small and positive impact uncovered in the randomized evaluation on on-time graduation, as discussed in Section 3.4. Comparing Panels A and B in Figure 9 suggests that some of the negative consequences of scaling up on graduation rates among high-achieving and low-SES applicants may be determined by changes in their preferences rather than through an equilibrium effect. Yet, a comparison of the magnitude of the effect across the two panels suggests that the latter may explain a substantial portion of the overall policy impact on education outcomes.

6 Conclusion

The aim of this paper is to characterize the effect of an information intervention in an education setting, while taking into account the equilibrium effects that often arise during a large-scale implementation. We study a randomized experiment designed and implemented within a centralized school assignment mechanism in Mexico City. The intervention consists of providing a sample of 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. Evidence drawn from the experimental data 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 three years post-assignment (i.e., by the end of the twelfth grade.)

While these findings point toward a positive assessment of the intervention under study, it is not clear whether the results are also informative in a different situation where the same intervention is scaled up to reach a larger and more diverse population of recipients. The key challenge to scaling-up in our setting is that providing information about test scores to all applicants would necessarily trigger aggregate sorting and congestion effects within the centralized mechanism. These indirect effects would necessarily alter placement outcomes and, through those, subsequent schooling trajectories. The "light-touch" and ex-ante scalable nature of the provision of the performance feedback allows us to abstract away from a host of implementation issues, which would otherwise threaten the feasibility of a model-based evaluation.

The experimental design enables us to flexibly incorporate the provision of information into a model of school choice by allowing for different valuations over the schooling alternatives across applicants in the treatment and the control groups. We validate the out-of-sample predictions based on the estimated preferences in the control group using the realized allocation for the universe of the assigned applicants. Based on this evidence, we postulate that the corresponding predictions of the model based on the estimated preferences of the treated applicants likely approximate a counterfactual scenario in which the broader population of applicants would be given performance feedback.

Comparing the matching equilibria between the status quo and the policy counterfactual reveals that providing students with information about their skills increases the share of students assigned to their most preferred option. We also document substantial heterogeneity in the school choice responses across the diverse spectrum of applicants, which unlocks an equilibrium effect within the centralized system that ultimately hinders the educational trajectories of the sub-population of applicants targeted by the small-scale randomized evaluation.

Beyond the specific context of our analysis, the various pieces of evidence presented here may possibly offer a novel perspective of the impact of information provision in education settings characterized by fixed school capacities. Access to more accurate information about skills is shown to enhances the ex-ante efficiency of the student-school allocation. However, the distributional consequences crucially depend on the extent of the congestion externalities across schools. While the underlying sources of this equilibrium effect may differ from one setting to another, it is crucial to take those into account for the design of scalable education policies.

References

Abdulkadiroglu, Atila, Parag A. Pathak, Jonathan Schellenberg, and Christopher R. Walters, "Do Parents Value School Effectiveness?," *American Economic Review*, May 2020, *110* (5), 1502–1539.

Agarwal, Nikhil and Paulo Somaini, "Revealed Preference Analysis of School Choice Models," *Annual Review of Economics*, 2020, *12* (1), 471–501.

Agostinelli, Francesco, Ciro Avitabile, and Matteo Bobba, "Enhancing Human Capital in Children: A Case Study on Scaling," NBER Working Papers 31407, National Bureau of Economic Research, Inc June 2023.

Ajayi, Kehinde F., "School Choice and Educational Mobility: Lessons from Secondary School Applications in Ghana," *Journal of Human Resources*, 2022.

Akyol, Pelin, Kala Krishna, and Jinwen Wang, "Taking PISA seriously: How Accurate are Low-Stakes Exams?," *Journal of Labor Research*, 2021, *42*, 184–243.

Al-Ubaydli, Omar, John A. List, and Dana Suskind, "2017 Klein Lecture: The Science of Using Science: Toward an Understanding of the Threats to Scalability," *International Economic Review*, 2020, *61* (4), 1387–1409.

Allcott, Hunt, "Site Selection Bias in Program Evaluation," *The Quarterly Journal of Economics*, 2015, *130* (3), 1117–1165.

Anderson, Michael L., "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*, 2008, *103* (484), 1481–1495.

Andrabi, Tahir, Jishnu Das, and Ijaz Khwaja, "Report Cards: The Impact of Providing School and Child Test Scores on Educational Markets," *American Economic Review*, 2017, *107* (6), 1535–1563.

Angrist, Joshua and Miikka Rokkanen, "Wanna Get Away? Regression Discontinuity Estimation of Exam School Effects Away From the Cutoff," *Journal of the American Statistical Association*, 2015, *110* (512), 1331–1344.

Arcidiacono, Peter, Esteban M. Aucejo, and V. Joseph Hotz, "University Differences in the Graduation of Minorities in STEM Fields: Evidence from California," *American Economic Review*, 2016, *106* (3), 525–562.

Artemov, Georgy, Yeon-Koo Che, and YingHua He, "Stable Matching with Mistaken Agents," *Journal of Political Economy Microeconomics*, 2023, (Forthcoming).

August, Gerald, Michael Bloomquist, Susanne Lee, George Realmuto, and Joel Hektner, "Can Evidence-Based Prevention Programs be Sustained in Community Practice Settings? The Early Risers? Advanced-Stage Effectiveness Trial," *Prevention science : the official journal of the Society for Prevention Research*, 07 2006, *7*, 151–65.

Avery, Christopher and Caroline Hoxby, "The Missing "One-Offs": The Hidden Supply of High Achieving, Low Income Students," 2012. NBER Working paper No. 18586.

Azevedo, Eduardo M. and Jacob D. Leshno, "A Supply and Demand Framework for Two-Sided Matching Markets," *Journal of Political Economy*, 2016, *124* (5), 1235–1268.

Azmat, Ghazala, Manuel Bagues, Antonio Cabrales, and Nagore Iriberri, "What You Donât KnowÂ CanâÂt Hurt You? A Natural Field Experiment on Relative Performance Feedback in Higher Education," *Management Science*, August 2019, *65* (8), 3714–3736.

Banerjee, Abhijit V., Rukmini Banerji, James Berry, Esther Duflo, Harini Kannan, Shobhini Mukerji, Marc Shotland, and Michael Walton, "From Proof of Concept to Scalable Policies: Challenges and Solutions, with an Application," *Journal of Economic Perspectives*, November 2017, *31* (4), 73–102.

Bergman, Peter, "Parent-Child Information Frictions and Human Capital Investment: Evidence from a Field Experiment," *Journal of Political Economy*, 2021, *129* (1), 286–322.

Beuermann, Diether W. and C. Kirabo Jackson, "The Short and Long-Run Effects of Attending The Schools that Parents Prefer," *Journal of Human Resources*, 2020.

Beuermann, Diether W, C Kirabo Jackson, Laia Navarro-Sola, and Francisco Pardo, "What is a Good School, and Can Parents Tell? Evidence on the Multidimensionality of School Output," *The Review of Economic Studies*, 06 2022, *90* (1), 65–101.

Bobba, Matteo and Veronica Frisancho, "Perceived Ability and School Choices," TSE Working Papers 16-660, Toulouse School of Economics (TSE) June 2016.

_ **and** _ , "Self-perceptions about academic achievement: Evidence from Mexico City," *Journal of Econometrics*, 2022, *231* (1), 58–73.

Burks, Stephen V., Jeffrey P. Carpenter, Lorenz Goette, and Aldo Rustichini, "Overconfidence and Social Signalling," *The Review of Economic Studies*, 2013, *80* (3), 949–983.

Calsamiglia, Caterina, Guillaume Haeringer, and Flip Klijn, "Constrained School Choice: An Experimental Study," *American Economic Review*, 2010, *100* (4), 1860–1874.

Carlana, Michela, Eliana La Ferrara, and Paolo Pinotti, "Goals and Gaps: Educational Careers of Immigrant Children," *Econometrica*, January 2022, *90* (1), 1–29.

Caron, EB, Kristin Bernard, and Allison Metz, "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, 2021.

Chen, Yan and Onur Kesten, "Chinese College Admissions and School Choice Reforms: A Theoretical Analysis," *Journal of Political Economy*, 2017, *125* (1), 99–139.

CONEVAL, "Medición de Pobreza 2008-2018, Estados Unidos Mexicanos," 2018.

Davis, Jonathan, Jonathan Guryan, Kelly Hallberg, and Jens Ludwig, "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, 2021.

Delavande, Adeline, Xavier Giné, and David McKenzie, "Eliciting probabilistic expectations with visual aids in developing countries: how sensitive are answers to variations in elicitation design?," *Journal of Applied Econometrics*, 2011, *26* (3), 479–497.

Dizon-Ross, Rebecca, "Parents Beliefs about Their Children Academic Ability: Implications for Educational Investments," *American Economic Review*, August 2019, *109* (8), 2728–65.

Duflo, Esther, "Field Experiments and the Practice of Policy," *American Economic Review*, July 2020, *110* (7), 1952–73.

Dustan, Andrew, "Can large, untargeted conditional cash transfers increase urban high school graduation rates? Evidence from Mexico City's Prepa Si," *Journal of Development Economics*, 2020, *143* (C).

_ , Alain de Janvry, and Elisabeth Sadoulet, "Flourish or Fail?: The Risky Reward of Elite High School Admission in Mexico City," *Journal of Human Resources*, 2017, *52* (3), 756–799.

Estrada, Ricardo and Jeremie Gignoux, "Benefits to elite schools and the expected returns to education: Evidence from Mexico City," *European Economic Review*, 2017, *95*, 168–194.

Ewers, Mara and Florian Zimmermann, "Image And Misreporting," *Journal of the European Economic Association*, 2015, *13* (2), 363–380.

Fack, Gabrielle, Julien Grenet, and Yinghua He, "Beyond Truth-Telling: Preference Estimation with Centralized School Choice and College Admissions," *American Economic Review*, April 2019, *109* (4), 1486–1529.

Goux, Dominique, Marc Gurgand, and Eric Maurin, "Adjusting Your Dreams? High School Plans and Dropout Behaviour," *Economic Journal*, June 2017, *127* (602), 1025–1046.

Haaland, Ingar K., Christopher Roth, and Johannes Wohlfart, "Designing Information Provision Experiments," *Journal of Economic Literature*, 2023, (forthcoming).

Haeringer, Guillaume and Flip Klijn, "Constrained School Choice," *Journal of Economic Theory*, 2009, *144* (5), 1921–1947.

Hassidim, Avinatan, DÃÂ©borah Marciano, Assaf Romm, and Ran I. Shorrer, "The Mechanism Is Truthful, Why Aren't You?," *American Economic Review*, May 2017, *107* (5), 220–224.

Hastings, Justine S. and Jeffrey M. Weinstein, "Information, School Choice, and Academic Achievement: Evidence from Two Experiments," *The Quarterly Journal of Economics*, 2008, *123* (4), 1373–1414.

Heckman, James J., "Randomization and Social Policy Evaluation," in "Evaluating Welfare and Training Programs. Edited by C. F. Manski and I. Garfinkel," Harvard University Press, 1992.

_ , **John Eric Humphries, and Gregory Veramendi**, "Returns to Education: The Causal Effects of Education on Earnings, Health, and Smoking," *Journal of Political Economy*, 2018, *126* (S1), 197–246.

Jensen, Robert, "The (Perceived) Returns to Education and The Demand for Schooling," *Quarterly Journal of Economics*, 2010, *125* (2), 515–548.

Lavecchia, A.M., H. Liu, and P. Oreopoulos, "Chapter 1 - Behavioral Economics of Education: Progress and Possibilities," in Eric A. Hanushek, Stephen Machin, and Ludger Woessmann, eds., *Eric A. Hanushek, Stephen Machin, and Ludger Woessmann, eds.*, Vol. 5 of *Handbook of the Economics of Education*, Elsevier, 2016, pp. 1–74.

Lee, David S., "Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects," *The Review of Economic Studies*, 07 2009, *76* (3), 1071–1102.

List, John A., *The Voltage Effect: How to Make Good Ideas Great and Great Ideas Scale*, Penguin Books, 2022.

_ , **Azeem M. Shaikh, and Yang Xu**, "Multiple Hypothesis Testing in Experimental Economics," *Experimental Economics*, December 2019, *22* (4), 773–793.

Lucas, Adrienne M. and Isaac M. Mbiti, "Effects of School Quality on Student Achievement: Discontinuity Evidence from Kenya," *American Economic Journal: Applied Economics*, July 2014, *6* (3), 234?63.

Muralidharan, Karthik and Paul Niehaus, "Experimentation at Scale," *Journal of Economic Perspectives*, 2017, *31* (4), 103–124.

Neilson, Christopher, Claudia Allende, and Francisco Gallego, "Approximating the Equilibrium Effects of Informed School Choice," Working Papers 628, Princeton University, Department of Economics, Industrial Relations Section. July 2019.

Psacharopoulos, George and Harry Anthony Patrinos, "Returns to investment in education: a decennial review of the global literature," *Education Economics*, 2018, *26* (5), 445–458.

Romano, Joseph P. and Michael Wolf, "Exact and Approximate Stepdown Methods for Multiple Hypothesis Testing," *Journal of the American Statistical Association*, March 2005, *100*, 94–108.

_ **and** _ , "Stepwise Multiple Testing as Formalized Data Snooping," *Econometrica*, July 2005, *73* (4), 1237–1282.

_ **and** _ , "Efficient Computation of Adjusted p-Values for Resampling-Based Stepdown Multiple Testing," *Statistics & Probability Letters*, 2016, *113* (C), 38–40.

Roth, Alvin E. and Marilda Sotomayor, "Two-sided matching," in R.J. Aumann and S. Hart, eds., *Handbook of Game Theory with Economic Applications*, Vol. 1 of *Handbook of Game Theory with Economic Applications*, Elsevier, 1992, chapter 16, pp. 485–541.

Shapiro, S. S. and M. B. Wilk, "An analysis of variance test for normality (complete samples)," *Biometrika*, 1965, *52* (3-4), 591–611.

Walters, Christopher R, "The demand for effective charter schools," *Journal of Political Economy*, 2018, *126* (6), 2179–2223.

Zhao, Anqi and Peng Ding, "To Adjust or not to Adjust? Estimating the Average Treatment Effect in Randomized Experiments with Missing Covariates," *Journal of the American Statistical Association*, January 2024, *119* (545), 450–460.

Zimmerman, Seth D., "Elite Colleges and Upward Mobility to Top Jobs and Top Incomes," *American Economic Review*, January 2019, *109* (1), 1–47.

Appendices

A Experimental Instructions

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 (authors' translation from Spanish):

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?

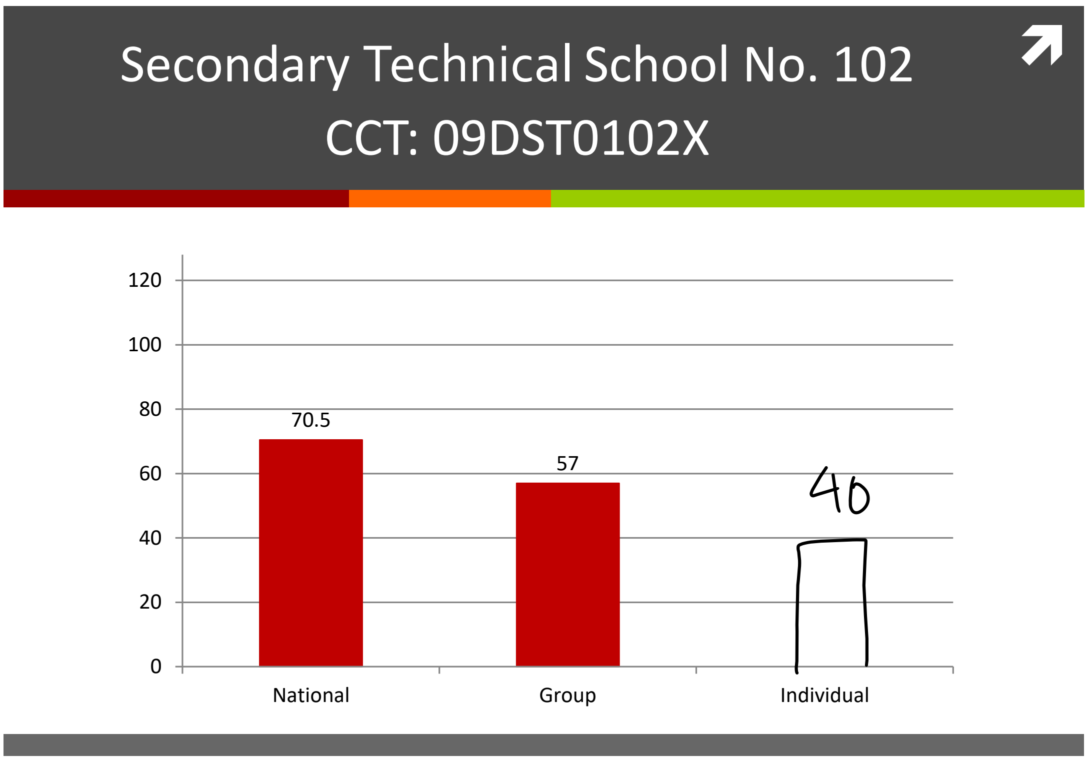
If respondents grasp the intuition behind our approach, they should provide an answer for question 2 that is larger than or equal to the answer in question 1, since the latter event is nested in the former. Similarly, respondents should report fewer beans in question 3 (close to zero probability event) 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 more or less discrete categories and/or more or fewer beans in order to assess 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 and Wilk, 1965) and reject it for only 11.4% of the respondents.

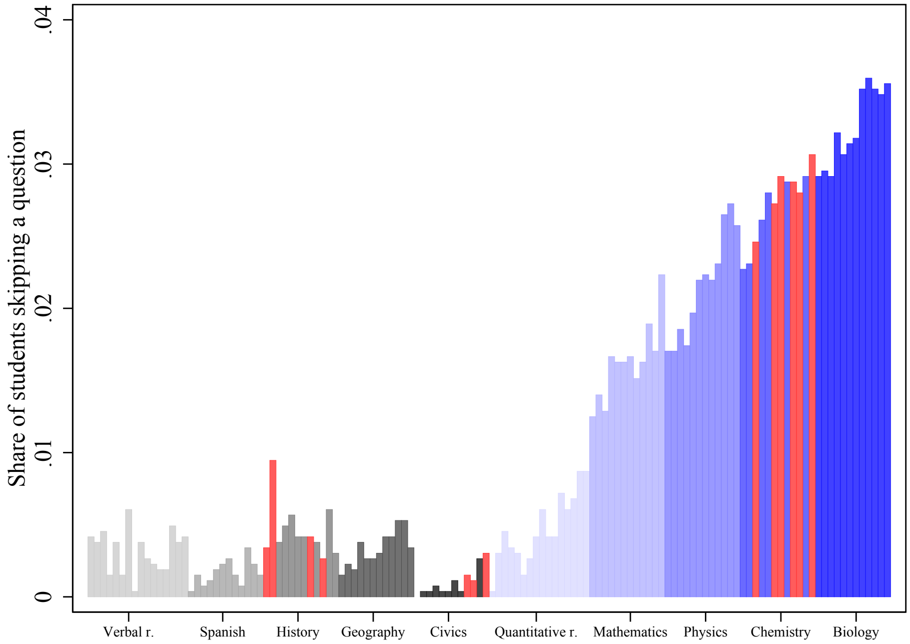
The delivery of individual scores takes place at the beginning of the follow up survey. Surveyors show the student a personalized graph with two pre-printed bars: the average score among the universe of applicants during the 2013 round and the average mock exam score of his classmates. During the delivery, the surveyors plotted a third bar corresponding to the individual's score in the mock test. Figure A.1 depicts a sample of the sheets used to deliver information to the students in the experiment.

Figure A.1: Sample of the Performance Delivery Sheet



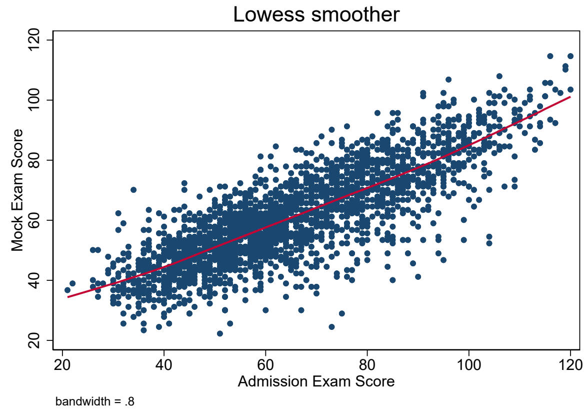
B Additional Figures and Tables

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

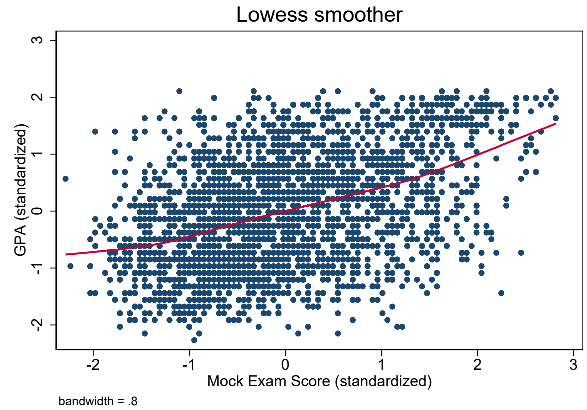


Note: The x-axis orders the 128 questions of the exam in order of appearance. Different colors are used to group together questions from the same section in the exam. Questions in red are the ones excluded from grading since the school curriculum did not cover those subjects by the time of the application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

Figure B.2: Correlates of the Mock Exam Score



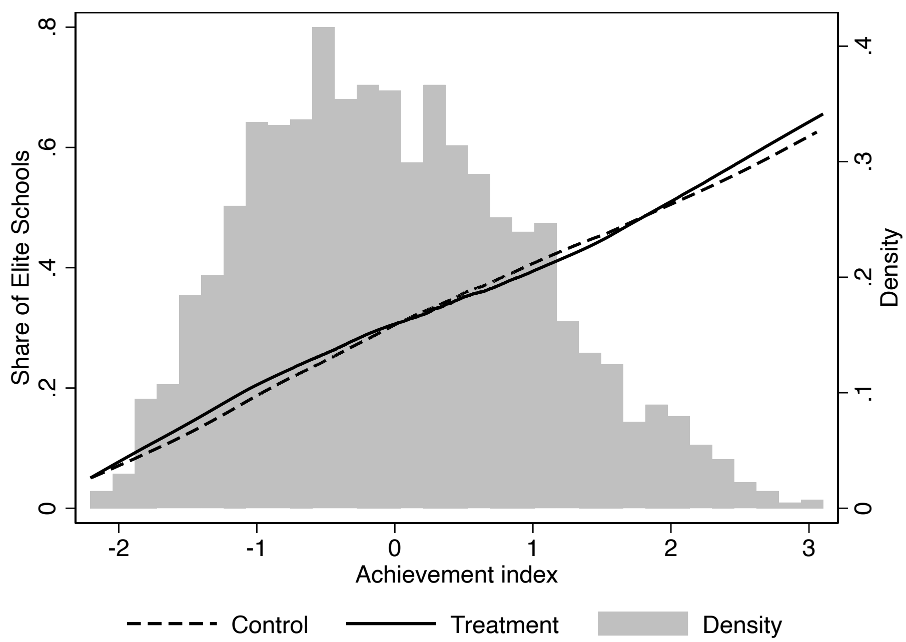
(a) Admission Exam Score



(b) Middle-school GPA

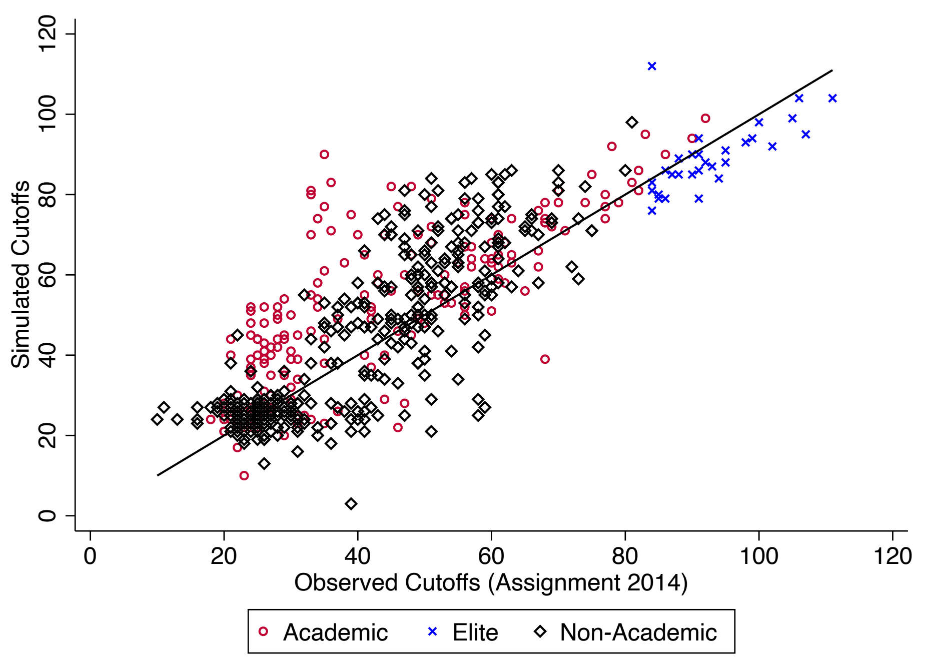
Note: This figure depicts scatter plots of the bi-variate relationship between the mock exam score and the admission exam score (Panel A), as well as between the mock exam score and the (standardized) Grade Point Average in middle school (Panel B). Overlaid on the scatters, we show non-parametric locally weighted estimates of the same relationships.

Figure B.3: Performance Feedback and Share of Elite Schools in Ranking



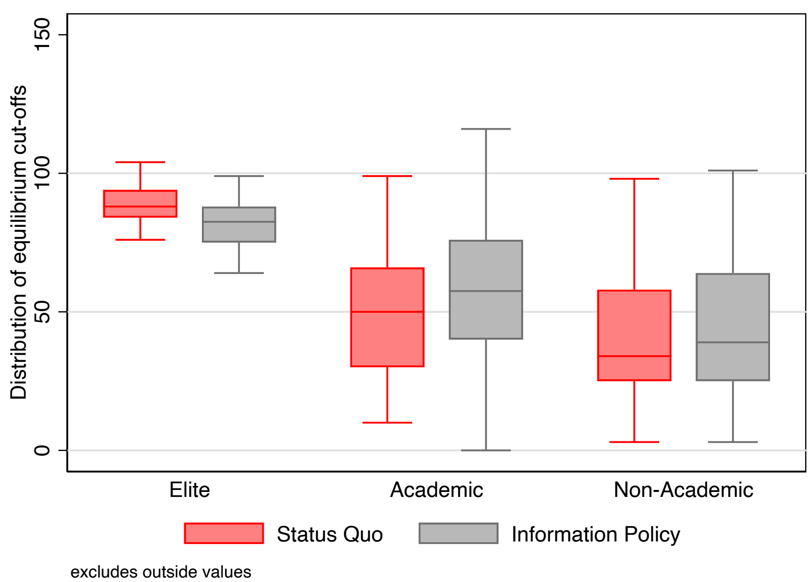
Note: This figure depicts the density of the performance index (y-axis on the right), which is a GLS-weighted average of the GPA in middle school, mock exam score, and exam score (Anderson, 2008). Overlaid on the density, we show non-parametric locally weighted estimates of the relationship between the share of elite schools in the applicant's ranking and the performance index separately for applicants in the treatment and control groups (y-axis on the left).

Figure B.4: Model Fit on Schools' Cutoff Scores for All Schools



NOTE: The observed cutoffs are computed from the data of the assignment mechanism in the year 2014 (see Section 2). The simulated cutoff scores displayed in the scatter plot are computed by running the Serial Dictatorship algorithm that is in place for the COMIPEMS system using the extrapolated school valuations from the experimental control group, the individual scores in the admission exam, and the school capacities as inputs.

Figure B.5: The Effect of Providing Performance Feedback on Cutoff Scores



NOTE: The simulated cutoff scores are computed by running the Serial Dictatorship algorithm that is in place for the COMIPEMS system using the predicted school valuations based on the control group (red bars) and the treatment group (grey bars). The corresponding estimates of the school choice model (2) are reported in Table B.9 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).

Table B.1: Performance in the Mock or Admission Exam and On-time Graduation

	Control Group	Control Group	All Applicants
Mock exam score (standardized)	0.072		
	[0.001]		
Admission exam score (standardized)		0.055	0.061
		[0.009]	[0.001]
Mean Dependent Variable	0.447	0.447	0.407
Number of Observations	1130	1207	195824
R-squared	0.019	0.011	0.015

NOTE: This Table shows OLS estimates of the relationship between the individual scores in the mock test or the admission exam and an indicator variable of whether students have completed upper secondary education in the statutory three years since enrollment in 10th grade. p-values reported in brackets refer to the conventional asymptotic standard errors, which take into account the clustering of the error terms at the high school level.

Table B.2: Treatment Effects on Application Outcomes

	Participates COMIPEMS	Exam Score	Length of ROL	Max cutoff in ROL	Min cutoff in ROL
Treatment	0.000	-0.669	0.126	1.641	-0.366
	[0.987]	[0.348]	[0.564]	[0.247]	[0.637]
	{0.983}	{0.908}	{0.982}	{0.777}	{0.982}
Achievement index	0.023	16.147	0.079	4.019	4.368
	[0.000]	[0.000]	[0.475]	[0.000]	[0.000]
	{0.001}	{0.001}	{0.978}	{0.001}	{0.001}
Treatment × Achievement index	-0.002	0.223	-0.108	0.262	0.483
	[0.777]	[0.582]	[0.489]	[0.757]	[0.510]
	{0.982}	{0.982}	{0.978}	{0.982}	{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. The achievement index is a GLS-weighted average (Anderson, 2008) 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 described in Romano and Wolf (2005a,b, 2016). Both inference procedures take into account clustering of the error terms at the middle school level and the block randomization design.

Table B.3: Summary Statistics and Randomization Check

	Control Group	Treatment Group	Treatment-Control
Mock exam score	60.540	62.366	1.496
	(15.416)	(16.290)	[0.163]
Exam score	65.541	65.248	-0.169
	(19.516)	(19.284)	[0.893]
GPA (middle school)	8.116	8.122	-0.013
	(0.846)	(0.846)	([0.777]
Scholarship in MS	0.106	0.115	0.007
	(0.308)	(0.319)	[0.642]
Grade retention in MS	0.263	0.233	-0.026
	(0.440)	(0.423)	[0.294]
Does not skip classes	0.971	0.971	-0.001
	(0.169)	(0.169)	[0.944]
Plans to go to college	0.670	0.671	-0.003
	(0.470)	(0.470)	[0.903]
Male	0.444	0.461	0.016
	(0.497)	(0.499)	[0.427]
Disabled student	0.142	0.148	0.006
	(0.349)	(0.355)	[0.719]
Indigenous student	0.085	0.101	0.017
	(0.278)	(0.302)	[0.219]
Does not give up	0.878	0.889	0.015
	(0.327)	(0.315)	[0.279]
Tries his best	0.735	0.722	-0.016
	(0.442)	(0.448)	[0.462]
Finishes what he starts	0.720	0.712	-0.015
	(0.449)	(0.453)	[0.442]
Works hard	0.725	0.739	0.010
	(0.447)	(0.439)	[0.644]
Experienced bullying	0.142	0.152	0.010
	(0.349)	(0.359)	[0.429]
Parental background and supervision	0.032	0.058	0.011
	(0.786)	(0.760)	[0.751]
High SES (asset index)	0.463	0.480	0.015
	(0.499)	(0.500)	[0.573]
Took prep courses	0.488	0.467	-0.026
	(0.500)	(0.499)	[0.314]
Exam Preparation	0.421	0.443	0.027
	(0.494)	(0.497)	[0.405]
Previous mock exam	0.269	0.290	0.017
	(0.444)	(0.454)	[0.649]
Previous mock exam with feedback	0.133	0.166	0.028
	(0.340)	(0.372)	[0.408]
Observations	1,290	1,203	2,493

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 p-values (in brackets) for the null hypothesis of zero effect.

Table B.4: On-time and Delayed Graduation Rates (Percentage Points)

	On-time Graduation	1-year delayed	2-year delayed	3-year delayed
Elite	47.0	54.6	58.4	60.8
Academic	37.6	44.2	47.4	49.6
Non-Academic	38.6	44.9	48.5	50.8
All	39.2	45.7	49.2	51.5

NOTE: The columns show graduation rates for each school track from 3 to 6 years after admission. The statutory high-school duration in Mexico is three years.

Table B.5: School Choices: Share in School Rankings

	Non-Academic	Academic	Elite
Treatment	0.001	-0.001	-0.000
	[0.936]	[0.928]	[0.999]
	{0.997}	{0.997}	{0.998}
Achievement index	-0.031	-0.054	0.084
	[0.002]	[0.000]	[0.000]
	{0.003}	{0.001}	{0.001}
Treatment $\times$ Achievement index	-0.032	0.030	0.002
	[0.011]	[0.008]	[0.894]
	{0.015}	{0.013}	{0.997}
Mean Control	0.365	0.336	0.299
Number of Observations	2493	2493	2493
Number of Clusters	90	90	90
R-squared	0.154	0.129	0.266

NOTE: The dependent variable is 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 achievement index is a GLS-weighted average (Anderson, 2008) 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 and Wolf (2005a,b, 2016). Both inference procedures take into account the clustering of the error terms at the middle school level.

Table B.6: Treatment Effects on Admission Outcomes

	Placed in 1st Round	Placed Any	Ranking of placement school
Treatment	-0.004	-0.006	0.141
	[0.796]	[0.719]	[0.411]
	{0.963}	{0.961}	{0.804}
Achievement index	0.068	0.064	-0.690
	[0.000]	[0.000]	[0.000]
	{0.001}	{0.001}	{0.001}
Treatment × Achievement index	-0.007	-0.005	-0.000
	[0.647]	[0.751]	[1.000]
	{0.934}	{0.963}	{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 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. The achievement index is a GLS-weighted average (Anderson, 2008) 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 described in Romano and Wolf (2005a,b, 2016). Both inference procedures take into account clustering of the error terms at the middle school level and the block randomization design.

Table B.7: Lee Bounds for the Effect of the Performance Feedback on Graduation on Time

	All Sample		Below Median		Above Median	
	Lower	Upper	Lower	Upper	Lower	Upper
Lee Bounds	0.016	0.039	0.041	0.063	-0.016	0.009
	[0.504]	[0.088]	[0.137]	[0.02]	[1.547]	[0.806]
Number of Observations	2493		1171		1322	
% Observations Trimmed	0.022		0.025		0.022	

NOTE: This table reports Lee bounds (Lee, 2009) in order to account for potentially non-random sample selection in the indicator variable for whether or not students graduate from secondary education three years post-assignment. The column 'Below Median' considers the sub-sample of applicants with a value of the achievement index below the median in the sample. The column 'Above Median' considers the sub-sample of applicants with a value of the achievement index above the median in the sample. p-values reported in brackets refer to the conventional asymptotic standard errors.

Table B.8: Treatment Effects on High-School Graduation Adjusted for Skills and Preferences

	Preferences	Placement
Treatment	0.001	0.001
	[0.906]	[0.939]
	{0.970}	{0.970}
Achievement index	-0.013	-0.016
	[0.000]	[0.000]
	{0.001}	{0.001}
Treatment $\times$ Achievement index	0.005	0.009
	[0.101]	[0.115]
	{0.164}	{0.164}
Mean Control	0.417	0.428
Number of Observations	2484	2236
Number of Clusters	90	90
R-squared	0.413	0.217

NOTE: The dependent variable "Preferences" is the estimated average graduation rate for the school programs in the students' school rankings, as predicted by the value added model (3). Analogously, the dependent variable "Placement" is the estimated graduation rate of the assigned school. Data for UNAM-sponsored high school programs is not available, hence the discrepancy in the number of observations in both columns when compared to the Tables in the main text (N=2,493). The achievement index is a GLS-weighted average (Anderson, 2008) 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 and Wolf (2005a,b, 2016). Both inference procedures take into account the clustering of the error terms at the middle school level.

Table B.9: Estimates of the School Choice Model

	Control Sample	Treatment Sample
Cole1-Aca	2.306	-0.0909
	(0.000)	(0.882)
Cole2-NonAca	-0.447	-1.901
	(0.453)	(0.002)
Cole3-Aca	2.143	-2.597
	(0.292)	(0.340)
Cole4-NonAca	1.499	-13.12
	(0.189)	(0.997)
Cole5-NonAca	2.160	0.999
	(0.002)	(0.182)
Cole6-NonAca	1.572	-0.345
	(0.001)	(0.540)
Cole7-Elite	3.124	0.101
	(0.000)	(0.923)
Cole8-Elite	10.47	-0.209
	(0.047)	(0.943)
Cole9-nonAca	0.274	-0.415
	(0.580)	(0.520)
Cole10-NonAca	-0.183	-1.621
	(0.768)	(0.029)
Cole11-Aca	1.246	-0.694
	(0.026)	(0.333)
Cole12-NonAca	-0.237	-0.657
	(0.629)	(0.261)
Cole13-Aca	0.741	0.397
	(0.075)	(0.438)
Cole14-Aca	-13.09	-2.049
	(0.998)	(0.761)
Cole15-Elite	4.380	-0.222
	(0.000)	(0.843)
Cole16-Elite	3.511	-2.062
	(0.006)	(0.178)
Cole1-Aca×Mock Score	-0.122	-0.401
	(0.565)	(0.063)
Cole2-NonAca×Mock Score	0.269	0.0496
	(0.249)	(0.817)
Cole3-Aca×Mock Score	0.338	-0.268
	(0.637)	(0.821)
Cole4-NonAca×Mock Score	0.154	0.511
	(0.714)	(1.000)

Continued on next page

Table B.9 Estimates of the School Choice Model – Continued from Previous Page

	Control Sample	Treatment Sample
Cole5-NonAca×Mock Score	0.595	-0.459
	(0.067)	(0.090)
Cole6-NonAca×Mock Score	0.261	-0.0889
	(0.199)	(0.660)
Cole7-Elite×Mock Score	-0.232	-0.862
	(0.533)	(0.024)
Cole8-Elite×Mock Score	-1.732	-0.938
	(0.060)	(0.321)
Cole9-nonAca×Mock Score	0.267	-0.745
	(0.204)	(0.002)
Cole10-NonAca×Mock Score	0.271	-0.628
	(0.267)	(0.006)
Cole11-Aca×Mock Score	0.167	-0.309
	(0.471)	(0.183)
Cole12-NonAca×Mock Score	0.194	-0.352
	(0.362)	(0.080)
Cole13-Aca×Mock Score	0.122	-0.487
	(0.502)	(0.006)
Cole14-Aca×Mock Score	0.788	0.716
	(1.000)	(0.755)
Cole15-Elite×Mock Score	-0.0397	-0.609
	(0.925)	(0.126)
Cole16-Elite×Mock Score	-0.181	-0.558
	(0.716)	(0.317)
Cole1-Aca×GPA	-0.453	-0.366
	(0.011)	(0.053)
Cole2-NonAca×GPA	-0.396	-0.515
	(0.049)	(0.010)
Cole3-Aca×GPA	0.0169	-0.248
	(0.976)	(0.765)
Cole4-NonAca×GPA	-0.508	0.0465
	(0.166)	(1.000)
Cole5-NonAca×GPA	-0.663	-0.187
	(0.018)	(0.450)
Cole6-NonAca×GPA	-0.516	-0.760
	(0.003)	(0.000)
Cole7-Elite×GPA	-0.337	-0.549
	(0.238)	(0.079)
Cole8-Elite×GPA	-0.445	-0.843
	(0.527)	(0.246)
Cole9-nonAca×GPA	-0.143	-0.195

Continued on next page

Table B.9 Estimates of the School Choice Model – Continued from Previous Page

	Control Sample	Treatment Sample
	(0.437)	(0.338)
Cole10-NonAca×GPA	-0.745	-0.200
	(0.001)	(0.351)
Cole11-Aca×GPA	-0.312	-0.103
	(0.102)	(0.619)
Cole12-NonAca×GPA	-0.528	-0.367
	(0.002)	(0.043)
Cole13-Aca×GPA	-0.446	-0.175
	(0.004)	(0.282)
Cole14-Aca×GPA	-0.329	0.311
	(1.000)	(0.840)
Cole15-Elite×GPA	-0.261	0.0693
	(0.408)	(0.836)
Cole16-Elite×GPA	-0.187	0.491
	(0.630)	(0.318)
Cole1-Aca×SES Index	-0.0335	-1.163
	(0.913)	(0.003)
Cole2-NonAca×SES Index	-0.431	-1.181
	(0.224)	(0.003)
Cole3-Aca×SES Index	0.350	-2.105
	(0.802)	(0.270)
Cole4-NonAca×SES Index	-0.665	-0.122
	(0.206)	(1.000)
Cole5-NonAca×SES Index	0.506	-0.103
	(0.202)	(0.815)
Cole6-NonAca×SES Index	-0.184	-0.976
	(0.517)	(0.006)
Cole7-Elite×SES Index	-0.522	-2.060
	(0.297)	(0.000)
Cole8-Elite×SES Index	4.464	-1.837
	(0.314)	(0.225)
Cole9-nonAca×SES Index	0.406	0.307
	(0.181)	(0.448)
Cole10-NonAca×SES Index	-0.119	-0.824
	(0.732)	(0.054)
Cole11-Aca×SES Index	0.0254	-0.936
	(0.937)	(0.029)
Cole12-NonAca×SES Index	-0.238	-0.268
	(0.404)	(0.459)
Cole13-Aca×SES Index	0.215	0.271
	(0.390)	(0.388)

Continued on next page

Table B.9 Estimates of the School Choice Model – Continued from Previous Page

	Control Sample	Treatment Sample
Cole14-Aca×SES Index	-0.249	-0.500
	(1.000)	(0.872)
Cole15-Elite×SES Index	-0.0760	-2.415
	(0.900)	(0.000)
Cole16-Elite×SES Index	-0.226	-2.522
	(0.747)	(0.002)
Cole1-Aca×Middle-School Math Score	0.155	-0.0929
	(0.495)	(0.724)
Cole2-NonAca×Middle-School Math Score	0.415	-0.330
	(0.088)	(0.198)
Cole3-Aca×Middle-School Math Score	0.403	-0.375
	(0.587)	(0.769)
Cole4-NonAca×Middle-School Math Score	0.176	0.357
	(0.750)	(1.000)
Cole5-NonAca×Middle-School Math Score	-0.418	0.544
	(0.308)	(0.081)
Cole6-NonAca×Middle-School Math Score	0.113	0.170
	(0.604)	(0.472)
Cole7-Elite×Middle-School Math Score	-0.168	-0.838
	(0.650)	(0.045)
Cole8-Elite×Middle-School Math Score	0.492	1.222
	(0.709)	(0.254)
Cole9-nonAca×Middle-School Math Score	0.0751	0.460
	(0.750)	(0.106)
Cole10-NonAca×Middle-School Math Score	0.728	0.296
	(0.017)	(0.290)
Cole11-Aca×Middle-School Math Score	0.234	-0.255
	(0.348)	(0.412)
Cole12-NonAca×Middle-School Math Score	0.122	-0.188
	(0.616)	(0.489)
Cole13-Aca×Middle-School Math Score	0.0217	0.589
	(0.910)	(0.008)
Cole14-Aca×Middle-School Math Score	0.408	1.953
	(1.000)	(0.332)
Cole15-Elite×Middle-School Math Score	0.810	-0.390
	(0.049)	(0.343)
Cole16-Elite×Middle-School Math Score	0.415	-0.157
	(0.390)	(0.780)
Cole1-Aca×Parent Higher Education	-0.508	0.209
	(0.247)	(0.682)
Cole2-NonAca×Parent Higher Education	-0.409	-0.0186

Continued on next page

Table B.9 Estimates of the School Choice Model – Continued from Previous Page

	Control Sample	Treatment Sample
	(0.408)	(0.972)
Cole3-Aca×Parent Higher Education	-0.948	-13.59
	(0.451)	(0.994)
Cole4-NonAca×Parent Higher Education	-0.636	0.606
	(0.573)	(1.000)
Cole5-NonAca×Parent Higher Education	-1.391	0.329
	(0.121)	(0.595)
Cole6-NonAca×Parent Higher Education	-0.829	0.0533
	(0.066)	(0.910)
Cole7-Elite×Parent Higher Education	-1.399	0.741
	(0.032)	(0.296)
Cole8-Elite×Parent Higher Education	-15.81	-13.96
	(0.991)	(0.988)
Cole9-nonAca×Parent Higher Education	-0.915	-0.658
	(0.088)	(0.302)
Cole10-NonAca×Parent Higher Education	-1.268	-0.191
	(0.039)	(0.747)
Cole11-Aca×Parent Higher Education	-0.453	-0.283
	(0.421)	(0.661)
Cole12-NonAca×Parent Higher Education	-0.910	0.0912
	(0.118)	(0.852)
Cole13-Aca×Parent Higher Education	-0.506	0.177
	(0.223)	(0.677)
Cole14-Aca×Parent Higher Education	-0.627	-13.84
	(1.000)	(0.994)
Cole15-Elite×Parent Higher Education	-2.705	1.004
	(0.000)	(0.165)
Cole16-Elite×Parent Higher Education	-1.668	1.256
	(0.044)	(0.171)
Distance (Km)	-0.271	-0.202
	(0.000)	(0.000)
Distance (Km)×Mock Score	0.0144	0.0169
	(0.046)	(0.023)
Distance (Km)×GPA	0.00383	0.00763
	(0.585)	(0.312)
Distance (Km)×SES Index	0.000575	0.0457
	(0.961)	(0.003)
Distance (Km)×Middle-School Math Score	0.0185	0.0215
	(0.035)	(0.033)
Distance (Km)×Parent Higher Education	0.0239	0.00427
	(0.142)	(0.808)

Continued on next page

Table B.9 Estimates of the School Choice Model – Continued from Previous Page

	Control Sample	Treatment Sample
Cutoff Score	1.063	1.510
	(0.000)	(0.000)
Cutoff Score×Mock Score	0.103	0.170
	(0.151)	(0.019)
Cutoff Score×GPA	0.126	0.154
	(0.052)	(0.023)
Cutoff Score×SES Index	0.0763	0.227
	(0.470)	(0.073)
Cutoff Score×Middle-School Math Score	0.110	0.262
	(0.212)	(0.002)
Cutoff Score×Parent Higher Education	0.519	-0.0602
	(0.003)	(0.727)
N	637,901	590,526

NOTE: This table displays the full set of maximum-likelihood estimates and standard errors (in parenthesis) for the parameters of the school choice model (2).

Table B.10: Estimates of the School Graduation Model

	On-time graduation
Cole2-NonAca	-0.174
	(0.023)
Cole3-Aca	0.056
	(0.052)
Cole4-NonAca	-0.028
	(0.201)
Cole5-NonAca	0.095
	(0.030)
Cole6-NonAca	0.055
	(0.017)
Cole7-Elite	-0.109
	(0.023)
Cole8-Elite	-0.050
	(0.055)
Cole9-nonAca	0.085
	(0.037)
Cole10-NonAca	0.249
	(0.035)
Cole11-Aca	-0.030
	(0.032)
Cole12-NonAca	0.060
	(0.028)

Continued on next page

Table B.10 Estimates of the School Graduation Model – Continued from Previous Page

	On-time graduation
Cole13-Aca	0.146
	(0.023)
Cole14-Aca	0.130
	(0.164)
GPA	0.160
	(0.005)
SES index	0.017
	(0.006)
Parent Education	0.004
	(0.004)
Middle-School Math Score	0.015
	(0.004)
Cole2-NonAcaXGPA	-0.105
	(0.010)
Cole3-AcaXGPA	0.031
	(0.022)
Cole4-NonAcaXGPA	0.054
	(0.048)
Cole5-NonAcaXGPA	0.045
	(0.015)
Cole6-NonAcaXGPA	-0.002
	(0.008)
Cole7-EliteXGPA	0.054
	(0.008)
Cole8-EliteXGPA	0.098
	(0.027)
Cole9-nonAcaXGPA	-0.015
	(0.017)
Cole10-NonAcaXGPA	0.024
	(0.015)
Cole11-AcaXGPA	-0.006
	(0.015)
Cole12-NonAcaXGPA	0.002
	(0.011)
Cole13-AcaXGPA	-0.008
	(0.008)
Cole14-AcaXGPA	0.044
	(0.081)
Cole2-NonAcaXSES index	-0.006
	(0.012)
Cole3-AcaXSES index	0.021
	(0.040)

Continued on next page

Table B.10 Estimates of the School Graduation Model – Continued from Previous Page

	On-time graduation
Cole4-NonAcaXSES index	-0.045
	(0.049)
Cole5-NonAcaXSES index	0.037
	(0.019)
Cole6-NonAcaXSES index	-0.022
	(0.009)
Cole7-EliteXSES index	0.003
	(0.010)
Cole8-EliteXSES index	0.014
	(0.034)
Cole9-nonAcaXSES index	0.005
	(0.019)
Cole10-NonAcaXSES index	-0.015
	(0.018)
Cole11-AcaXSES index	-0.002
	(0.016)
Cole12-NonAcaXSES index	-0.019
	(0.011)
Cole13-AcaXSES index	-0.007
	(0.009)
Cole14-AcaXSES index	0.111
	(0.097)
Cole2-NonAcaXParent Education	0.009
	(0.012)
Cole3-AcaXParent Education	0.028
	(0.018)
Cole4-NonAcaXParent Education	0.075
	(0.059)
Cole5-NonAcaXParent Education	0.009
	(0.017)
Cole6-NonAcaXParent Education	0.010
	(0.008)
Cole7-EliteXParent Education	0.009
	(0.006)
Cole8-EliteXParent Education	0.034
	(0.018)
Cole9-nonAcaXParent Education	0.016
	(0.020)
Cole10-NonAcaXParent Education	0.000
	(0.019)
Cole11-AcaXParent Education	0.019
	(0.020)

Continued on next page

Table B.10 Estimates of the School Graduation Model – Continued from Previous Page

	On-time graduation
Cole12-NonAcaXParent Education	-0.008
	(0.016)
Cole13-AcaXParent Education	0.002
	(0.007)
Cole14-AcaXParent Education	0.008
	(0.043)
Cole2-NonAcaXMiddle-School Math Score	-0.021
	(0.008)
Cole3-AcaXMiddle-School Math Score	0.024
	(0.019)
Cole4-NonAcaXMiddle-School Math Score	-0.004
	(0.058)
Cole5-NonAcaXMiddle-School Math Score	0.052
	(0.020)
Cole6-NonAcaXMiddle-School Math Score	0.017
	(0.007)
Cole7-EliteXMiddle-School Math Score	0.004
	(0.007)
Cole8-EliteXMiddle-School Math Score	0.006
	(0.024)
Cole9-nonAcaXMiddle-School Math Score	0.027
	(0.020)
Cole10-NonAcaXMiddle-School Math Score	0.047
	(0.020)
Cole11-AcaXMiddle-School Math Score	0.064
	(0.020)
Cole12-NonAcaXMiddle-School Math Score	0.055
	(0.013)
Cole13-AcaXMiddle-School Math Score	0.016
	(0.008)
Cole14-AcaXMiddle-School Math Score	-0.194
	(0.115)
School-Average GPA	0.024
	(0.010)
School-Average SES index	-0.021
	(0.011)
School-Average Parent Education	-0.011
	(0.012)
School-Average Middle-School Math Score	0.000
	(0.012)
Constant	0.378

Continued on next page

Table B.10 Estimates of the School Graduation Model – Continued from Previous Page

	On-time graduation
	(0.012)
N	182,824

Note: This table displays the full set of OLS estimates and standard errors (in parenthesis) of the parameters of the school effectiveness model (3). The ROL fixed effects are included in the regression but they are not reported. The sample includes all the assigned applicants to the centralized system in the year 2010 except for the 15% of applicants who are assigned to the UNAM-sponsored high-schools (2 colleges out of 16 participating colleges).