Preferences, access, and the STEM gender gap in centralized high school assignment

Diana Ngo *
Department of Economics
Occidental College

Andrew Dustan
Department of Economics
Vanderbilt University

July 2022

Abstract

The gender gap in science, technology, engineering, and mathematics (STEM) widens during high school, due both to differences in student choices and institutional barriers to accessing STEM education. Using rich data from Mexico City and a structural model of high school choice, we show that most of the gap is due to gendered choices but that test-based assignment also restricts access to selective STEM schools. Simulations show that affirmative action policies targeting STEM programs eliminate the gap with small welfare impacts, while untargeted policies widen the gap substantially. Policy impacts are heterogeneous by student placement test score and school competitiveness.

Keywords: School choice, STEM education, gender disparities

*E-mail addresses: dngo@oxy.edu, andrew.dustan@vanderbilt.edu. Author names are in reverse alphabetical order.
We thank Mexico's Secretariat of Public Education for providing the data. We also thank Kevin Grier, Robin Grier, and participants at the Liberal Arts College Development Conference, Association for Education Finance and Policy Conference, Association for Public Policy and Management Conference, Western Economics Association Conference, Southern Economic Association, Pacific Conference for Development Economics, Latin American and Caribbean Economic Association Meetings, Universidad del Rosario, University of Hawaii, and Texas Tech University for helpful feedback. Aleeza Toribio provided excellent research assistance.

1 Introduction

Despite substantial interest in understanding and reducing the gender gap in science, technology, engineering, and mathematics (STEM), this gap remains large throughout the world and continues to have important welfare consequences. Within the context of school choice, the gap arises due to a combination of gendered choices (Zafar 2013; Shi 2018) and admissions structures that constrain access to quality schools, particularly for females (Saygin 2020).[1] Differential access may be a particular concern in contexts with test-based admissions criteria.[2] This interaction between access and choice raises questions about how much each contributes to the gap and questions about the effectiveness of policies aimed at changing the gap. Due to gendered choices, policies that raise priority for females at all schools relative to the status quo (e.g., GPA-based priorities or lotteries) may exacerbate the STEM gap. At the same time, policies that reduce the gap by increasing female placement in STEM schools (e.g., STEM affirmative action) may come at a large welfare cost. Moreover, both sets of policies may have differential effects on students across the admissions priority distribution. The extent to which these concerns manifest in practice remains an important empirical question for designing effective policy.

We examine these issues by estimating a rich model of student choices over STEM and non-STEM options, including both highly-demanded elite STEM programs and non-elite technical STEM programs. We study Mexico City high schools in 2007 and 2008, where females were 11 percentage points (p.p.) less likely to attend STEM programs than males. Mexico City has a centralized high school choice policy in which students submit a ranked list of educational options (either schools or specific programs within some schools) and are given assignment priorities based on their performance on a placement test. We use a conditional logit model to estimate willingness to travel for different program types by gender, geographic location, and other student characteristics including math achievement, Spanish achievement, and parental education. Importantly, we identify

1. While "access" can be understood to encompass a wide range of both institutionalized and de facto contributors, we use the term "access" in this paper to more narrowly refer to admissions criteria that are formally considered in granting entry into a field or school.

2. Examples of explicit merit-based tracking into schools include Ghana (Ajayi 2013), Romania (Pop-Eleches and Urquiola 2013), Trinidad and Tobago (Jackson 2010), and the U.S. (Corcoran and Baker-Smith 2018).

student ability using performance on a low-stakes exam that is separate and imperfectly correlated with the placement test. This allows us to estimate preferences for the full distribution of students over all program types under relatively weak assumptions, even when low-scoring students are never admitted to highly-demanded schools. We then use the model to simulate the effects of various policy changes on the STEM gender gap and welfare. These include changes to the priority structure used for assignment, such as using lottery-based priorities and general and STEM-specific affirmative action (AA).

Several key findings emerge from the model estimates. First, both males and females have high willingness to travel to elite STEM programs relative to traditional programs and lower willingness to travel to non-elite STEM programs. The majority of females include elite and/or non-elite STEM programs in their choice lists, and higher math ability predicts stronger preferences for both elite and non-elite STEM programs. Second, preferences for STEM programs are gendered, with females expressing a lower willingness to travel to STEM programs relative to males, after controlling for a variety of other factors. These gendered preferences contribute substantially to the STEM gender gap, with the overall gap disappearing under simulations where observably similar males and females have the same preferences. Third, access constraints are important for assignment to the elite STEM programs. Approximately 20% of the gender gap in elite STEM programs remains when males and females are simulated to have the same preferences, indicating that the test-based admissions policy is a barrier for females, who score an average 0.2 standard deviations lower on the placement test. Together, the results highlight a nuanced interaction between preferences and access across the achievement distribution and by school competitiveness, suggesting that the effects of any proposed policy changes may be complex and heterogeneous.

The simulations confirm this, showing that policies that address general access concerns without consideration for preference differences are ineffective for closing the STEM gap in this context. Specifically, due to gendered choices, policies that increase female admission priorities to all programs (e.g., using lotteries, prior grades, or general AA) actually increase the overall size of the gap. This is consistent with females differentially preferring non-STEM programs and being given higher chances of admission into those schools. However, the overall effects obscure more heterogeneous impacts—the overall increase in the STEM gender gap comes from a reduction in the elite STEM gap coupled with a larger countervailing increase in the non-elite STEM gap. With respect to welfare, these policies have relatively large welfare costs, particularly for both males and females with high placement test scores under a lottery system and for males with high placement scores under general AA or priority based on prior grades.

On the other hand, policies that only increase female admission priorities at STEM programs can reduce the gap substantially. Specifically, a STEM AA policy awarding females points equivalent to the male-female placement test score differential decreases the STEM gap by 36%, and a 50% female minority reserve quota at STEM programs decreases the STEM gap by 69%. Both forms of STEM affirmative action have minimal impacts on aggregate student welfare and comparatively small gender-specific welfare effects.

Again, though, these policies have heterogeneous effects across the placement score distribution and by program competitiveness. The largest gap reductions occur for elite STEM programs and among students at the margin of admission to elite programs. The non-elite STEM gap falls at the lower end of the score distribution but actually increases at the higher end of the distribution, as higher scoring females are moved from non-elite STEM programs into elite STEM programs. The welfare effects tend to be small across the score distribution, with slightly larger impacts at scores close to elite STEM cutoffs. Overall, STEM-specific AA policies are effective in reducing the gap with small welfare costs, and point-based and quota forms of STEM AA have similar impacts in our context.

This study relates to a large body of literature focused on factors contributing to the STEM gender gap. Consistent with other work, we find gendered preferences for STEM. These studies show that males and females have different preferences over coursework (Zafar 2013), with females preferring work that is perceived as more pro-social and cooperative (Shi 2018). Underlying these gendered preferences are environmental factors such as culture, role models, family expectations, and chilly climates (Kahn and Ginther 2017). With respect to issues of access, we find admissions constraints that are similar to other studies where test-based admissions limit entry to highly demanded schools for particular groups (Corcoran and Baker-Smith 2018). Test-based admissions criteria constraints are likely to contribute to the STEM gender gap elsewhere, with lower female test scores appearing throughout the world (Hyde et al. 2008; Stoet and Geary 2013). This is a particular concern when there is gender bias in standardized test scores (Saygin 2020).

Leveraging Mexico City's unique institutional setting, we contribute to this literature by examining two important aspects of the STEM gender gap that have received less attention. First, while most prior work has focused on the gap at the university level (Card and Payne 2021; Lichtenberger and George-Jackson 2013), we examine STEM high school programs and younger students' preferences. Here, we show that the STEM gender gap arises early, with underlying gaps in both choices and access emerging early in the pipeline. Second, we examine the interaction between access and choice in an environment where STEM options are heterogeneous, with large differences between academic and technical programs. Technical programs are less studied, but they are important labor market entry points for many economies (Rothwell 2013). We find that academic STEM programs are highly demanded by both males and females, and access barriers are particularly important for females in the middle to high region of the placement test distribution. Technical programs account for the entirety of the gap at the lower end of the distribution and a sizable portion of the gap among the higher end of the distribution. These programs are less demanded, with preference differences playing a larger role in generating the gap.

Our simulated preference structure changes directly address many proposed policies that are designed to increase access for particular groups. These include changing admissions criteria to include relative measures based on levels of disadvantage (Jaschik 2019), test scores aligned with problem-solving curriculum (Treschan et al. 2013), prior course-taking and grades (California Community Colleges 2018), lotteries (Shapiro 2020), and various combinations of the above (Karp 2021). Furthermore, there are ongoing debates around AA in admissions (Fryer Jr and Loury 2005; Bertrand, Hanna, and Mullainathan 2010; Francis-Tan and Tannuri-Pianto 2018; Frisancho and Krishna 2016) and related policies, such as reserving seats for top performers from each lower level school (Cullen, Long, and Reback 2013; Treschan et al. 2013). Our policy counterfactual analyses are similar to Corcoran and Baker-Smith (2018), but we pair the simulations with a rich model of student preferences. Our results suggest that policies like lotteries or general AA can widen the STEM gender gap (or other areas of inequality) when preference differences are important. Instead, targeted changes to the priority structure (e.g., STEM AA) can be effective in achieving specific policy goals and can be implemented with small welfare costs.

This paper adds to an active literature that structurally estimates student preferences over school characteristics within centralized assignment mechanisms.[3] The mechanism and empirical context present many of the conceptual and practical issues reviewed in Agarwal and Somaini (2020). Our approach follows Fack, Grenet, and He (2019) in assuming that students are assigned to their most-preferred feasible program conditional on their assignment priority. We also allow for rich preference heterogeneity with respect to observable student characteristics, following Abdulkadirolu et al. (2020).

The rest of the paper proceeds as follows. Section 2 provides more background on the definition of STEM and the returns to STEM fields of study. Section 3 describes the model and simulation strategy. Sections 4 and 5 present the data and results, respectively, and Section 6 concludes.

2 Context

2.1 STEM and STEM high schools

STEM categories

Before proceeding, we present our definition of STEM for this analysis. Various STEM classifications exist, differing in their approaches to categorizing certain groups of occupations such as educators, managers, technicians, healthcare professionals, and social scientists (Beede et al. 2011). The definition of STEM has implications for the size of the gender gap, as female underrepresenta-

3. See, for example, Agarwal and Somaini (2018), Beuermann et al. (2018), Bordón, Canals, and Mizala (2020), Calsamiglia, Fu, and Güell (2020), Glazerman and Dotter (2017), Hastings, Kane, and Staiger (2009), Kapor, Neilson, and Zimmerman (2020), and Pathak and Shi (2021).

tion is mostly found in the math-intensive or "high-status" fields, while more women are found in fields such as life science, psychology, and social science (George-Jackson 2011; Ceci et al. 2014).

In this analysis, we apply the classification developed by the Brookings Institute (Rothwell 2013). This classification uses data from the U.S. Department of Labor to categorize occupations based on the level of STEM knowledge they require, identifying high-STEM occupations as those which require substantial knowledge (at least 1.5 standard deviations above the mean knowledge score) in any one STEM field (science, technology, engineering, or math). Overall, this generates a broad classification of STEM. It includes non-math intensive areas such as nursing in its STEM classification and therefore represents a more inclusive definition of STEM. In addition, it includes a set of occupations, like equipment maintenance and building technicians, that do not require higher education. Including these as STEM fields allows us to compare nonprofessional STEM areas to professional STEM areas. These nonprofessional STEM occupations have received less attention despite the fact that they account for a large fraction of STEM jobs and pay a wage premium relative to nonprofessional non-STEM jobs (Rothwell 2013). Moreover, these fields are particularly relevant in Mexico and other developing countries. Here, higher education participation is growing but not universal, and large parts of the economy are driven by manufacturing.

STEM high school programs

We also identify STEM high school programs using a broad definition of STEM. We define STEM programs as programs that focus on providing specific science and math coursework to their students, including schools that offer more academic curricula and schools that offer more technical curricula. Both types of STEM high school programs exist in a variety of contexts. The more academically focused STEM programs, for example The Bronx High School of Science in New York City, are often highly selective and are pathways to elite education (Corcoran and Baker-Smith 2018). The technical/vocational STEM programs are more widely accessible and have garnered increasing interest in recent years as options for job preparation and alternative pathways to higher education (Gewertz 2018; Dougherty and Harbaugh Macdonald 2019).

Given the range of different features and options, prior research shows mixed results on the effectiveness of STEM high school programs. Some studies find positive impacts of STEM program attendance on STEM course-taking and interest (Means et al. 2016) and on STEM test scores and rates of STEM test-taking (Wiswall et al. 2014), while others find no positive impacts (Eisenhart et al. 2015; Bottia et al. 2018) or heterogeneous results across different STEM programs (Gnagey and Lavertu 2016).

As the returns to STEM high school programs appear to be context-dependent, we discuss the likelihood of positive returns in Mexico City. Dustan, de Janvry, and Sadoulet (2017) show that attending selective STEM programs in Mexico City high schools increases end-of-high-school test scores. There is little evidence available regarding the effectiveness of technical STEM programs.[4] However, using labor force data from Mexico, we compute and present some motivating statistics that suggest positive returns from technical STEM high school programs. Specifically, there is an 11% wage premium for STEM occupations among females with no higher education (22% among females with STEM occupations requiring higher education). In addition, females who study STEM in high school are 17% more likely to study STEM in college, and females who study STEM in college are 7.5 times more likely to work in STEM occupations. Females who study STEM in high school but do not continue to college are also twice as likely to work in STEM compared to females who do not continue to college and do not study STEM in high school. The details of this analysis are included in the Appendix, with the results in Appendix Tables A.1 and A.2. Together, these results suggest that, in Mexico City, there are positive returns to STEM high schools for females, motivating further study on ways to increase the proportion of females studying in STEM high schools.

4. See Ortega Hesles and Dougherty (2017) for a working paper related to the returns to technical schools in Mexico City.

2.2 School choice in Mexico City

Here, we describe more specifically the types of STEM programs in Mexico City high schools and the assignment mechanism that places students in programs. This description applies to the 2007 and 2008 rounds of assignment, the cohorts analyzed in this paper. Students select from over 600 academic programs available throughout Mexico City. We delineate five different types of programs: elite STEM, non-elite STEM, elite non-STEM, technical non-STEM, and traditional public school programs (which are non-technical, non-elite, and non-STEM).

Two subsystems of programs are considered elite. Both are highly demanded and draw students from all areas of Mexico City despite being clustered near the city center (Dustan and Ngo 2018). The twelve elite STEM programs focus on science and engineering and are affiliated with the Instituto Politécnico Nacional (IPN), a prestigious national polytechnic university. The IPN also offers four elite non-STEM programs, focused on social sciences and business administration. The large majority of elite non-STEM programs (14) are affiliated with the Universidad Nacional Autónoma de México (UNAM), offering broader, liberal arts-style curricula.[5] The remaining non-elite programs tend to be less competitive, drawing most of their students from nearby neighborhoods. The non-elite STEM programs exist within a set of technical and vocational schools, primarily in the Colegio Nacional de Educación Profesional Técnica (CONALEP) and Dirección General de Educación Tecnológica Industrial (DGETI). These technical schools offer training intended to prepare students for specific occupations. Students taking part in the COMIPEMS process choose specializations within these schools, with many schools offering both STEM-focused and non-STEM specializations. We categorize students as assigned to non-elite STEM programs if their specializations are STEM specializations (e.g., mechanical engineering) and as assigned to technical non-STEM schools if their specializations are non-STEM specializations (e.g., business administration), following a procedure we will make more concrete in the next section.[6] Thus

5. See Dustan, de Janvry, and Sadoulet (2017) for a more detailed description of the elite programs as well as their associated returns.

6. See Avitabile, Bobba, and Pariguana (2017) and Ortega Hesles and Dougherty (2017) for additional discussion on the constraints and returns to these educational tracks.

students assigned to technical schools may still be considered non-STEM because of their specialization. Finally, the non-elite non-STEM programs consist of a set of traditional schools offering general courses.

Students from any type of high school may go on to higher education, regardless of whether or not their secondary schooling was geared towards general education or vocational training. UNAM-affiliated high schools differ in one important way; students attending the UNAM-affiliated elite non-STEM schools are guaranteed entrance into the UNAM university conditional on meeting certain academic requirements. No such guarantee exists for other schools.

During the final year of middle school, students interested in attending these public high schools participate in a competitive choice-based assignment process. The choice process occurs during and after students' final year of middle school (grade nine). First, students rank their preferred programs, listing up to a maximum of twenty options. Students then take a comprehensive placement test, covering content on verbal reasoning, Spanish, history, geography, philosophy/civilization/ethics, quantitative reasoning, math, physics, chemistry, and biology.[7] Each of the 128 questions is worth one point, with no penalty for incorrect answers. Finally, students are assigned to programs using a serial dictatorship algorithm that is a special case of the student-proposing deferred acceptance algorithm (SPDA) characterized by Gale and Shapley (1962). A computer orders all students according to their placement test scores, and moving from highest scoring to lowest scoring, assigns each student to her highest-ranked program with a remaining seat when her turn arrives. Programs that fill up thus have a "cutoff score" equal to the placement test score of the student assigned to the final seat.[8]

There is full compliance with the assignment mechanism, in the sense that students cannot enroll at a public school to which they were not assigned, though students can opt out of the system entirely. Most participating students (85%) are assigned through this algorithm, while the remaining

7. Due to historical and political reasons, there is an alternate version of the test for students who rank an UNAM-affiliated program as their first choice. The two versions of the test are written in collaboration and are perceived to be substantively equivalent.

8. The elite IPN and UNAM programs require that students have a middle school GPA of 7/10 or above, but this standard is low enough that it almost never binds.

students are not assigned because they have chosen only schools with cutoff scores higher than their own score. These students can later choose from schools that have seats remaining after the computerized assignment. We do not observe this outcome for most years, but show in Appendix Table B.1 that accounting for assignments in this later round does not substantively affect the size of the STEM gender gap using data from an earlier cohort for which we have full assignment data.

During the same school year in which students participate in the COMIPEMS process, they also complete a low-stakes standardized exam called the ENLACE 9. The ENLACE 9 is taken after students have submitted their choice lists, but before they take the placement test, although they do not know their ENLACE 9 scores at the time of the placement test. The ENLACE 9 has no bearing on middle school graduation or any other student outcome. It consists of math and Spanish subsections, with other subjects rotating annually.

3 Methods

This section explains the discrete choice model used to estimate student preferences, the method used to estimate marginal willingness to travel, and the counterfactual simulations that show how the STEM gap responds to policy changes.

3.1 Student preferences and their estimation

We model student i as having preferences over all programs $j \in \{1, ..., J\}$, which may depend on student-program-specific characteristics X_{ij}, distance from home, and its interactions with student characteristics, collected in D_{ij}. Allowing mean utilities to differ at the program level and assuming linear preferences with additive separability gives the following expression for utility from assignment to program j:

$$U_{ij} = \delta_j + X_{ij}\beta + D_{ij}\tau + \eta_{ij} = V_{ij} + \eta_{ij}, \quad (1)$$

where η_{ij} are unobserved idiosyncratic tastes.

Estimating the δ_j, β, and τ parameters requires making assumptions about preferences and students' submitted portfolios.[9] We estimate student preferences over programs under the assumption that the matching resulting from implementing the COMIPEMS assignment mechanism is stable. This follows Fack, Grenet, and He (2019), who show that stability is a weaker assumption than alternatives such as "truth-telling," which may not hold in our context. Specifically, "weak truth-telling" assumes that students truthfully rank programs and that all unlisted programs are less-preferred than all listed programs. In Mexico City, though elite options are highly demanded, students rarely exhaust all elite options in their ranked lists. Instead, they often list a few elite programs and then list less-competitive, neighborhood programs. This is more likely due to students' beliefs about their likelihood of acceptance into the marginal elite school, conditional on scoring too low for assignment to those listed, rather than students having lower preferences for unlisted elite options (Ali and Shorrer 2021).

In contrast to truth-telling approaches that place strong assumptions on the relationship between choice *portfolios* and the student preferences of the students submitting them, the stability assumption regards the *matching* that results from implementing the assignment mechanism. Stability assumes that students are assigned to their most-preferred program that was *ex post* feasible given their assignment priority. This is plausible in Mexico City, where information on program feasibility is available and students can list enough programs that they are assigned to their most-preferred feasible program. Specifically, program cutoff scores are stable from year to year, and these cutoffs are public for applicants to observe. Appendix Figure B.1 plots program-level cutoffs in 2008 compared to their cutoffs in 2007, the years to be analyzed. With few exceptions, these cutoff pairs lie close to the 45-degree line, and their correlation is 0.95, implying that students can forecast future cutoffs using existing information. Students do not know their placement test scores at the time of application, and Bobba and Frisancho (2020) show that COMIPEMS participants imperfectly predict and are uncertain about their scores. However, this uncertainty is mitigated by

9. Agarwal and Somaini (2020) provide a thorough overview of empirical approaches to revealed preference analysis in the context of school choice.

the ability for students to rank up to 20 programs at zero marginal monetary cost and low marginal time cost. Only 2.7% of students in the data exhaust all 20 program choices in their lists, suggesting that students are not constrained in their ability to submit a portfolio large enough that it results in assignment to their most-preferred feasible program, regardless of their assignment priority.

These institutional features suggest that estimating preferences under the stability assumption is reasonable. Fack, Grenet, and He (2019) show that, asymptotically, stability nests truth-telling—that is, even when students rank programs in order of preference and prefer listed programs to unlisted ones, empirical approaches based on stability alone still consistently recover preference parameters.[10]

Estimating student preference parameters is straightforward under the stability assumption. Here, student i's feasible programs are those whose cutoff scores c_j are less than or equal to her placement test score s_i: $\mathcal{J}_i = \left\{ j : s_i \geq c_j \right\}$.[11] Stability implies that among all programs in this "personalized choice set," her most-preferred is the one to which she was assigned, denoted by A_i:

$$A_i = \underset{j \in \mathcal{J}_i}{\operatorname{argmax}} \; U_{ij}.$$

We first construct each student's personalized choice set $\mathcal{J}_i$ using their placement test score and program-level cutoffs from the year in which they participated in COMIPEMS. Assuming that η_{ij} is i.i.d. extreme value type I, Fack, Grenet, and He (2019) show that Equation 1 can be estimated using a conditional logit with the personalized choice sets. We follow this approach while accounting for features specific to the institutional and empirical context. First, we allow for flexible preference heterogeneity across seven geographical regions r and exam years t by estimating preferences separately by region-year cell. We denote the region-year-specific observable utilities by $V_{ijrt} = \delta_{jrt} + X_{ij}\beta_{rt} + D_{ij}\tau_{rt}$. Second, non-elite programs of the same type (STEM, non-STEM

10. The drawback of assuming stability alone when truth-telling holds is that empirical approaches based on truth-telling use students' full portfolios for preference estimation, making them more efficient than stability-based approaches that only use the assigned program.

11. A 7.0 GPA or above is also required to be considered at UNAM and IPN programs, so it is part of assignment priority at those programs. We account for this in the empirical analysis by including the GPA cutoff in the feasibility requirement for these programs.

technical, and traditional) outside of a geographical buffer surrounding the region are aggregated into region-year-specific alternatives.[12] Finally, because some students remain unassigned by the computerized assignment process, we consider "unassigned" to be the outside option available in every student's choice set and normalize its utility to zero.

The conditional likelihood for student i's assignment outcome is:

$$\mathcal{L}\left(A_i | \boldsymbol{X}_i, \boldsymbol{D}_i, \boldsymbol{c}_t, s_i\right) = \frac{exp\left(V_{iA_irt}\right)}{\sum_{j \in \mathcal{J}_i} exp\left(V_{ijrt}\right)}, \tag{2}$$

where $\boldsymbol{X}_i$ and $\boldsymbol{D}_i$ are matrices containing the student-program covariates and student-distance covariates for all programs $j \in \mathcal{J}_i$, respectively; $\boldsymbol{c}_t$ contains all program cutoff scores in year t, and s_i is the student's placement test score. Student characteristics include middle school GPA, cubic functions of math and Spanish subscores from the low-stakes ENLACE 9 exam, a parental high school education indicator, indicators for missing ENLACE 9 score and missing parental education indicator, an indicator for whether the applicant has already graduated from middle school, each of which is de-meaned with respect to females in the region-year-specific estimation sample; a male indicator; and interactions between the male indicator and all aforementioned characteristics. Each student characteristic is interacted with the following program characteristics to form X_{ij}: a constant (to allow utility relative to the outside option to vary with student characteristics), and indicators for elite STEM, elite non-STEM, non-elite STEM, non-STEM technical, and traditional program types.[13] Similarly, D_{ij} consists of distance and its interactions with each student characteristic. Parameters are estimated via maximum likelihood.

12. The width of the buffer varies by region-year and is chosen such that 90% of non-elite assignments are to schools within the buffer. Buffer widths range from 2.7 km to 10.1 km. These two modeling choices are similar to those taken by Abdulkadirolu et al. (2020), who estimate preferences separately by borough and aggregate all schools outside a student's borough into an outside option.

13. Following Parsons and Needelman (1992) and Lupi and Feather (1998), we account for the inclusion of the aggregated far-away alternatives by including as a covariate $ln(n_{ij})$ with coefficient constrained to 1, where n_{ij} is the number of programs that were aggregated to form alternative j. For non-aggregated alternatives, $n_{ij} = 1$.

3.2 Willingness to travel

The estimated student preference parameters need to be transformed in order to aggregate them across region-years, summarize female preferences over STEM and non-STEM program types, and estimate how they differ from male preferences. Because the student characteristics are de-meaned, the estimated alternative-specific coefficients $\widehat{\delta}_{jrt}$ represent average program-specific utilities at the mean of female region-year characteristics. Denoting the estimated uninteracted distance coefficient by $\widehat{\tau}_{rt}^{0}$, willingness to travel for program j at the mean of female characteristics is thus:

$$\text{WTT}_{jrt} = -\frac{\widehat{\delta}_{jrt}}{\widehat{\tau}_{rt}^{0}}$$

Projecting these program-specific WTT on program type indicators allows us to estimate average marginal WTT for program types, pooling all region-years:

$$\begin{aligned} \text{WTT}_{jrt} = \alpha_{rt} + \gamma_1(\text{Elite STEM}_j) + \gamma_2(\text{Non-elite STEM}_j) + \\ \gamma_3(\text{Elite non-STEM}_j) + \gamma_4(\text{Technical non-STEM}_j) + \varepsilon_{jrt}, \end{aligned} \tag{3}$$

where the α_{rt} fixed effects allow the value of the outside option to differ by region-year and the γ coefficients give the average marginal WTT to the indicated program type, compared to the excluded category, traditional programs. This regression is weighted by the number of students in the region-year cell.

The gender difference in marginal WTT for program type p (compared to traditional programs) is denoted by ΔWTT^p and is estimated as a weighted average of region-year gender differences:

$$\Delta\text{WTT}^p = \sum_r \sum_t \frac{N_{rt}}{N} \Delta\text{WTT}_{rt}^p = \sum_r \sum_t \frac{N_{rt}}{N} \left(-\frac{\widehat{\beta}_{rt}^{\text{male} \times \text{p}}}{\widehat{\tau}_{rt}^{0} + \widehat{\tau}_{rt}^{\text{male}}} \right), \tag{4}$$

where $\widehat{\beta}_{rt}^{\text{male} \times \text{p}}$ is the estimated gender difference in marginal utility from program type p in region-year rt, $\widehat{\tau}_{rt}^{\text{male}}$ is the estimated gender difference in marginal utility from distance in region-year rt, N_{rt} is the number of students in region-year rt, and $N \equiv \sum_{rt} N_{rt}$. The variance of each component $Var(\Delta\text{WTT}^p_{rt})$ is computed using the delta method and the variance of the sum of these components is $Var(\Delta\text{WTT}^p) = \sum_r \sum_t \left(\frac{N_{rt}}{N}\right)^2 Var(\Delta\text{WTT}^p_{rt})$. These gender differences in WTT are at the mean of female characteristics. The approach is also extended to estimate gender differences at different points on the ENLACE 9 math and Spanish subscore distributions.

3.3 Counterfactual simulations

Simulating counterfactual STEM assignment gaps under alternative priority structures requires two steps. First, the estimated preference parameters and student-program characteristics are used to simulate preference orderings over *all* programs and the outside option of non-assignment, for each student in the sample. We draw unobserved tastes $\widehat{\eta}_{ij}$ from the extreme value type I distribution and rank programs by simulated utilities $\widehat{U}_{ijrt} = \widehat{V}_{ijrt} + \widehat{\eta}_{ijrt}$.[14]

Second, these preference orderings are submitted to the assignment mechanism, which is then implemented using the actual or counterfactual priority structure to obtain simulated program assignments.[15] This simulation differs from the actual COMIPEMS assignment process because it uses students' full preference orderings rather than shorter choice portfolios. This follows the approach of Artemov, Che, and He (2017), who note that this simulation approach is valid under the stability assumption.

We set program capacities (seat counts) to match those used in the respective year's actual assignment process.[16] While in the true COMIPEMS assignment process, ties are resolved by school system representatives in real-time, who either accept all tied applicants (exceeding capacity) or reject them all (leaving excess capacity), we keep the capacities fixed and use a single random tiebreaker. Once the assignment algorithm terminates, we summarize the simulated assignment outcomes, in particular the STEM gender gap.

14. Once these rankings are generated for the estimation sample, we combine them with the observed portfolios of the small sample of assignable students who were not included in the estimation sample. This results in "portfolios" for the universe of assignable students observed in the data.

15. For this step, the "programs" that aggregate far-away alternatives are replaced with a randomly drawn feasible alternative among those that were aggregated.

16. Capacities are unobserved for programs that did not fill up. We assume unlimited capacities for these programs, but results are similar if we fix capacities at the number of seats that were filled in that year.

After simulating the assignment gap under the status quo priority structure, we simulate several counterfactual priority structures, such as lottery assignment and STEM-specific affirmative action. In all cases, the mechanism remains SPDA, so that the assignment simulation is unchanged except for the assignment priorities that students have at programs.[17]

Finally, we simulate distance-denominated changes in overall and gender-specific average consumer surplus to show how counterfactual priority structures affect the level and distribution of welfare.[18] Expected surplus for each student i is a function of their estimated preference parameters and the set of programs that were *ex post* feasible given the results of implementing the assignment mechanism. Given preferences and a feasible choice set $\mathcal{J}_i$, Small and Rosen (1981) and Williams (1977) show that expected consumer surplus can be computed up to a constant:

$$\mathrm{E}\left(CS_{irt}\right) = \frac{1}{-\widehat{\tau}_{irt}} \ln\left(\sum_{j\in\mathcal{J}_i} \exp(\widehat{V}_{ijrt})\right) + C,$$

where $\widehat{\tau}_{irt}$ is individual i's estimated marginal utility from distance.[19] Differences in $\mathrm{E}\left(CS_{irt}\right)$ between two priority structures result from differences in the feasible sets they generate, $\mathcal{J}_i^0$ and $\mathcal{J}_i^1$:

$$\Delta\mathrm{E}\left(CS_{irt}\right) = \frac{1}{-\widehat{\tau}_{irt}} \left[\ln\left(\sum_{j\in\mathcal{J}_i^1} \exp(\widehat{V}_{ijrt})\right) - \ln\left(\sum_{j\in\mathcal{J}_i^0} \exp(\widehat{V}_{ijrt})\right)\right]. \tag{5}$$

Equation 5 is used to compute student-level changes in distance-denominated expected consumer surplus, which are then averaged in the full or gender-specific sample.

The results we present account for estimation error in the preference parameters, as well as simulation error, by reporting the means and standard deviations of 200 independent simulations. Each simulation takes independent draws from the joint preference parameter $(\widehat{\delta}_j, \widehat{\beta}, \widehat{\tau})$, idiosyncratic taste $(\widehat{\eta}_{ij})$, and tiebreaker distributions.[20]

17. Implementation of the minority reserve quota in a SPDA mechanism follows Hafalir, Yenmez, and Yildirim (2013).

18. The general approach and exposition here follow Train (2009).

19. Letting Z_i be the vector of all student characteristics (including a constant) that are interacted with distance to form D_{ij}, we have $\widehat{\tau}_{irt} = Z_i\widehat{\tau}_{rt}$.

20. This approach is similar to Pathak and Shi (2021), except that student characteristics are treated as fixed rather

4 Data

The data used to analyze students' school choices and assignments are drawn from the 2007 and 2008 COMIPEMS administrative databases, corresponding to the cohorts for which we also have matched student-level data from the low-stakes ENLACE 9 exam. The databases include each participating student's reported school choice list, with all programs (including specialties within technical schools) and their ranks. The data also include the exam scores for both the placement test and ENLACE test, with subcomponents for each subject, middle school GPAs, students' home postal codes, students' final assigned programs, and an explanation code if students were not assigned by the mechanism. In addition, we use demographic data on parental education levels from a survey completed by each student and turned in with the school choice list. In studying the gender gap, we use the male/female classifications in the COMIPEMS data, which are extracted from each student's national identification record. We do not have data on students' self-identified genders, and we do not have data on any non-binary gender categories. Thus, our analysis produces results for the STEM gap by sex, but we follow the literature in referring to this gap as the gender gap.

The analytic sample consists of students eligible for assignment and for whom we are able to estimate preferences.[21] We are unable to estimate preferences for students whose address is either invalid (missing home and middle school) or outside the COMIPEMS geographical boundary. For students with no home address but a middle school, we use the location of their middle school as their address. We also exclude adults who are returning to school, as they are substantially different from the majority of the students. Excluded students are 3.7% of all assignable students. The final sample includes students who are finishing middle school at the time of the exam and (overwhelmingly recent) middle school graduates, who are usually re-taking the exam. Analyses of the gap in assigned programs only include students in our analytic sample who were successfully assigned to a program (in actuality or according to the policy simulations).

than resampling them in each simulation.

21. Students are not eligible for assignment if they failed middle school or scored below a minimum score on the placement exam. They are dropped from our analysis.

We classify the high school programs as STEM or non-STEM by creating a cross-walk between the choice dataset and the Brookings STEM occupational classifications. Specifically, for each program, we identify all possible occupational matches and take the average STEM classification for them. We categorize the program as STEM if the average STEM classification is 0.5 or more. Each classification was done independently by two people, with discrepancies reconciled by a third individual. The program-level STEM classifications are presented in Appendix Table B.2. Examples of STEM programs include electricity and electronics, agro-industrial technician, and nursing, while examples of non-STEM programs include social work, tourism, business, and administration.

We supplement school choice data with geographic information system (GIS) data, which provides the location of each postal code and school.[22]

The student-level ENLACE 9 matching, performed on the basis of unique student identifiers, has a 96% match rate for students who participate in the COMIPEMS process in ninth grade. Students who have already graduated middle school do not take the exam, so they are left unmatched. Math and Spanish subscores are normalized with respect to the national score distribution.

Table 1 presents summary statistics for academic outcomes, school choices, and demographics by gender. With respect to academics, males significantly outperform females on the placement test on average, scoring 3.9 points (0.2 SD) higher than females. The low-stakes ENLACE 9 does not show the same disparity: males score 0.14 SD higher in math, while females score 0.21 SD higher in Spanish, such that overall performance is quite similar. Appendix Table B.3 presents regression results showing that the placement test and ENLACE 9 scores are highly correlated, with ENLACE 9 subscores explaining about 65% of the variation in placement test scores. Conditional on ENLACE 9 subscores, males have an even larger advantage in placement test results (0.24 SD) than in the uncontrolled comparison. Females obtain 0.54 SD higher middle school grade point averages than males. Figure 1 displays the densities of these four indicators. Panel A shows that the average difference in placement test scores is not driven by outliers, but rather by large differences

22. The sources include Mexico's National Institute of Statistics and Geography (INEGI) and the official COMIPEMS website for the location of each school, http://opciones.comipems.org.mx.

throughout the distribution. Panel D shows strikingly different gender-specific middle school GPA distributions.

With respect to choices, males choose STEM programs more than females. Specifically, on average 31.5% of male choice portfolios are STEM programs (10.5% elite STEM and 21.0% non-elite STEM), while only 21.6% of female choice portfolios are STEM programs (6.6% elite STEM and 15.0% non-elite STEM). In total, females list slightly more programs than males (9.5 versus 9.4), being more likely to list elite non-STEM and technical non-STEM schools on average. Males and females both list schools that are 8.8 km away on average. Males are slightly more likely to come from households in which at least one parent has a high school degree.

Table 2, Panel A summarizes program assignment by gender for students in the analysis sample who were assigned by the placement algorithm. 38.7% of males are assigned to STEM programs compared to 27.5% of females, generating a STEM gender gap of 11.1 p.p. This comes from 5.7 and 5.5 p.p. gender gaps in elite STEM and non-elite STEM programs, respectively. Females are 1.3 p.p. more likely to be assigned to elite non-STEM programs, 2.9 p.p. more likely to be assigned to technical non-STEM programs, and assigned to schools 0.2 km closer than males. Panel B shows that females are 5.7 p.p. more likely to be left unassigned by the algorithm, compared to a male non-assignment rate of 11.7%.

Figure 2 examines heterogeneity in the STEM gender gap by placement test score percentile. The total gap is higher than the 11.1 p.p. average for very low scorers and for high scorers, exceeding 15 p.p. at the top of the score distribution. The gap in non-elite STEM assignment dominates for low scores, which are below elite programs' assignment cutoffs. The gap in elite STEM assignment grows at higher scores, as assignment to more competitive programs becomes feasible.

5 Results

This section presents the results from estimating the school choice model. First, we briefly assess model fit. Next, we describe student preferences, followed by the counterfactual simulations of the

STEM gap and welfare under alternative priority structures.

5.1 Model fit

Estimated student preferences from the conditional logit model are summarized in Appendix Table B.4. We report preference parameters averaged over all region-year cells and weighted by number of students. Column 1 reports estimates from a model assuming stability. As a point of comparison, column 2 reports estimates from a model assuming weak truth-telling.[23] Performing the Hausman test described in Fack, Grenet, and He (2019) to evaluate these two models against each other, we overwhelmingly reject the truth-telling model ($\chi^2(4226) = 169050$, $p < 0.0001$).

Using the estimated student preferences to simulate preference orderings, and then implementing the assignment mechanism, results in assignment patterns that are nearly identical to those in the data. Table 3 compares the simulated and actual proportions assigned to different school types by gender and the resulting gender assignment gaps. The simulated overall STEM gap is 11.0 p.p., compared to the true 11.1 p.p. gap. The simulated proportions of male and female students assigned to STEM programs also match the data closely. Disaggregating the STEM gap, the simulated elite STEM gap of 5.6 p.p. nearly matches the actual 5.7 p.p. gap, while the simulated non-elite STEM gap of 5.4 p.p. is similar to the true 5.5 p.p. gap. The elite and technical non-STEM gaps are also similar to those in the data.

The simulated program-level cutoff scores correspond closely to those observed in the data. Figure 3 plots simulated program-by-year cutoff scores—the score of the lowest-scoring student assigned to programs that filled their capacities, or 31 (the minimum score for assignment) for programs that did not—against true cutoff scores. The points closely follow the 45 degree line, indicating that simulated cutoffs are close to those observed in the data. An assignment count-weighted univariate regression of simulated cutoffs on actual cutoffs yields a slope coefficient of 0.99 ($SE = 0.005$).

23. The weak truth-telling model corresponds to a rank-ordered logit on the full choice set (not limited by feasibility), with remaining unassigned as the outside option ranked $k + 1$ for any student ranking $k < 20$ programs.

5.2 Student preferences

We begin summarizing female students' preferences by presenting WTT estimates from equation 3, presented in Table 4. Column 1 includes a single intercept term, while column 2 includes region-year fixed effects that allow programs to be valued differentially compared to the outside option by estimation cell. The excluded category indicator is traditional schools, so that program type coefficients are one-way distance-denominated marginal willingness to travel (in kilometers) compared to traditional schools. Focusing on column 2, females value elite STEM schools highly compared to traditional schools, being on average willing to travel 13.2 km farther to them. The WTT for elite non-STEM programs is even higher, at 20.2 km. Preferences for non-elite programs go in the opposite direction: marginal WTT for non-elite STEM programs is -6.9 km, slightly more negative than the -5.4 km WTT for technical non-STEM programs.

Figure 4 illustrates estimated preference heterogeneity with respect to student math and Spanish ENLACE 9 subscores. Panel A shows that above-mean math subscores predict higher elite STEM preferences: all else equal, increasing the math subscore by 1 SD from the mean increases WTT for elite STEM schools by 1.2 km. This pattern does not hold for Spanish subscores, where there is no clear pattern. Similarly, in Panel B, higher math subscores predict higher preference for non-elite STEM programs: a 1 SD increase over the mean predicts a 0.34 km increase in WTT. Higher Spanish subscores, on the other hand, predict lower preference for non-elite STEM programs. In summary, higher math ability predicts stronger preferences for STEM programs, while higher Spanish ability does not.

The relationship between subscores and preferences for non-STEM programs is less straightforward, as shown in Panels C and D. Higher math scores predict higher WTT for both elite and technical non-STEM programs, recalling that the base category is traditional programs. The math subscore-preference gradient for elite non-STEM schools is flatter than for elite STEM programs: a 1 SD increase over the mean predicts a 0.3 km increase in WTT for elite non-STEM programs, compared to 1.2 km for elite STEM programs. Low Spanish subscores actually predict higher elite non-STEM preferences, while high subscores predict higher WTT to technical non-STEM programs. Overall, we note that the relationship between exam subscores and elite preference is modest, except in the case of math subscores and elite STEM programs, and that higher math subscores generally predict moderately higher preferences for programs other than those categorized as traditional academic.

Significant gender differences exist in preferences for STEM programs. Table 5 reports the results of estimating gender differences in WTT following equation 4.[24] Males have 4.0 km higher WTT to elite STEM programs than females, while WTT for non-elite STEM programs is 2.1 km higher for males. Preference differences for non-STEM programs are more moderate: males have 0.64 km lower WTT for elite non-STEM programs and 0.26 km higher WTT for technical non-STEM programs. Compared to the base category of traditional programs, males have higher preferences for the outside option of remaining unassigned. We find little evidence for large changes in gendered preferences along the math and Spanish subscore distributions, as illustrated in Appendix Figure B.2.

5.3 Counterfactual simulations

This section describes hypothetical changes to the priority structure and their simulated impacts on the STEM gender gap and welfare. In general, we find that priority structure changes only close the STEM gap when they explicitly advantage females in STEM programs (and not in other programs). Impacts on the gender gap and its composition of elite and non-elite components vary significantly across the placement test score distribution, requiring a nuanced evaluation of policy impacts on different types of students.

5.3.1 Alternative priority structures

Columns 1 through 3 of Table 6 present simulated changes in the STEM gap and its elite and non-elite components when the status quo priority structure is altered. Baseline gaps are given at

24. These comparisons are made at the mean of female covariates, so preference differences reported here are not due to gender differences in observable characteristics.

the bottom of the table for comparison. Simulated changes in gender-specific and overall welfare, averaging the expected surplus changes in equation 5, are in columns 4 through 6. Figure 5 shows how the gap and its components change at each point in the placement test score distribution, while Figure 6 gives changes in welfare along the score distribution.

First, determining assignment priorities using a single lottery *increases* the gender gap in STEM assignment by 2.2 p.p., an increase of 19.8% over the baseline gap of 11.0 p.p. While the elite component of the STEM gap decreases by 24.2% because females' exam score disadvantage is removed, the non-elite gap increases by 65.9%. This increase is consistent with the preference parameter estimates, which show that females on average prefer non-STEM programs. Panel A of Figure 5 shows that the elite STEM gap falls among students with high placement test scores, as males in these programs are disproportionately displaced by lower-scoring females.

Due to the removal of males' placement score advantage, average male welfare declines by the equivalent of 0.74 km, while female welfare increases by 0.33 km. Overall welfare declines by 0.19 km. Panel A of Figure 6 shows the significant redistribution of welfare from students with high scores, most of whom lose access to some programs because of their lottery draw, toward students with low scores, many of whom can now be assigned to elite and other high-cutoff programs. The reallocation of welfare across genders is almost imperceptible compared to the reallocation across the score distribution.

Modifying the priority structure to rank students based on an "academic index" equally weighting normalized placement test score and normalized middle school GPA, thus advantaging females on average, increases the STEM gap by 37.8%. Male welfare losses (0.97 km) and female welfare gains (0.85 km) are both larger in magnitude than under the lottery, but the overall welfare loss is small, at 0.03 km. The pattern of effects along the placement test score distribution in Figure 5, Panel B, illustrates the roles of preferences and access in generating the STEM gap. For students scoring too low for elite access, using the academic index increases the STEM gap by allowing some females to access their more-preferred non-STEM programs. But this priority structure also allows some females with scores at the margin of admission to elite STEM programs to be assigned there, reducing the elite STEM gap in this region of the score distribution. Welfare effects follow a similar pattern: female-male welfare changes are largest for scores near elite cutoffs. We note that lower-scoring students benefit on average from this policy because it dilutes the impact of the placement test on assignment priority, allowing them displace high-scoring students with consequent welfare reductions.

The effects of affirmative action policies depend on whether they advantage females at all programs ("point-based AA"), or only at STEM programs ("point-based STEM AA,"). Point-based AA is implemented by adding, for each female applicant, a constant number of points to the test score that determines assignment priority to all programs. A general point-based AA program granting females 4 extra points (enough to equalize means between genders) would widen the STEM gap by 1.8 p.p. (16.8%), with declines in the elite gap and large increases in the non-elite gap. A more aggressive 10-point (roughly 1/2 standard deviation of the exam score) AA program would widen the gap by 4.5 p.p. (40.9%), reducing the elite gap by 33.7% while more than doubling the non-elite gap. Thus, even when highly favoring females in the assignment process, AA programs are unable to close the elite STEM gap, while at the same time widening the overall gap. Overall welfare impacts are essentially zero, avoiding the welfare losses induced by lottery and academic index priority structures due to a positive correlation between placement test scores and WTT for high-cutoff programs. Reallocation of welfare from males to females is of course still present: in the 10-point AA structure, gender-specific effects exceed 1 km.

Examining STEM gap impacts along the placement test distribution, we see a similar but more pronounced pattern to those seen with the academic index priority structure. Lower-scoring females are able to access more-preferred programs, which are disproportionately non-STEM, increasing the gap. In score ranges that are marginal for elite program assignment, AA allows females to access elite (sometimes STEM) programs, displacing males. In very high score ranges, the preference effect again dominates, as females move out of elite STEM and into elite non-STEM. Welfare effects are again largest for students near the margin of assignment to elite programs.

In contrast to general AA, STEM AA closes the STEM gap and its components. We first model a point-based STEM AA, which gives students different priorities to STEM and non-STEM programs. Priority at non-STEM programs relies on the unmodified test score, but priority at STEM programs is determined by the test score plus a constant number of points added to females' scores. A 4-point STEM AA program would close the STEM gap by 3.9 p.p. (35.8%), reducing the elite gap by 2.5 p.p. (-45.3%) and the non-elite gap by 1.4 p.p. (-25.8%). The larger 10-point STEM AA program is sufficient to eliminate the gap almost entirely (-89.8%), while actually reversing the elite STEM gap from 5.6 p.p. to -0.9 p.p. This policy prevents the preference channel from directly moving females out of STEM programs and into more-preferred non-STEM programs, while still allowing the benefits of increased access to elite (STEM) programs. This is seen in the impacts along the placement test score distribution: in contrast to the general AA policy, STEM AA strongly reduces the STEM gap for low scorers, maintains elite STEM gap reductions for marginal scorers, and prevents flow out of elite STEM to elite non-STEM programs among the highest scorers. Welfare reallocation between genders is much smaller than under general AA: the aggressive 10-point STEM AA policy reduces male welfare by 0.34 km instead of 1.19 km in its general AA counterpart, increasing female welfare by 0.24 km instead of 1.09 km. Reallocation is strongest for scores close to elite STEM cutoffs, as marginal females displace marginal males.

Quota-based versions of STEM AA are also effective in closing the STEM gap. Here, we simulate minority reserves where programs give higher priority to females to the point where females fill up the allocated quota slots, and males and females compete for non-quota slots according to their placement score priority. In cases where there is insufficient female demand to fill up the allocated quota, the residual slots are given to remaining males according to their priority.[25] Reserving 50% of seats at STEM programs for females reduces the STEM gap by 7.5 p.p. (-67.8%), nearly closing the elite STEM gap (-98.9%) and closing the non-elite STEM gap by 35.3%.[26] Limiting the 50% reserve policy to elite STEM programs reduces the gap by 4.6 p.p.

25. In contrast to majority quota policies, minority reserves policies continue to satisfy weak group strategy-proofness and stability under SPDA mechanisms (Hafalir, Yenmez, and Yildirim 2013), allowing us to reasonably apply our preference model estimates.

26. The gap need not close entirely because female demand for certain programs may be insufficient to fill up the reserve slots. Additionally, there are more females assigned to programs than males.

(-41.6%), more than closing the elite STEM gap (-100.4%) while increasing the non-elite STEM gap as girls who list STEM schools are moved from non-elite STEM programs into elite STEM programs.[27] The welfare and distributional impacts are very similar to point-based STEM AA, with a 50% STEM quota and elite STEM quota resulting in overall welfare losses of 0.03 km and 0.02 km, respectively. When comparing similarly scaled policies, there are marginally larger aggregate welfare costs under minority reserves as females benefit slightly less (Appendix Figure B.3).

Overall, AA policies succeed in maintaining the importance of the placement test score in determining assignment (in contrast to lottery and academic index approaches), but the direction of their impact on the STEM gap depends crucially on whether they are targeted specifically toward STEM programs. Both point-based and quota versions of STEM AA are effective, with similar distributional effects.

5.3.2 Scaling STEM AA

In the previous sections, we explore specific levels of STEM AA that are effective in reducing the STEM gap to different extents. However, administrators can set the scale of STEM AA to achieve any target STEM gap change. Figure 7 shows the decrease in the total, elite, and non-elite STEM gaps associated with different levels of STEM AA. Panel A shows increasing levels of point-based STEM AA expressed in standard deviations of the placement test score. The total, elite, and non-elite STEM gaps decrease linearly with additional STEM AA points for females, indicating that there are no increasing or decreasing returns in STEM gap reductions for these values of STEM AA. Increasing the level of STEM AA by 0.1 standard deviations of the placement test score induces a 3.3% decrease in the total STEM gap compared to the baseline. To close the total STEM gap fully, females must be given points equivalent to nearly 0.6 standard deviations of the placement test score. At any given level, STEM AA causes larger reductions in the elite STEM gap relative to the total STEM gap, consistent with reducing the score-based access barriers to elite STEM programs. STEM AA also reduces that gap at non-elite STEM programs but has a smaller effect since demand

27. The elite STEM gap can close more than 100% when females fill up the reserve slots and are admitted into some of the non-reserve slots.

for these programs is lower and entrance is less constrained by placement exam scores.

Panel B shows increasing levels of minority reserve quotas expressed in percentage of seats reserved. Here, minority reserves set at 25% of seats or less have no impact on the STEM gap. This occurs because each program is already 25% female in the absence of a reserve or has insufficient female demand for the female reserve to cause a change. Increasing minority reserves above 25% decreases the elite STEM gap rapidly and linearly. Reserving an additional 1 p.p. of seats for females generates a 4 p.p. decrease in the elite STEM gap, with the elite STEM gap closing fully at around a 50% reserve. As with point-based STEM AA, minority reserve STEM AA has a smaller effect on the non-elite STEM gap. At these programs, the gap actually increases slightly at minority reserves between 25% and 45% as females at the margin of elite STEM admission move from non-elite STEM programs into elite STEM programs. Above 45%, increasing minority reserves also decreases the non-elite STEM gap rapidly and linearly, closing fully at a minority reserve level of approximately 60%. Taken together, the total STEM gap begins closing at a minority reserve of approximately 25% and closes fully at a minority reserve of approximately 55%.

To contextualize the scale of these policies, we highlight demand-side interventions shown to be effective at reducing the STEM gap and discuss the corresponding scales of STEM AA that generate the same total STEM gap reductions in our context.[28] Burgess et al. (2022) show that semi-external assessments in math generate a 20% reduction in the STEM gap, a reduction that is roughly equivalent to a 0.1 standard deviation point-based policy or a 40% minority reserve in Mexico City. Several policies, centered around exposure to female role models (Porter and Serra 2019), female teachers (Bottia et al. 2015), and female science advisors (Canaan and Mouganie 2021), reduce the STEM gap by 45% to 52%.[29] In Mexico City, the equivalent gap reduction is achieved by implementing a point-based STEM AA policy set at 0.3 standard deviations of the placement test score or a 48% minority reserve. Finally, other studies examine policies that can fully close the STEM gap in their respective contexts. These include a curriculum reform that bundled

28. See Appendix C for details on the specific outcomes and calculations of STEM gap reduction in these other studies.

29. There is variation in the effectiveness of related policies, with Breda et al. (2021) finding a much smaller reduction in the STEM gap following exposure to female role models in French high schools.

advanced math with a wider range of other classes (Joensen and Nielsen 2009)[30] and a scenario where high-performing females are placed with female teachers in their mandatory introductory STEM courses (Carrell, Page, and West 2010). Again, these effects are equivalent to setting a STEM AA policy of between 0.5 and 0.6 standard deviations or a 55% minority reserve. Together, these results show that, in contexts where STEM access remains a constraint, modest to moderately-sized STEM AA policies have the potential to generate STEM gap reductions commensurate with successful demand-side interventions. This is true even in contexts where preferences as gendered, as long as females continue to list STEM programs somewhere in their choice portfolios.

6 Discussion

We used a structural model of preferences and policy simulations to show that the STEM gender gap is generated by a complex interaction between choices and admissions constraints. Much of the overall gap is attributable to gendered preferences, but females choosing highly-demanded STEM schools continue to be constrained by test-based admissions. Together, this implies that various policies have the potential to decrease the STEM gap substantially, but with heterogeneous effects by student achievement and school type.

The findings have important implications for the design of policies that aim to increase females' participation in STEM education, especially in the context of centralized admissions systems. As the general affirmative action results show, when the STEM gap is driven by gendered preferences, giving females higher assignment priority can actually widen the gap. On the other hand, STEM-specific affirmative action has the potential to close the gap. This strategy is particularly effective for elite STEM schools, for which there is significant demand among females who score just below the margin for admission to them. Gender-specific (and minority-based) affirmative action already exists in some places (Dahl, Rooth, and Stenberg 2021; Francis-Tan and Tannuri-Pianto 2018) suggesting that it could be feasibly applied in others, though political considerations may present

30. Several other studies examining curricular reforms have not been effective in reducing the STEM gap. See Biewen and Schwerter (2022) for a review.

large barriers in certain contexts.

Still, when differences in assignment outcomes are deeply rooted in gendered preferences, using changes to the priority structure to address the resulting gap is likely to be insufficient. Here, we point to a large body of complementary studies that suggest that choices are highly mutable, and demand-side interventions can be effective. For example, interventions providing information about school quality (Hastings and Weinstein 2008), student-specific achievement (Bobba and Frisancho 2020), and major-specific wage returns (Wiswall and Zafar 2014) have been shown to change students choices, and interventions with more interactive curricula have increased students interest in STEM (Polikoff et al. 2018). Studies also show that females can be encouraged to choose and continue in STEM fields with more exposure to applied science courses (Gottfried and Bozick 2016), female role models (Carrell, Page, and West 2010; Breda et al. 2018; Lim and Meer 2019; Porter and Serra 2019), female peers (Schneeweis and Zweimüller 2012), and positive feedback with respect to course grades (Owen 2010).

This paper also illustrates the utility and feasibility of using microsimulation to approximate the effects of priority structure changes. Because assignments (and the resulting STEM gap) depend on the gender-specific joint distributions of stated preferences and placement scores, it is difficult to predict the impact of even simple interventions without simulating them. This is particularly true when attempting to understand policy effects along the achievement distribution, where average effects for low-achieving students will be very different than for high-achieving ones. Furthermore, the structural model of student preferences allows us to estimate policy impacts on welfare in addition to impacts on school assignments. The usefulness of the modeling and simulation approach goes beyond the STEM gap studied here: it can be used to forecast a variety of school-, group-, or neighborhood-level changes in assignments and welfare resulting from proposed policies.

There are limitations to this approach, however. Specifically, the model and simulations that we estimate are based on choices made under an existing system. In practice, choices are not innate and may change in response to policies, as students have more experiences in STEM or change their demand for schools in response to changes in the gender composition of schools. Thus, the long-run general equilibrium effects of the analyzed policies may differ markedly from the simulations. Furthermore, despite the fact that the mechanism in this paper is strategy-proof, a change in the priority structure may induce changes in preferences beyond what we predict (Bobba and Frisancho 2020). While the findings in Pathak and Shi (2021) suggest that preferences are stable in response to assignment policy reforms, whether this holds for affirmative action policies is an interesting avenue for future research.

References

Abdulkadirolu, Atila, Parag A Pathak, Jonathan Schellenberg, and Christopher R Walters. 2020. “Do parents value school effectiveness?” *American Economic Review* 110 (5): 1502–39.

Agarwal, Nikhil, and Paulo Somaini. 2018. “Demand analysis using strategic reports: An application to a school choice mechanism.” *Econometrica* 86 (2): 391–444.

———. 2020. “Revealed preference analysis of school choice models.” *Annual Review of Economics* 12:471–501.

Ajayi, Kehinde F. 2013. “School choice and educational mobility: Lessons from secondary school applications in Ghana.” *IED Working Paper* 259.

Ali, S Nageeb, and Ran I Shorrer. 2021. *The College Portfolio Problem.* Technical report. Working Paper.

Artemov, Georgy, Yeon-Koo Che, and Yinghua He. 2017. “Strategic mistakes: Implications for market design research.” *NBER Working Paper.*

Avitabile, Ciro, Matteo Bobba, and Marco Pariguana. 2017. *High School Track Choice and Liquidity Constraints: Evidence from Urban Mexico.* Technical report. IZA Discussion Papers.

Beede, David N, Tiffany A Julian, David Langdon, George McKittrick, Beethika Khan, and Mark E Doms. 2011. “Women in STEM: A gender gap to innovation.” *Economics and Statistics Administration Issue Brief,* nos. 04-11.

Bertrand, Marianne, Rema Hanna, and Sendhil Mullainathan. 2010. “Affirmative action in education: Evidence from engineering college admissions in India.” *Journal of Public Economics* 94 (1-2): 16–29.

Beuermann, Diether, C. Kirabo Jackson, Laia Navarro-Sola, and Francisco Pardo. 2018. *What is a Good School, and Can Parents Tell? Evidence on the Multidimensionality of School Output.* Working Paper, Working Paper Series 25342. National Bureau of Economic Research, December. https://doi.org/10.3386/w25342. http://www.nber.org/papers/w25342.

Biewen, Martin, and Jakob Schwerter. 2022. “Does more maths and natural sciences in high school increase the share of female STEM workers? Evidence from a curriculum reform.” *Applied Economics* 54 (16): 1889–1911.

Bobba, Matteo, and Veronica Frisancho. 2020. “Self-perceptions about academic achievement: Evidence from Mexico City.” *Journal of Econometrics,* ISSN: 0304-4076. https://doi.org/https://doi.org/10.1016/j.jeconom.2020.06.009. https://www.sciencedirect.com/science/article/pii/S0304407620302724.

Bordón, Paola, Catalina Canals, and Alejandra Mizala. 2020. “The gender gap in college major choice in Chile.” *Economics of Education Review* 77:102011.

Bottia, Martha Cecilia, Elizabeth Stearns, Roslyn Arlin Mickelson, and Stephanie Moller. 2018. “Boosting the numbers of STEM majors? The role of high schools with a STEM program.” *Science Education* 102 (1): 85–107.

Bottia, Martha Cecilia, Elizabeth Stearns, Roslyn Arlin Mickelson, Stephanie Moller, and Lauren Valentino. 2015. “Growing the roots of STEM majors: Female math and science high school faculty and the participation of students in STEM.” *Economics of Education Review* 45:14–27.

Breda, Thomas, Julien Grenet, Marion Monnet, and Clémentine Van Effenterre. 2018. “Can female role models reduce the gender gap in science? Evidence from classroom interventions in French high schools.”

———. 2021. “Do female role models reduce the gender gap in science? Evidence from French high schools.”

Burgess, Simon, Daniel Sloth Hauberg, Beatrice Schindler Rangvid, and Hans Henrik Sievertsen. 2022. “The importance of external assessments: High school math and gender gaps in STEM degrees.” *Economics of Education Review* 88:102267.

California Community Colleges. 2018. *What is AB 705?* Technical report.

Calsamiglia, Caterina, Chao Fu, and Maia Güell. 2020. "Structural estimation of a model of school choices: The boston mechanism versus its alternatives." *Journal of Political Economy* 128 (2): 642–680.

Canaan, Serena, and Pierre Mouganie. 2021. "The impact of advisor gender on female students STEM enrollment and persistence." *Journal of Human Resources,* 0320–10796R2.

Card, David, and A Abigail Payne. 2021. "High school choices and the gender gap in STEM." *Economic Inquiry* 59 (1): 9–28.

Carrell, Scott E, Marianne E Page, and James E West. 2010. "Sex and science: How professor gender perpetuates the gender gap." *The Quarterly Journal of Economics* 125 (3): 1101–1144.

Ceci, Stephen J, Donna K Ginther, Shulamit Kahn, and Wendy M Williams. 2014. "Women in academic science: A changing landscape." *Psychological Science in the Public Interest* 15 (3): 75–141.

Corcoran, Sean Patrick, and E Christine Baker-Smith. 2018. "Pathways to an elite education: Application, admission, and matriculation to New York City's specialized high schools." *Education Finance and Policy* 13 (2): 256–279.

Cullen, Julie Berry, Mark C Long, and Randall Reback. 2013. "Jockeying for position: Strategic high school choice under Texas' top ten percent plan." *Journal of Public Economics* 97:32–48.

Dahl, Gordon B, Dan-Olof Rooth, and Anders Stenberg. 2021. "High school majors and future earnings." *American Economic Journal: Applied Economics.*

Dougherty, Shaun, and Isabel Harbaugh Macdonald. 2019. "Can growth in the availability of STEM technical education improve equality in participation?: evidence from Massachusetts." *Journal of Vocational Education & Training,* 1–24.

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

Dustan, Andrew, and Diana KL Ngo. 2018. "Commuting to educational opportunity? School choice effects of mass transit expansion in Mexico City." *Economics of Education Review* 63:116–133.

Eisenhart, Margaret, Lois Weis, Carrie D Allen, Kristin Cipollone, Amy Stich, and Rachel Dominguez. 2015. “High school opportunities for STEM: Comparing inclusive STEM-focused and comprehensive high schools in two US cities.” *Journal of Research in Science Teaching* 52 (6): 763–789.

Fack, Gabrielle, Julien Grenet, and Yinghua He. 2019. “Beyond truth-telling: Preference estimation with centralized school choice and college admissions.” *American Economic Review* 109 (4): 1486–1529.

Francis-Tan, Andrew, and Maria Tannuri-Pianto. 2018. “Black Movement: Using discontinuities in admissions to study the effects of college quality and affirmative action.” *Journal of Development Economics* 135:97–116.

Frisancho, Veronica, and Kala Krishna. 2016. “Affirmative action in higher education in India: targeting, catch up, and mismatch.” *Higher Education* 71 (5): 611–649.

Fryer Jr, Roland G, and Glenn C Loury. 2005. “Affirmative action and its mythology.” *Journal of Economic Perspectives* 19 (3): 147–162.

Gale, David, and Lloyd S Shapley. 1962. “College admissions and the stability of marriage.” *The American Mathematical Monthly* 69 (1): 9–15.

George-Jackson, Casey. 2011. “STEM switching: Examining departures of undergraduate women in STEM fields.” *Journal of Women and Minorities in Science and Engineering* 17 (2).

Gewertz, Catherine. 2018. “What is Career and Technical Education, Anyway?” Accessed August 23, 2019, *Education Week* (July). https://www.edweek.org/ew/issues/career-technical-education/index.html.

Glazerman, Steven, and Dallas Dotter. 2017. “Market signals: Evidence on the determinants and consequences of school choice from a citywide lottery.” *Educational Evaluation and Policy Analysis* 39 (4): 593–619.

Gnagey, Jennifer, and Stéphane Lavertu. 2016. “The impact of inclusive STEM high schools on student achievement.” *AERA Open* 2 (2): 2332858416650870.

Gottfried, Michael A, and Robert Bozick. 2016. “Supporting the STEM pipeline: Linking applied STEM course-taking in high school to declaring a STEM major in college.” *Education Finance and Policy* 11 (2): 177–202.

Hafalir, Isa E, M Bumin Yenmez, and Muhammed A Yildirim. 2013. “Effective affirmative action in school choice.” *Theoretical Economics* 8 (2): 325–363.

Hastings, Justine, Thomas J Kane, and Douglas O Staiger. 2009. “Heterogeneous preferences and the efficacy of public school choice.” *NBER Working Paper* 2145:1–46.

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

Hyde, Janet S, Sara M Lindberg, Marcia C Linn, Amy B Ellis, and Caroline C Williams. 2008. “Gender similarities characterize math performance.” *Science* 321 (5888): 494–495.

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

Jaschik, Scott. 2019. “New SAT Score: Adversity.” Accessed August 22, 2019, *Inside Higher Ed* (May). https://www.insidehighered.com/admissions/article/2019/05/20/college-board-will-add-adversity-score-everyone-taking-sat.

Joensen, Juanna Schrøter, and Helena Skyt Nielsen. 2009. “Is there a causal effect of high school math on labor market outcomes?” *Journal of Human Resources* 44 (1): 171–198.

Kahn, Shulamit, and Donna Ginther. 2017. *Women and STEM.* Technical report. National Bureau of Economic Research.

Kapor, Adam J, Christopher A Neilson, and Seth D Zimmerman. 2020. “Heterogeneous beliefs and school choice mechanisms.” *American Economic Review* 110 (5): 1274–1315.

Karp, Sarah. 2021. “Big changes in how students are picked for CPS elite high schools start today.” *WBEZ Chicago National Public Radio* (October 13, 2021). Accessed October 18, 2021. https://www.wbez.org/stories/big-changes-in-how-students-are-picked-for-cps-elite-high-schools-start-today/27f4345f-4ed6-4f4e-a215-d577956504fd.

Lichtenberger, Eric, and Casey George-Jackson. 2013. “Predicting High School Students’ Interest in Majoring in a STEM Field: Insight into High School Students’ Postsecondary Plans.” *Journal of Career and Technical Education* 28 (1): 19–38.

Lim, Jaegeum, and Jonathan Meer. 2019. “Persistent effects of teacher-student gender matches.” *Journal of Human Resources,* 0218–9314R4.

Lupi, Frank, and Peter M Feather. 1998. “Using partial site aggregation to reduce bias in random utility travel cost models.” *Water resources research* 34 (12): 3595–3603.

Means, Barbara, Haiwen Wang, Viki Young, Vanessa L Peters, and Sharon J Lynch. 2016. "STEM-focused high schools as a strategy for enhancing readiness for postsecondary STEM programs." *Journal of Research in Science Teaching* 53 (5): 709–736.

Ortega Hesles, Maria Elena, and Shaun M. Dougherty. 2017. *Academic Program Choice in Secondary Education: Regression Discontinuity Evidence from Mexico City.* Technical report. American Education Finance and Policy Conference Paper.

Owen, Ann L. 2010. "Grades, gender, and encouragement: A regression discontinuity analysis." *The Journal of Economic Education* 41 (3): 217–234.

Parsons, George R, and Michael S Needelman. 1992. "Site aggregation in a random utility model of recreation." *Land economics,* 418–433.

Pathak, Parag A, and Peng Shi. 2021. "How well do structural demand models work? Counterfactual predictions in school choice." *Journal of Econometrics* 222 (1): 161–195.

Polikoff, Morgan, Q Tien Le, Robert W. Danielson, Gale M. Sinatra, and Julie A. Marsh. 2018. "The impact of speedometry on student knowledge, interest, and emotions." *Journal of Research on Educational Effectiveness* 11 (2): 217–239.

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

Porter, Catherine, and Danila Serra. 2019. "Gender differences in the choice of major: The importance of female role models." *American Economic Journal: Applied Economics.*

Rothwell, Jonathan. 2013. *The hidden STEM economy.* Metropolitan Policy Program at Brookings.

Saygin, Perihan O. 2020. "Gender bias in standardized tests: evidence from a centralized college admissions system." *Empirical Economics* 59 (2): 1037–1065.

Schneeweis, Nicole, and Martina Zweimüller. 2012. "Girls, girls, girls: Gender composition and female school choice." *Economics of Education Review* 31 (4): 482–500.

Shapiro, Eliza. 2020. "New York City Will Change Many Selective Schools to Address Segregation." *The New York Times* (December 18, 2020). Accessed October 18, 2021. https://www.nytimes.com/2020/12/18/nyregion/nyc-schools-admissions-segregation.html.

Shi, Ying. 2018. "The puzzle of missing female engineers: Academic preparation, ability beliefs, and preferences." *Economics of Education Review* 64:129–143.

Small, Kenneth A, and Harvey S Rosen. 1981. "Applied welfare economics with discrete choice models." *Econometrica: Journal of the Econometric Society,* 105–130.

Stoet, Gijsbert, and David C Geary. 2013. "Sex differences in mathematics and reading achievement are inversely related: Within-and across-nation assessment of 10 years of PISA data." *PloS one* 8 (3): e57988.

Train, Kenneth E. 2009. *Discrete choice methods with simulation.* Cambridge university press.

Treschan, Lazar, Apurva Mehrotra, Damon Hewitt, Rachel Kleinman, NAACP Legal Defense, and Educational Fund Inc. 2013. *The Meaning of Merit: Alternatives to Determining Admission to NYC's Specialized High Schools.* Technical report. Community Service Society.

Williams, Huw CWL. 1977. "On the formation of travel demand models and economic evaluation measures of user benefit." *Environment and planning A* 9 (3): 285–344.

Wiswall, Matthew, Leanna Stiefel, Amy Ellen Schwartz, and Jessica Boccardo. 2014. "Does attending a STEM high school improve student performance? Evidence from New York City." *Economics of Education Review* 40:93–105.

Wiswall, Matthew, and Basit Zafar. 2014. "Determinants of college major choice: Identification using an information experiment." *The Review of Economic Studies* 82 (2): 791–824.

Zafar, Basit. 2013. "College major choice and the gender gap." *Journal of Human Resources* 48 (3): 545–595.

Tables

Table 1: Academic, choice, and demographic summary statistics, by gender

	(1) Male	(2) Female	(3) Difference
Academic achievement			
Placement test score	67.85	63.95	3.90***
	(19.52)	(18.74)	(0.05)
ENLACE 9 math subscore (normalized)	0.54	0.39	0.14***
	(1.04)	(0.99)	(0.00)
ENLACE 9 Spanish subscore (normalized)	0.44	0.65	-0.21***
	(0.92)	(0.87)	(0.00)
Middle school GPA (normalized)	-0.28	0.26	-0.54***
	(0.96)	(0.97)	(0.00)
School choices			
Total number of choices listed	9.36	9.54	-0.18***
	(3.73)	(3.75)	(0.01)
Percent of choices that are elite STEM	10.53	6.56	3.97***
	(18.02)	(13.01)	(0.04)
Percent of choices that are non-elite STEM	20.98	15.01	5.97***
	(24.31)	(20.93)	(0.06)
Percent of choices that are elite non-STEM	30.10	35.21	-5.11***
	(30.41)	(31.35)	(0.09)
Percent of choices that are technical non-STEM	9.35	10.45	-1.10***
	(15.33)	(16.44)	(0.05)
Percent of choices that are traditional academic	29.03	32.77	-3.74***
	(26.89)	(27.85)	(0.08)
Average distance to all choices (km)	8.78	8.76	0.02*
	(4.70)	(4.59)	(0.01)
Demographics			
High parental education (high school or more)	51.02	48.37	2.65***
	(49.99)	(49.97)	(0.15)
Observations	238749	255875	494624

Note: Calculations are for all students in the analysis sample for the the 2007 and 2008 COMIPEMS cycles. Standard deviations are in parentheses in columns (1) and (2); standard errors are in parentheses in column (3). Placement test score is raw score out of 128. ENLACE 9 subscores are nationally normed. Middle school GPA is normalized by year within the analysis sample.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 2: School assignment summary statistics

	(1) Male	(2) Female	(3) Difference
Panel A. Program type conditional on assignment			
STEM assigned program (elite or non-elite)	38.7	27.5	11.1***
	(48.7)	(44.7)	(0.1)
Elite STEM assigned program	10.3	4.6	5.7***
	(30.3)	(20.9)	(0.1)
Non-elite STEM assigned program	28.4	23.0	5.5***
	(45.1)	(42.1)	(0.1)
Elite non-STEM assigned program	17.2	18.5	-1.3***
	(37.8)	(38.8)	(0.1)
Technical non-STEM assigned program	12.8	15.7	-2.9***
	(33.4)	(36.4)	(0.1)
Traditional academic assigned program	31.3	38.2	-6.9***
	(46.4)	(48.6)	(0.1)
Distance to assigned program	7.2	7.0	0.2***
	(6.2)	(6.1)	(0.0)
Observations	210773	211336	422109
Panel B. Assignment			
Unassigned by algorithm	11.7	17.4	-5.7***
	(32.2)	(37.9)	(0.1)
Observations	238749	255875	494624

Note: Calculations in Panel A are for all students in the analysis sample who were assigned to a program by the placement algorithm in the 2007 or 2008 COMIPEMS cycles. Calculations in Panel B do not condition on assignment. Indicator variables are percentages. Standard deviations are in parentheses in columns (1) and (2); standard errors are in parentheses in column (3).

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3: Model fit: simulated and actual gender gaps in choices and assignments

	Simulated			Actual		
	(1) Male	(2) Female	(3) Difference	(4) Male	(5) Female	(6) Difference
STEM	38.5	27.5	11.0	38.7	27.5	11.1
	(0.10)	(0.10)	(0.18)			
Elite STEM	10.2	4.6	5.6	10.3	4.6	5.7
	(0.05)	(0.05)	(0.10)			
Non-elite STEM	28.3	23.0	5.4	28.4	23.0	5.5
	(0.11)	(0.10)	(0.19)			
Elite non-STEM	17.1	18.3	−1.2	17.2	18.5	−1.3
	(0.05)	(0.05)	(0.10)			
Technical non-STEM	12.8	15.8	−3.0	12.8	15.7	−2.9
	(0.08)	(0.09)	(0.14)			
Traditional academic	31.5	38.3	−6.8	31.3	38.2	−6.9
	(0.10)	(0.10)	(0.19)			

Note: Columns 1 through 3 report the simulated gender-specific proportions of assigned students who were assigned to the indicated program type, and their difference. Proportions are means over 200 independent simulations of the assignment process accounting for uncertainty in student preference parameters, idiosyncratic student preferences, and random tie-breakers in assignment. Standard deviations of the 200 simulated proportions are in parentheses. Columns 4 through 6 show the actual proportions in the data. Proportions are reported in percentages. Simulations are as described in Section 3.3, using estimated student preferences from the procedure described in Section 3.1 and the status quo priority structure. Sample is 2007 and 2008 COMIPEMS cycles.

Table 4: Marginal willingness to travel to different program types, females

	(1)	(2)
Elite STEM	12.71***	13.16***
	(1.157)	(1.158)
Non-elite STEM	-7.06***	-6.90***
	(0.538)	(0.535)
Elite non-STEM	19.74***	20.23***
	(0.852)	(0.849)
Technical non-STEM	-5.66***	-5.36***
	(0.580)	(0.579)
Constant	-1.19**	
	(0.476)	
Observations	2578	2578
Adjusted R^2	0.690	0.734
Fixed effects	None	Region-year

Note: Dependent variable is estimated region-year program-level willingness-to-travel (WTT), in kilometers, using the estimated preferences from the procedure described in Section 3.1 and transforming them to WTT following Section 3.2. Regression weights by number of students in region-year estimation cell. Base category is traditional academic program (non-elite, non-STEM). Standard errors clustered at the region-year estimation cell level are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: Gender differences in marginal willingness to travel to different program types

	(1)
Elite STEM	3.97***
	(0.135)
Non-elite STEM	2.10***
	(0.057)
Elite non-STEM	-0.64***
	(0.142)
Technical non-STEM	0.26***
	(0.067)
Unassigned	0.45***
	(0.077)

Note: Entries are estimated differences between males and females in mean marginal willingness-to-travel (WTT) for for the indicated program characteristics, in kilometers. These are obtained by transforming the *male* × *characteristic* coefficient to a WTT estimate in each region-year estimation cell, then computing the mean of these estimates, weighting by the number of students in the cell. Standard errors of these estimates are in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01.

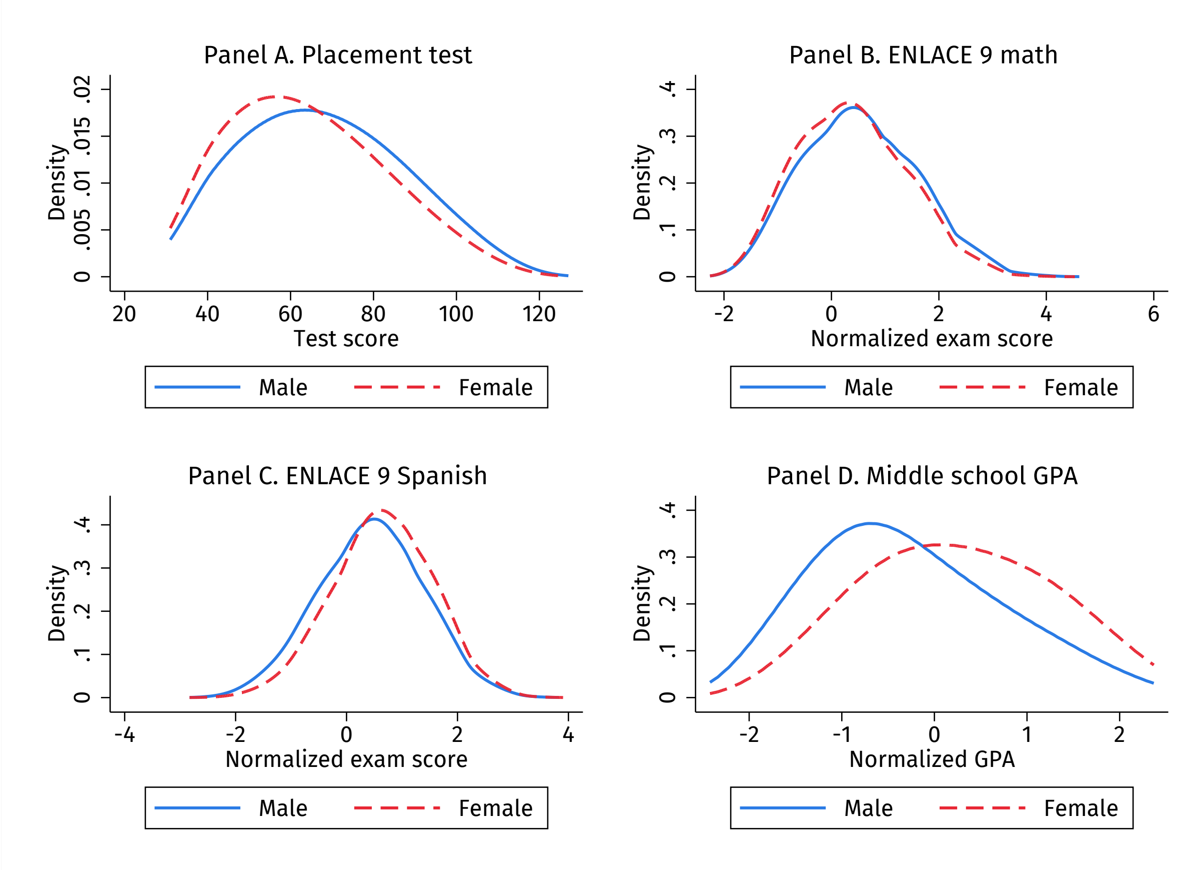
Table 6: Effects of simulated priority structure changes on STEM gender gap and welfare

	(1) Total gap	(2) Elite gap	(3) Non-elite gap	(4) Male welfare	(5) Female welfare	(6) Overall welfare
Lottery	2.2	−1.4	3.5	−0.74	0.33	−0.19
	(0.20)	(0.16)	(0.16)	(0.02)	(0.02)	(0.01)
	[19.8%]	[−24.2%]	[65.9%]			
Academic index	4.2	−1.7	5.9	−0.97	0.85	−0.03
	(0.13)	(0.08)	(0.11)	(0.01)	(0.01)	(0.01)
	[37.8%]	[−30.3%]	[109.1%]			
Point-based AA (4 pt)	1.8	−0.7	2.6	−0.47	0.44	0.00
	(0.07)	(0.04)	(0.05)	(0.00)	(0.00)	(0.00)
	[16.8%]	[−12.9%]	[47.9%]			
Point-based AA (10 pt)	4.5	−1.9	6.4	−1.19	1.09	−0.01
	(0.10)	(0.06)	(0.09)	(0.01)	(0.01)	(−0.00)
	[40.9%]	[−33.7%]	[119.2%]			
Point-based STEM AA (4 pt)	−3.9	−2.5	−1.4	−0.13	0.10	−0.01
	(0.05)	(0.03)	(0.04)	(0.00)	(0.00)	(0.00)
	[−35.8%]	[−45.3%]	[−25.8%]			
Point-based STEM AA (10 pt)	−9.9	−6.5	−3.4	−0.34	0.24	−0.04
	(0.07)	(0.05)	(0.06)	(0.00)	(0.00)	(0.00)
	[−89.8%]	[−115.1%]	[−63.3%]			
STEM quotas	−7.5	−5.6	−1.9	−0.25	0.18	−0.03
	(0.10)	(0.10)	(0.09)	(0.00)	(0.00)	(0.00)
	[−67.8%]	[−98.9%]	[−35.3%]			
Elite STEM quotas	−4.6	−5.6	1.1	−0.16	0.11	−0.02
	(0.08)	(0.09)	(0.03)	(0.00)	(0.00)	(0.00)
	[−41.6%]	[−100.4%]	[20.0%]			
Male preferences	−11.8	−4.7	−7.1			
	(0.16)	(0.09)	(0.15)			
	[−107.2%]	[−82.7%]	[−133.0%]			
Baseline gap	11.0	5.6	5.4			
	(0.18)	(0.10)	(0.19)			

Note: Rows represent alternative priority structures. Columns 1 through 3 are simulated changes in the respective STEM gap, compared to the baseline gap. Changes in STEM gaps are in percentage points. Simulated changes are means over 200 independent simulations of the assignment process accounting for uncertainty in student preference parameters, idiosyncratic student preferences, and random tie-breakers in assignment. Standard deviations of the 200 simulated changes are in parentheses. Percent changes compared to baseline are in brackets. Baseline gap is the simulated level of the respective STEM gap under the status quo priority structure. Simulated welfare changes in columns 4 through 6, estimated using the procedure described in Section 3.3, are in kilometers, with standard deviations of the 200 simulated changes in parentheses.

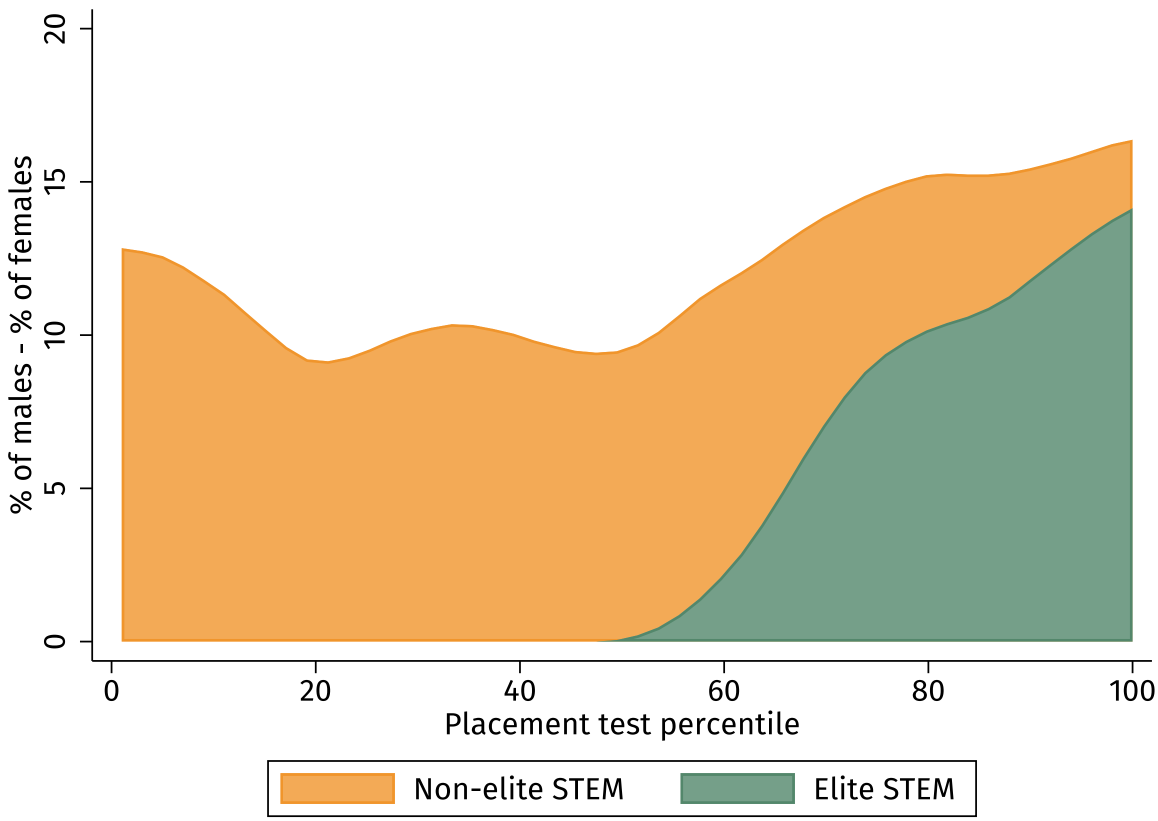
Figures

Figure 1: Academic achievement distributions, by gender



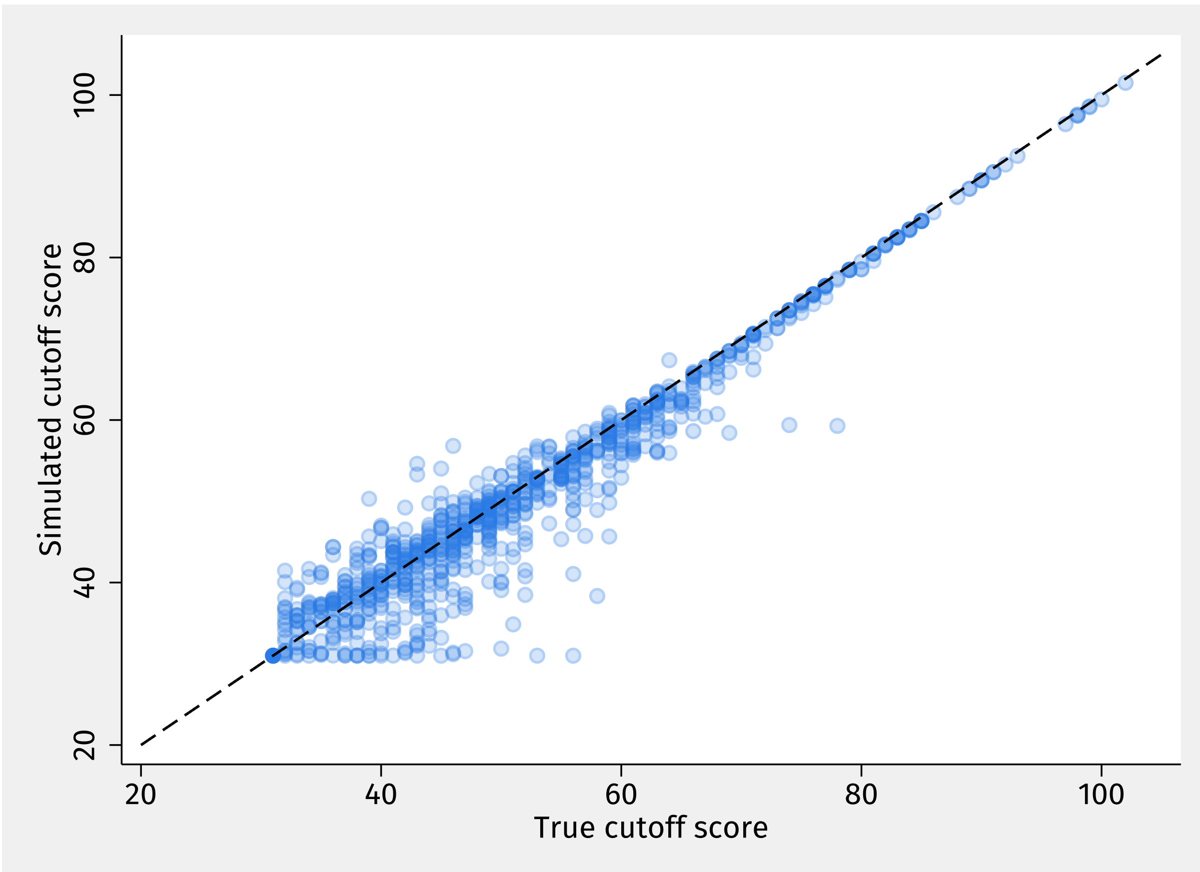
Note: Calculations are for all students in analysis sample in the 2007 and 2008 COMIPEMS cycles. Placement test score is raw score out of 128. ENLACE 9 subscores are nationally normed. Middle school GPA is normalized by year within the analysis sample.

Figure 2: STEM gender gap in program assignment by placement test percentile



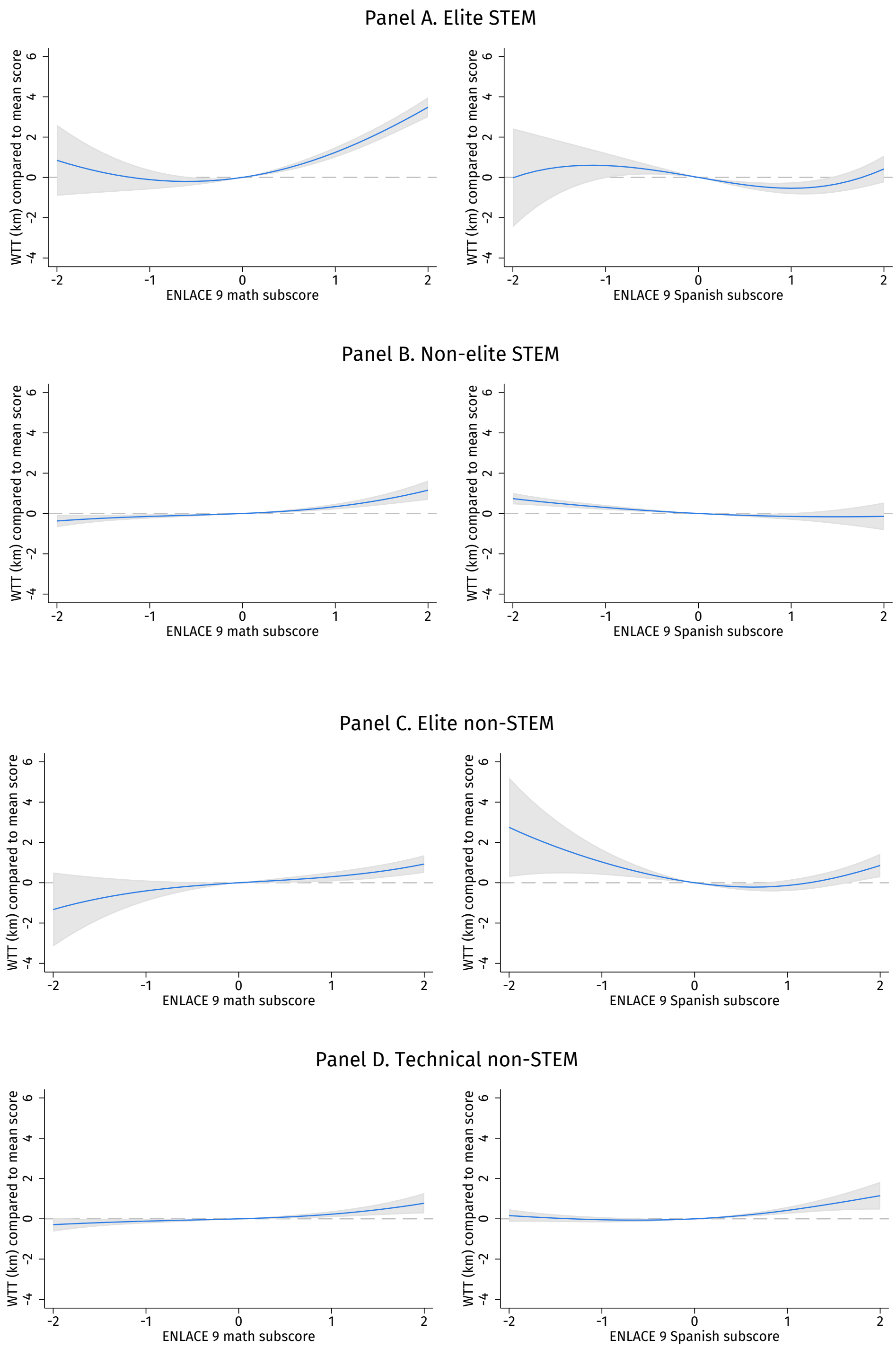
Note: The non-elite and elite STEM gap components are stacked, so that the overall STEM gap is represented by the top of the stack. Calculated percentages are for all students in analysis sample assigned to a program by the placement algorithm in the 2007 and 2008 COMIPEMS cycles. Test score percentiles are from the pooled (male and female) distribution of scores within each year.

Figure 3: Model fit: simulated versus actual program cutoffs



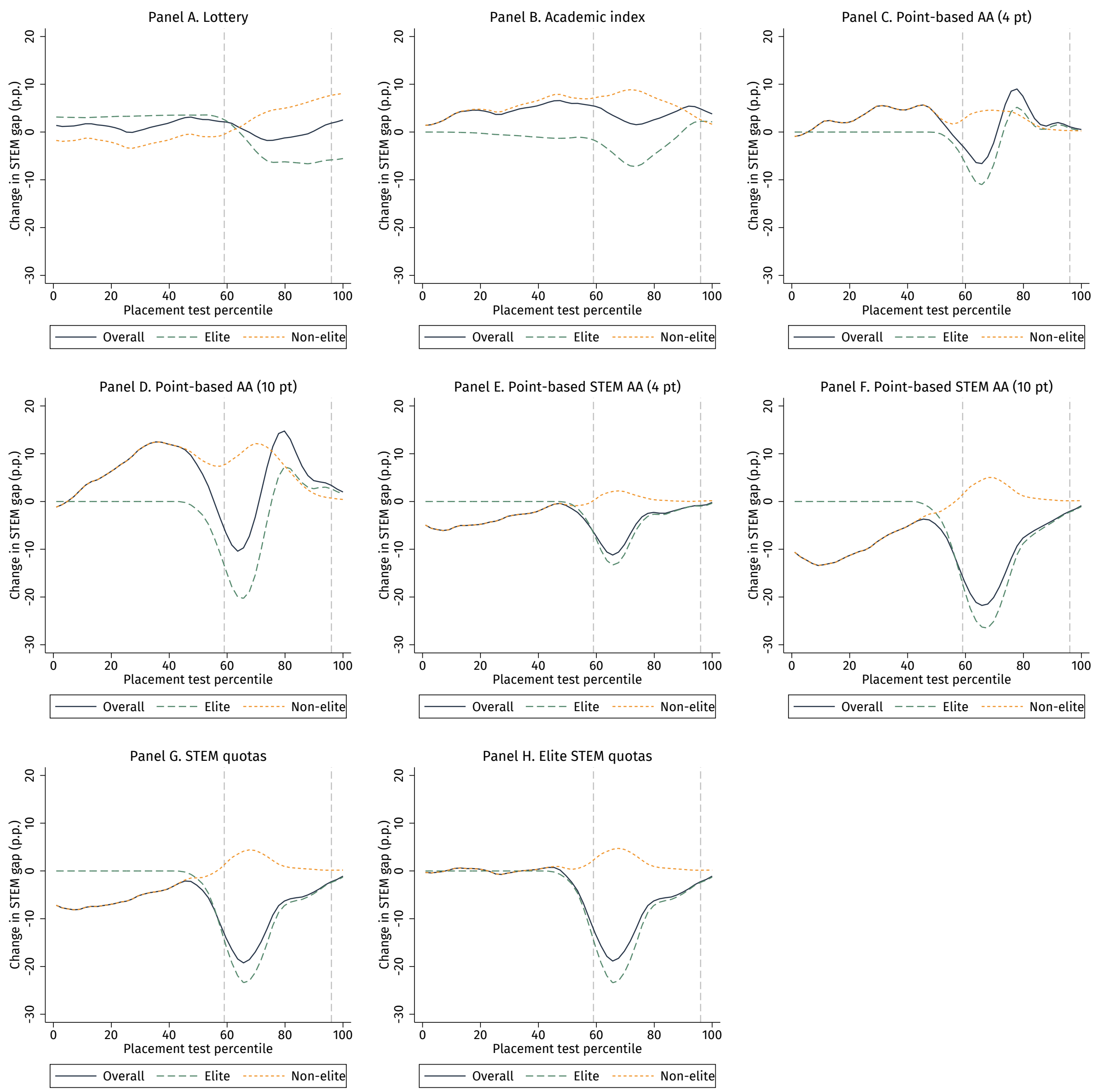
Note: Markers correspond to program-by-year cutoff score pairs, where the x-axis is the true cutoff score and the y-axis is the simulated cutoff score resulting from simulating assignment under the status quo priority structure as described in Section 3.3. Simulated cutoff scores are means over 200 independent simulations of the assignment process accounting for uncertainty in student preference parameters, idiosyncratic student preferences, and random tie-breakers in assignment. Cutoff scores are the lowest placement test score a student could obtain and be assigned, and are set to 31 (the minimum to be eligible for assignment) for programs that are not oversubscribed. Opacity is determined by true enrollment counts in the respective year, such that darker points indicate higher enrollment. Dashed line is a 45-degree line. The raw correlation between true and simulated cutoff scores is 0.98 and enrollment-weighted correlation is 0.99.

Figure 4: Female marginal willingness-to-travel for program characteristics with respect to standardized test subscore



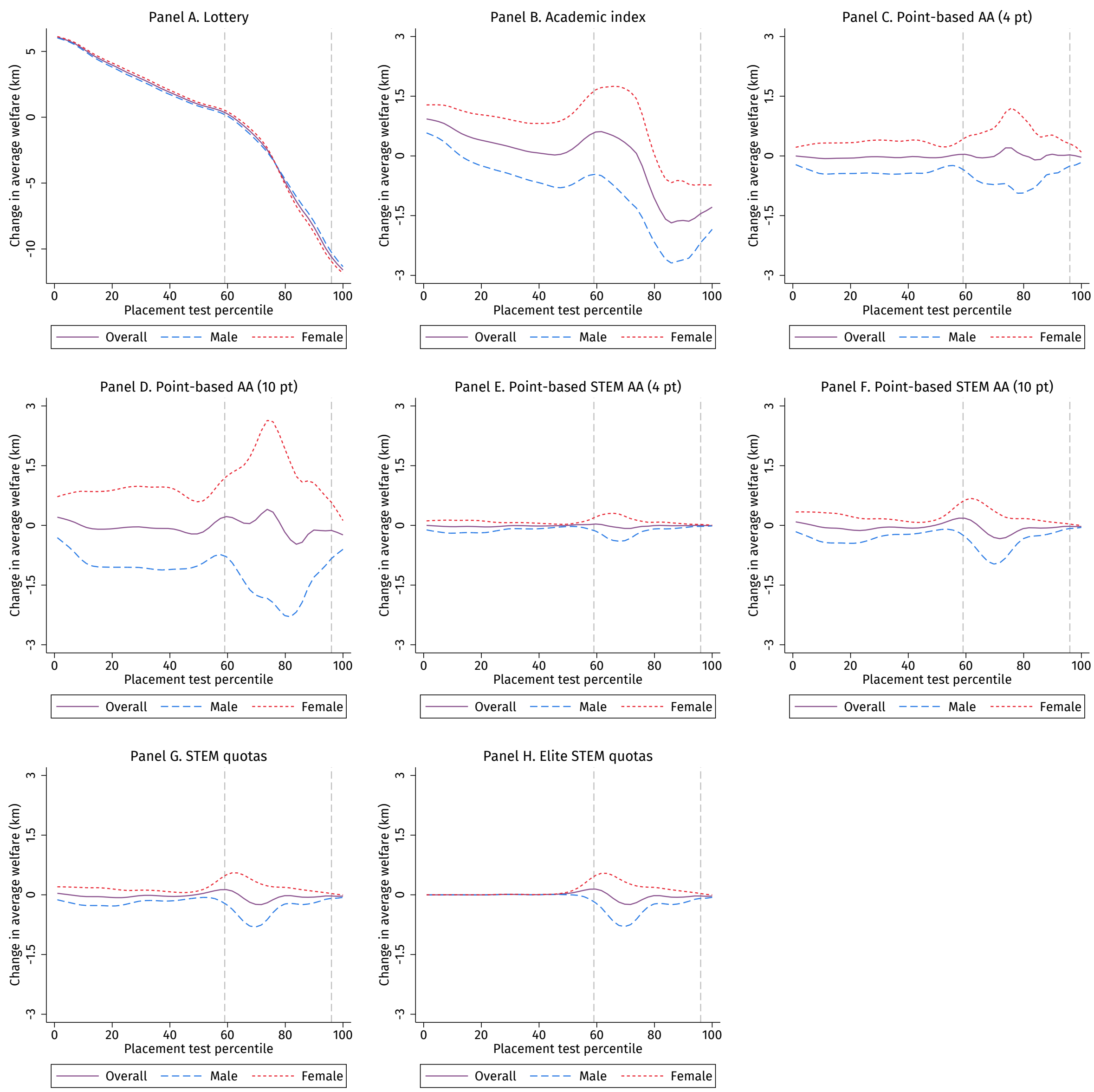
Note: Solid lines are estimated marginal willingness-to-travel (WTT), in kilometers, for the indicated program characteristic, compared to the WTT of a student at the sample mean of the indicated ENLACE 9 exam subscore. These are obtained by transforming the appropriate linear combination of *subscore* × *characteristic* coefficients to a WTT estimate in each region-year estimation cell, then computing the mean of these estimates, weighting by the number of students in the cell. Gray regions are 95% confidence intervals.

Figure 5: Simulated effects of priority structure changes on STEM gap and its components, by placement test percentile



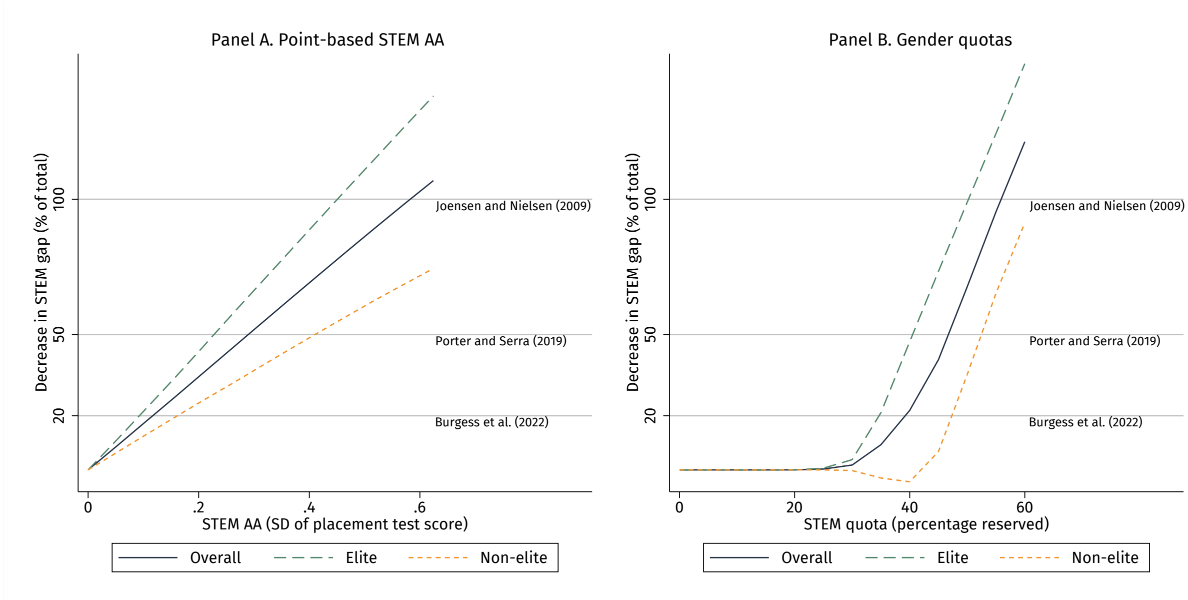
Note: Lines represent percentage point differences between the simulated STEM gaps under the status quo priority structure and the priority structure indicated in the panel title, conditional on the placement test percentile. Simulated changes are means over 200 independent simulations of the assignment process accounting for uncertainty in student preference parameters, idiosyncratic student preferences, and random tie-breakers in assignment. Simulations are as described in Section 3.3. Dashed vertical lines indicate the percentiles corresponding to the lowest and highest elite program cutoff scores.

Figure 6: Simulated welfare effects of priority structure changes, by gender and placement test percentile)



Note: Lines represent, for the indicated subsample, simulated differences in average welfare between the status quo priority structure and the priority structure indicated in the panel title, conditional on the placement test percentile. Simulated changes are means over 200 independent simulations of the assignment process accounting for uncertainty in student preference parameters, idiosyncratic student preferences, and random tie-breakers in assignment. Simulations and welfare computations are as described in Section 3.3. Dashed vertical lines indicate the percentiles corresponding to the lowest and highest elite program cutoff scores. Panel A has a different vertical scale than the other panels.

Figure 7: Simulated impacts of point-based STEM affirmative action and gender quotas on the STEM gap



Note: Lines represent the simulated decrease in the STEM gap and its components against the intensity of the indicated policy, compared to the baseline gap under the status quo priority structure. Vertical axis is the percent of the gap, so that 100% corresponds to fully closing the respective gap. Simulations are as described in Section 3.3. Decreases are means over 20 independent simulations of the assignment process accounting for uncertainty in student preference parameters, idiosyncratic student preferences, and random tie-breakers in assignment. Horizontal axis in Panel A is in standard deviations of placement test score, to facilitate interpretation. Horizontal gray lines are authors' calculations of policy-induced reductions in STEM gaps in the indicated studies, as described in Appendix C.

Online Appendix A: STEM education and the labor market

To estimate the labor market returns to STEM occupations, we use data from the third quarter of the 2010 Encuesta Nacional de Ocupación y Empleo (ENOE) conducted by the Instituto Nacional de Estadística y Geografía (INEGI). The ENOE is a nationally representative survey with information on employment, occupation, monthly income, and hours worked. In this same quarter, an additional module (Encuesta Nacional de Inserción Laboral de los Egresados de la Educación Media Superior, ENILEMS), collected information on recent high school graduates and their transition to higher education and/or the workforce, including their high school concentration, college major, and occupation, when applicable.[31]

We use the same Brookings classifications described in the main text to classify occupations as STEM or non-STEM. Specifically, INEGI provides a crosswalk between the Mexican occupation codes and the U.S. Bureau of Labor Statistics Standard Occupation Classification (SOC) codes, with each Mexican occupation matching one or more SOC codes. We compute the average STEM classification of all matched occupations and categorize the Mexican occupation/education track as STEM if the average STEM classification is 0.5 or more. For high school and college STEM tracks, two individuals separately code each as STEM or non-STEM using the Brookings classification as the guide. Discrepancies are reconciled by a third individual.

Using these data, we calculate the STEM wage premiums for females as well as the transition probabilities between various levels of STEM education and STEM work. We restrict all analyses to the sample of respondents 40 years old or younger, under the rationale that the relevant reference labor market for high school students is that of younger working adults. We calculate hourly wages from monthly income and average hours worked per week (asked of all individuals reporting current employment).

Table A.1 shows the hourly wages for females, by education and STEM classification. The data indicates that there is a 22.0% wage premium for STEM occupations among females with any higher education (25.9% of females). This wage difference is statistically significant and consistent with other work on the returns to STEM occupations. Among females with no higher education, those who work in STEM occupations receive wages that are 10.8% higher than those working in non-STEM occupations. Again, this is statistically significant and similar to the U.S. 10% all gender blue-collar STEM premium identified by Rothwell (2013). This suggests that there are returns to working in STEM occupations, even for individuals who do not proceed past high school.

To understand the relationship between STEM education and STEM work, we examine the

31. Since the larger ENOE module only asks about terminal degrees, we are unable to fully trace individuals' pathways through high school tracks to college majors to occupations. The ENILEMS module allows us to see both high school tracks and college majors for the small subsample; however, since these are recent high school graduates, we do not observe their post-college careers.

correlation between studying STEM in high school or college and transitioning to a STEM field of study or occupation afterward. Table A.2 shows the probability of three separate STEM transitions, where all analyses are for females only. Note, the ENILEMS survey respondents have not completed college and can report both enrolling in college and working. The first column estimates the probability of studying STEM in college for those females who enroll in college. The results indicate that females with STEM concentrations in high school are 7.7 percentage points (p.p) more likely to have a STEM college major than those who do not, a statistically significant difference (at the 10% level) of 44.8% over the base rate. Using the overall ENOE survey, the second column calculates the probability of working in a STEM occupation among females for whom college is their terminal degree. Again, females who study STEM in college are more likely to work in STEM occupations; STEM majors are 43.0 p.p. more likely to work in STEM, a difference (statistically significant) of more than sixfold over the base rate. Finally, using the ENILEMS module, the third column calculates the probability of working in STEM among recent high school graduates who report working. Again, females who study STEM in high school are 9.0 p.p. (98.0%) more likely to work in STEM occupations after high school. This difference is statistically significant at the 10% level.

Table A.1: Female hourly wage by education and STEM classification (pesos)

	(1) Non-STEM occupation	(2) STEM occupation	(3) Difference STEM – non-STEM
Higher education (any college)	42.84	52.25	9.41
	(33.18)	(42.90)	(0.95)
Lower education (no college)	20.45	22.66	2.21
	(23.79)	(15.53)	(0.58)
Observations	26325	3526	29851

Note: Sample is comprised of females aged 40 or younger from the 2010 ENOE. Hourly wages are computed from monthly income and average hours worked. We are unable to classify 8 occupation codes and exclude them from this analysis (columns 2 and 3, 0.5% of observations).

Table A.2: Transitions between STEM education and later STEM-related activities

	(1) STEM college major	(2) STEM occupation	(3) STEM occupation
STEM high school	0.077*		0.090
	(0.0445)		(0.0471)
STEM college major		0.430	
		(0.0148)	
Constant	0.448	0.066	0.092
	(0.0186)	(0.0059)	(0.0157)
Observations	2565	9635	1332
Adjusted R^2	0.003	0.240	0.013

Note: All data are from the 2010 ENOE. The samples for columns (1) and (3) are from the ENILEMS labor force insertion module collected from recent high school graduates. Column (1) includes individuals who transitioned from high school into higher education. Column (3) includes individuals who transitioned from high school into working. Individuals may report both working and participating in higher education. We are unable to classify 8 occupation codes and exclude them from this analysis (columns 2 and 3, 0.8% of observations). We also exclude individuals who list college majors using high school major codes (columns 1 and 3, 2.4% of observations).

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Online Appendix B: Additional tables and figures

Table B.1: STEM school assignment summary statistics for 2005, before and after post- computerized assignment phase

	(1) Male	(2) Female	(3) Difference
Panel A. Computerized assignment			
STEM assigned program (elite or non-elite)	39.3	27.3	12.0***
	(48.8)	(44.5)	(0.2)
Elite STEM assigned program	10.1	4.0	6.1***
	(30.1)	(19.6)	(0.1)
Non-elite STEM assigned program	29.2	23.3	5.9***
	(45.5)	(42.3)	(0.2)
Elite non-STEM assigned program	18.1	19.5	-1.4***
	(38.5)	(39.6)	(0.2)
Technical non-STEM assigned program	12.4	16.3	-3.9***
	(32.9)	(36.9)	(0.2)
Observations	102869	99714	202583
Panel B. Finalized assignment			
STEM assigned program (elite or non-elite)	39.6	27.9	11.6***
	(48.9)	(44.9)	(0.2)
Elite STEM assigned program	9.2	3.5	5.7***
	(29.0)	(18.5)	(0.1)
Non-elite STEM assigned program	30.3	24.4	6.0***
	(46.0)	(42.9)	(0.2)
Elite non-STEM assigned program	16.6	17.3	-0.7***
	(37.2)	(37.8)	(0.2)
Technical non-STEM assigned program	12.7	17.5	-4.8***
	(33.3)	(38.0)	(0.2)
Observations	112069	112747	224816

Note: Calculations in Panel A are for all students who were assigned to a program by the placement algorithm in the 2005 COMIPEMS cycle who resided within the COMIPEMS geographical boundary. Calculations in Panel B include students assigned either by the placement algorithm or during the post-assignment program selection process, in which unassigned students were able to choose a program that had not filled its quota (or remain unassigned). Indicator variables are percentages. Standard deviations are in parentheses in columns (1) and (2); standard errors are in parentheses in column (3).

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B.2: Education track STEM mapping

Program code	STEM classification	Program code	STEM classification
001 Administración	0	072 Mantenimiento automotriz	0
002 Refrigeración y aire acondicionado	1	073 Producción industrial	0
003 Análisis y tecnología de alimentos	1	074 Sistemas de impresión offset y serigrafía	0
006 Computación	1	075 Telecomunicaciones	1
007 Computación fiscal contable	1	076 Técnico en mecatrónica	1
008 Comunicación	0	077 Técnico en manufactura asistida por computadora	1
009 Construcción	0	078 Técnico en alimentos instituciones educativas	0
010 Contabilidad	1	203 Agencia de viajes	0
011 Dietética	0	208 Artes gráficas	0
012 Arquitectura	1	214 Contabilidad	0
013 Diseño gráfico	0	218 Cosmetología esteticista	0
014 Diseño de modas	0	220 Dibujo publicitario	0
015 Electricidad	1	222 Diseño arquitectónico	1
016 Electrónica	1	223 Diseño decorativo	0
017 Enfermería general	1	224 Diseño gráfico	0
018 Gericultura	0	225 Diseño industrial	1
019 Informática administrativa	0	226 Diseño industrial de patrones	1
020 Laboratorista clínico	1	227 Ediciones	0
021 Laboratorista químico	1	229 Electricidad industrial	1
022 Mantenimiento	0	237 Fotomecánica	0
023 Mantenimiento de equipo de computo	1	238 Gerencia y supervisión en la industria del vestido	0
024 Máquinas de combustión interna	0	246 Mecánica automotriz	0
025 Máquinas-herramienta	1	247 Mecánica industrial	1
026 Mecánica industrial	1	250 Modelismo y fundición	0
027 Producción	0	252 Paquetes de cómputo	1
028 Programador	1	260 Radiología e imagen	1
029 Prótesis dental	1	264 Sastrería industrial	0
030 Puericultura	0	265 Secretario bilingue	0
031 Secretario ejecutivo	0	266 Secretario ejecutivo	0
032 Supervisor en la industria del vestido	0	267 Servicio a equipo de cómputo	1
033 Técnico en agroindustrias	1	275 Telecomunicaciones	1
034 Técnico agropecuario	1	277 Trabajo social	0
035 Técnico en instrumentación dental	1	278 Secretario ejecutivo bilingue	0
036 Técnico en administración	0	301 Administración	0
037 Técnico en computacion fiscal contable	1	302 Alimentos y bebidas	0
038 Técnico en edificación	1	303 Asistente directivo	0
039 Técnico en contabilidad	1	304 Automotriz	0
040 Técnico en diseño industrial	1	305 Construcción	0
041 Técnico en diseño gráfico	0	306 Contaduría	1
042 Técnico en electricidad	1	307 Control de calidad	1
043 Técnico en electronica	1	308 Conservación del medio ambiente	1
044 Técnico en enfermería general	1	309 Dental	1
045 Técnico en industrializacion de lacteos	1	310 Electricidad industrial	1
046 Técnico en informática	1	311 Electromecánica	1
047 Técnico en informática agropecuaria	1	312 Electrónica industrial	1
048 Técnico en mantenimiento en equipo de computo	1	313 Enfermería general	1
049 Técnico en mantenimiento industrial	1	314 Hospitalidad turística	0
050 Técnico en maquinas-herramienta	1	315 Industria del vestido	0
052 Técnico laboratorista clinico	1	316 Informática	1
053 Técnico laboratorista químico-clínico	1	317 Mantenimiento de equipo de cómputo y control digital	1
054 Técnico en manufactura en la industria del vestido	0	318 Mantenimiento de motores y planeadores	1
055 Trabajo social	0	319 Mantenimiento de sistemas automáticos	1
056 Turismo	0	320 Máquinas herramienta	1
057 Técnico programador	1	321 Metalmecánica	0
058 Diseño decorativo	0	322 Optometría	1
059 Diseño industrial	1	323 Plásticos	0
060 Mecatrónica	1	324 Procesamiento industrial de alimentos	0
061 Técnico en horticultura	1	325 Producción y transformación de productos acuícolas	0
062 Técnico en sistemas electricos de control y automatizacion	1	326 Productividad industrial	0
063 Técnico asistente ejecutivo	0	327 Química industrial	1
064 Diseño y proyecto gráfico	0	328 Refrigeración y aire acondicionado	1
065 Asistente ejecutivo bilingüe	0	329 Sistemas electrónicos de aviación	1
066 Técnico en diseño asistido por computadora	1	330 Telecomunicaciones	1
067 Mantenimiento de equipo y sistemas	0	331 Terapia respiratoria	1
068 Informática	1	332 Laministería y recubrimiento de las aeronaves	0
069 Técnico en turismo	0	333 Seguridad e higiene y Protección civil	1
070 Técnico en gastronomía	0	334 Expresión gráfica digital	0
071 Técnico en mercadotecnia	0	335 Mecatrónica	1
		336 Autotrónica	1

Note: The guidelines for STEM classification come from Rothwell (2013), which identifies U.S. STEM occupations based on level of STEM knowledge required.

Table B.3: Relationship between placement test score and ENLACE 9 subscores

	(1)	(2)	(3)
ENLACE 9 math subscore (normalized)	9.63***		9.05***
	(0.026)		(0.026)
ENLACE 9 Spanish subscore (normalized)	8.55***		9.28***
	(0.030)		(0.030)
Male		3.98***	4.63***
		(0.065)	(0.039)
2008 cohort	1.32***	1.26***	1.40***
	(0.038)	(0.064)	(0.038)
Constant	56.60***	63.99***	54.29***
	(0.030)	(0.054)	(0.034)
Observations	371739	371739	371739
Adjusted R^2	0.647	0.011	0.660
Mean of dependent variable	66.46	66.46	66.46
SD of dependent variable	19.74	19.74	19.74

Note: Dependent variable is raw placement test score. Sample is the subset of the analysis sample from Table 1 that has ENLACE 9 scores available. Huber-White robust standard errors are in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01.

Table B.4: Average conditional logit estimates under alternative assumptions

	(1) Stability	(2) Truth-telling
Distance	-0.257***	-0.213***
	(0.0267)	(0.0179)
Distance × high parental education	0.009	0.009
	(0.0304)	(0.0208)
Distance × high parental education × male	0.003	0.002
	(0.0360)	(0.0247)
Distance × missing parental education	0.008	0.006
	(0.0361)	(0.0251)
Distance × missing parental education × male	0.000	0.001
	(0.0424)	(0.0297)
Distance × missing test score	0.012	0.019
	(0.0336)	(0.0235)
Distance × missing test score × male	0.001	-0.002
	(0.0394)	(0.0275)
Distance × middle school GPA	0.001	-0.002
	(0.0247)	(0.0170)
Distance × middle school GPA × male	0.002	0.001
	(0.0293)	(0.0202)
Distance × male	0.007	0.005
	(0.0300)	(0.0205)
Distance × math	0.009	0.004
	(0.0321)	(0.0214)
Distance × math × male	-0.003	0.001
	(0.0380)	(0.0254)
Distance × math^2	-0.003	0.001
	(0.0227)	(0.0150)
Distance × math^2× male	0.001	0.000
	(0.0272)	(0.0182)
Distance × math^3	-0.001	0.000
	(0.0183)	(0.0117)
Distance × math^3× male	0.000	-0.000
	(0.0212)	(0.0137)

	(1) Stability	(2) Truth-telling
Distance × Spanish	0.014	0.003
	(0.0330)	(0.0225)
Distance × Spanish × male	-0.007	0.000
	(0.0396)	(0.0269)
Distance × Spanish^2	-0.004	0.001
	(0.0245)	(0.0159)
Distance × Spanish^2× male	0.001	-0.000
	(0.0295)	(0.0192)
Distance × Spanish^3	-0.003	0.000
	(0.0197)	(0.0130)
Distance × Spanish^3× male	0.002	0.000
	(0.0229)	(0.0153)
Elite non-STEM × high parental education	0.441***	0.531***
	(0.1685)	(0.0889)
Elite non-STEM × high parental education × male	0.029	-0.018
	(0.1992)	(0.1083)
Elite non-STEM × missing parental education	0.231	0.316***
	(0.2183)	(0.1098)
Elite non-STEM × missing parental education × male	0.070	0.018
	(0.2540)	(0.1329)
Elite non-STEM × missing test score	-0.003	-0.182*
	(0.2141)	(0.1030)
Elite non-STEM × missing test score × male	-0.312	-0.310**
	(0.2432)	(0.1232)
Elite non-STEM × middle school GPA	0.363***	0.417***
	(0.1380)	(0.0733)
Elite non-STEM × middle school GPA × male	-0.009	0.140
	(0.1610)	(0.0891)
Elite non-STEM × male	-0.157	0.143
	(0.1839)	(0.0886)
Elite non-STEM × math	0.072	0.147
	(0.1939)	(0.0910)
Elite non-STEM × math × male	-0.020	-0.028
	(0.2268)	(0.1091)
Elite non-STEM × math^2	-0.011	-0.006
	(0.1844)	(0.0649)

	(1) Stability	(2) Truth-telling
Elite non-STEM × math^2× male	0.055	-0.006
	(0.2142)	(0.0799)
Elite non-STEM × math^3	0.019	-0.005
	(0.1126)	(0.0509)
Elite non-STEM × math^3× male	-0.013	-0.003
	(0.1312)	(0.0601)
Elite non-STEM × Spanish	-0.158	0.205**
	(0.2118)	(0.0955)
Elite non-STEM × Spanish × male	0.155	-0.046
	(0.2373)	(0.1159)
Elite non-STEM × Spanish^2	0.131	0.010
	(0.2066)	(0.0677)
Elite non-STEM × Spanish^2× male	-0.018	-0.010
	(0.2232)	(0.0832)
Elite non-STEM × Spanish^3	0.006	-0.014
	(0.1389)	(0.0561)
Elite non-STEM × Spanish^3× male	0.019	0.009
	(0.1571)	(0.0670)
Elite STEM × high parental education	0.232	0.326***
	(0.1770)	(0.1079)
Elite STEM × high parental education × male	0.153	0.101
	(0.2008)	(0.1273)
Elite STEM × missing parental education	0.108	0.214
	(0.2287)	(0.1330)
Elite STEM × missing parental education × male	0.062	0.018
	(0.2569)	(0.1565)
Elite STEM × missing test score	0.001	-0.140
	(0.2181)	(0.1253)
Elite STEM × missing test score × male	-0.565**	-0.536***
	(0.2418)	(0.1460)
Elite STEM × middle school GPA	0.291**	0.368***
	(0.1458)	(0.0885)
Elite STEM × middle school GPA × male	0.006	0.116
	(0.1641)	(0.1034)
Elite STEM × male	0.975***	0.888***
	(0.1802)	(0.1029)

	(1) Stability	(2) Truth-telling
Elite STEM × math	0.175	0.247**
	(0.1943)	(0.1082)
Elite STEM × math × male	0.055	0.068
	(0.2189)	(0.1265)
Elite STEM × math^2	0.142	0.040
	(0.1824)	(0.0775)
Elite STEM × math^2× male	-0.026	-0.010
	(0.2037)	(0.0929)
Elite STEM × math^3	-0.002	-0.003
	(0.1188)	(0.0590)
Elite STEM × math^3× male	0.005	-0.007
	(0.1324)	(0.0686)
Elite STEM × Spanish	-0.195	0.064
	(0.2139)	(0.1154)
Elite STEM × Spanish × male	0.036	-0.024
	(0.2347)	(0.1355)
Elite STEM × Spanish^2	0.020	0.002
	(0.2095)	(0.0806)
Elite STEM × Spanish^2× male	0.105	-0.024
	(0.2215)	(0.0954)
Elite STEM × Spanish^3	0.058	-0.001
	(0.1434)	(0.0665)
Elite STEM × Spanish^3× male	-0.033	-0.004
	(0.1579)	(0.0770)
Non-elite STEM × high parental education	-0.278**	-0.264***
	(0.1188)	(0.0896)
Non-elite STEM × high parental education × male	0.125	0.089
	(0.1398)	(0.1058)
Non-elite STEM × missing parental education	-0.035	-0.057
	(0.1406)	(0.1070)
Non-elite STEM × missing parental education × male	0.031	0.009
	(0.1646)	(0.1258)
Non-elite STEM × missing test score	0.083	0.015
	(0.1330)	(0.1001)
Non-elite STEM × missing test score × male	-0.268*	-0.205*
	(0.1552)	(0.1165)

	(1) Stability	(2) Truth-telling
Non-elite STEM × middle school GPA	-0.059	-0.076
	(0.0960)	(0.0731)
Non-elite STEM × middle school GPA × male	-0.018	-0.019
	(0.1146)	(0.0876)
Non-elite STEM × male	0.518***	0.459***
	(0.1178)	(0.0876)
Non-elite STEM × math	0.049	0.027
	(0.1278)	(0.0932)
Non-elite STEM × math × male	0.057	0.021
	(0.1516)	(0.1092)
Non-elite STEM × math^2	0.023	0.019
	(0.0980)	(0.0641)
Non-elite STEM × math^2× male	0.005	-0.007
	(0.1153)	(0.0762)
Non-elite STEM × math^3	0.013	0.006
	(0.0818)	(0.0523)
Non-elite STEM × math^3× male	-0.021	-0.005
	(0.0962)	(0.0605)
Non-elite STEM × Spanish	-0.057	-0.108
	(0.1286)	(0.0957)
Non-elite STEM × Spanish × male	0.052	0.028
	(0.1537)	(0.1135)
Non-elite STEM × Spanish^2	0.022	-0.001
	(0.1135)	(0.0703)
Non-elite STEM × Spanish^2× male	0.021	-0.001
	(0.1416)	(0.0857)
Non-elite STEM × Spanish^3	-0.001	0.002
	(0.0872)	(0.0569)
Non-elite STEM × Spanish^3× male	0.010	0.005
	(0.1035)	(0.0667)
Constant × high parental education	-0.648***	-0.256***
	(0.1303)	(0.0821)
Constant × high parental education	-0.032	0.006
	(0.1607)	(0.0992)
Constant × missing parental education	-0.409***	-0.181*
	(0.1568)	(0.0994)

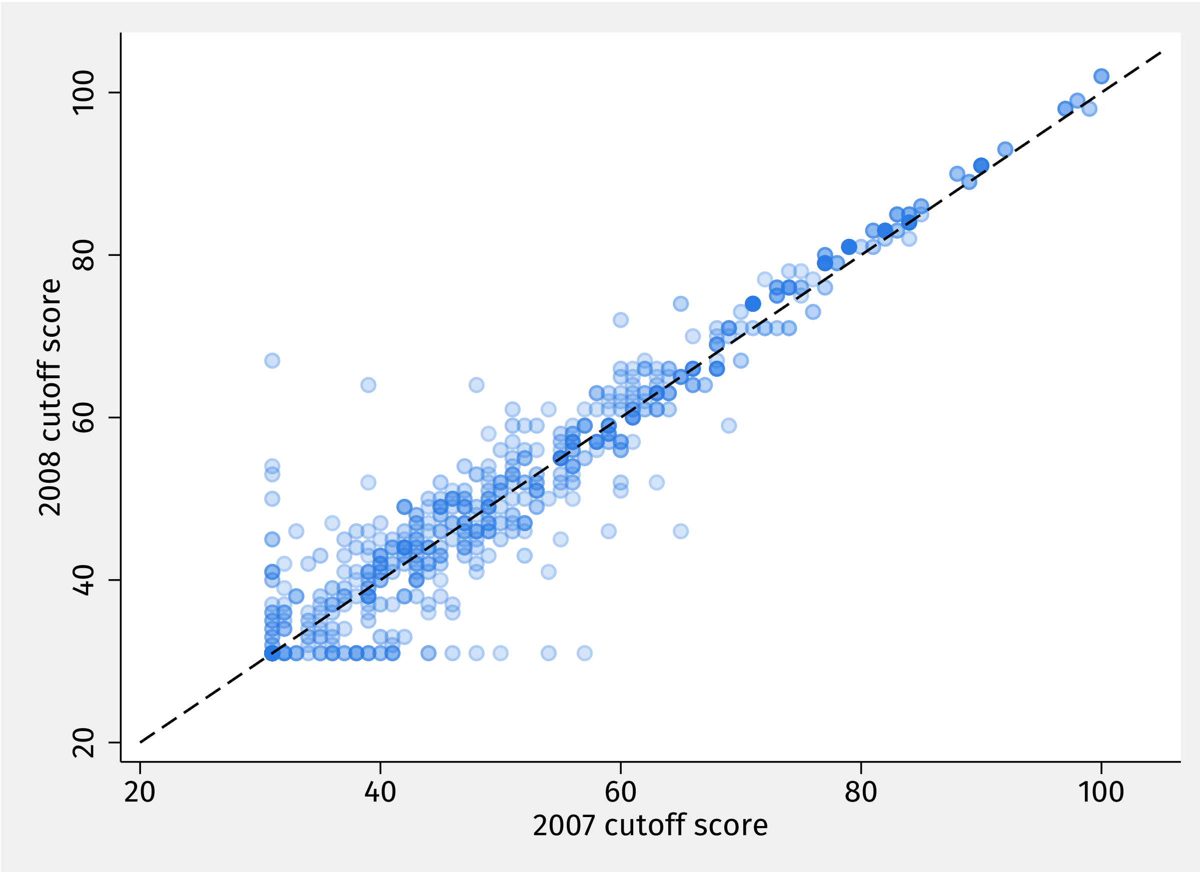
	(1) Stability	(2) Truth-telling
Constant × missing parental education × male	0.008	0.002
	(0.1911)	(0.1187)
Constant × missing test score	-0.381**	-0.326***
	(0.1493)	(0.0927)
Constant × missing test score × male	0.217	0.157
	(0.1815)	(0.1096)
Constant × middle school GPA	-0.163	-0.024
	(0.1064)	(0.0671)
Constant × middle school GPA × male	-0.200	-0.068
	(0.1314)	(0.0820)
Constant × male	-0.108	-0.253***
	(0.1367)	(0.0825)
Constant × math	-0.062	-0.100
	(0.1434)	(0.0858)
Constant × math × male	-0.049	-0.031
	(0.1793)	(0.1035)
Constant × math^2	0.049	-0.029
	(0.1083)	(0.0592)
Constant × math^2× male	-0.004	-0.008
	(0.1332)	(0.0726)
Constant × math^3	0.009	-0.004
	(0.0926)	(0.0484)
Constant × math^3× male	0.024	0.007
	(0.1151)	(0.0576)
Constant × Spanish	-0.131	-0.092
	(0.1422)	(0.0884)
Constant × Spanish × male	0.045	0.003
	(0.1788)	(0.1074)
Constant × Spanish^2	0.042	-0.053
	(0.1238)	(0.0645)
Constant × Spanish^2× male	0.010	0.012
	(0.1656)	(0.0805)
Constant × Spanish^3	0.035	-0.007
	(0.0950)	(0.0529)
Constant × Spanish^3× male	-0.010	-0.001
	(0.1176)	(0.0631)

	(1) Stability	(2) Truth-telling
Technical non-STEM × high parental education	-0.197	-0.197**
	(0.1245)	(0.0920)
Technical non-STEM × high parental education × male	-0.054	-0.064
	(0.1521)	(0.1114)
Technical non-STEM × missing parental education	0.018	-0.024
	(0.1475)	(0.1098)
Technical non-STEM × missing parental education × male	-0.039	-0.007
	(0.1777)	(0.1315)
Technical non-STEM × missing test score	-0.005	-0.060
	(0.1398)	(0.1029)
Technical non-STEM × missing test score × male	-0.254	-0.152
	(0.1686)	(0.1224)
Technical non-STEM × middle school GPA	-0.110	-0.123
	(0.1008)	(0.0751)
Technical non-STEM × middle school GPA × male	0.035	0.007
	(0.1246)	(0.0927)
Technical non-STEM × male	0.057	-0.051
	(0.1286)	(0.0931)
Technical non-STEM × math	0.035	-0.011
	(0.1342)	(0.0957)
Technical non-STEM × math × male	0.018	-0.024
	(0.1648)	(0.1155)
Technical non-STEM × math^2	0.013	0.003
	(0.1006)	(0.0662)
Technical non-STEM × math^2× male	0.004	-0.007
	(0.1234)	(0.0806)
Technical non-STEM × math^3	0.007	0.003
	(0.0854)	(0.0543)
Technical non-STEM × math^3× male	-0.017	-0.006
	(0.1040)	(0.0646)
Technical non-STEM × Spanish	0.055	-0.007
	(0.1341)	(0.0980)
Technical non-STEM × Spanish × male	0.002	-0.057
	(0.1658)	(0.1191)
Technical non-STEM × Spanish^2	0.047	-0.002
	(0.1134)	(0.0725)

	(1) Stability	(2) Truth-telling
Technical non-STEM × Spanish^2× male	-0.009	-0.014
	(0.1516)	(0.0918)
Technical non-STEM × Spanish^3	0.004	-0.001
	(0.0892)	(0.0588)
Technical non-STEM × Spanish^3× male	-0.007	0.003
	(0.1095)	(0.0703)
Observations	494624	494624
Hausman test statistic $\chi^2(4226)$	169049.54	
Hausman test p-value	0.0000	

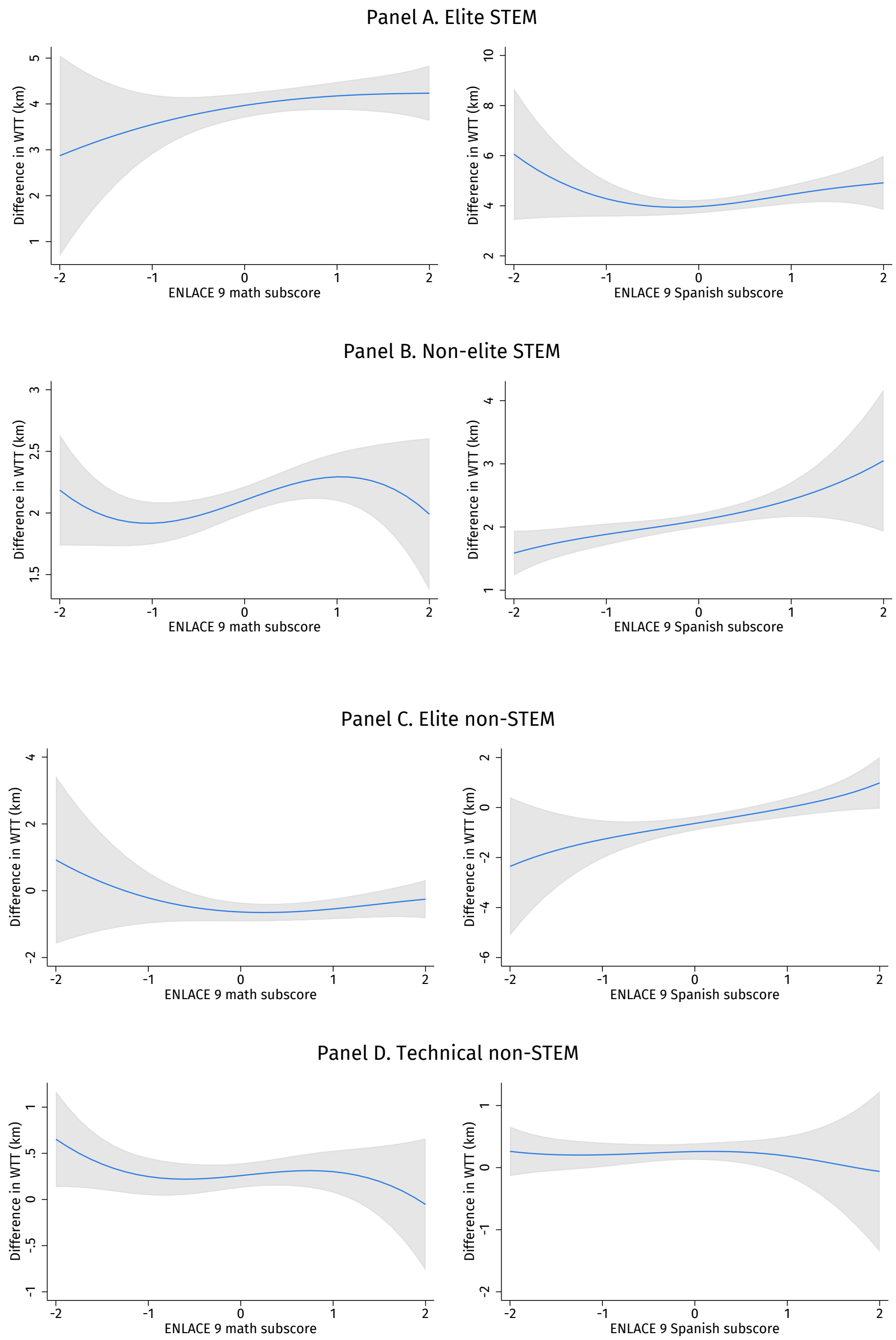
Note: Columns refer to assumptions under which conditional logit models were estimated. Column 1 assumes stability of the matching, such that the conditional logit estimates preferences using student assignments given personalized choice sets as described in Section 3.1. Column 2 assumes truth-telling, such that the rank-ordered logit estimates preferences using the full sequence of choices given the full choice set. Student count-weighted averages of region-year estimation cell coefficients are reported. Standard errors of these weighted averages are in parentheses. Sample is 2007 and 2008 COMIPEMS cycles. Hausman test follows Fack, Grenet, and He (2019) and tests for consistency of the truth-telling model when the stability model is consistent (but potentially inefficient). * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Figure B.1: Comparison of program-level cutoff scores in 2007 and 2008



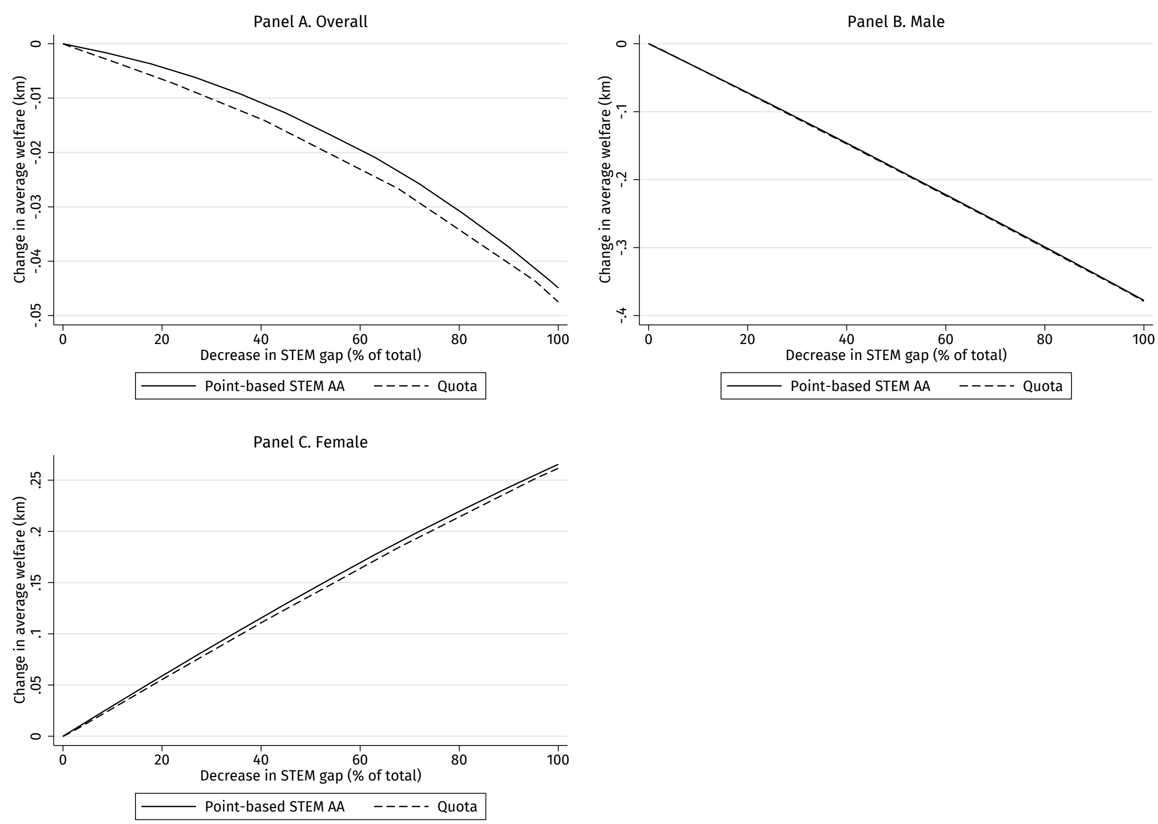
Note: Markers correspond to program-level cutoff scores in 2007 and 2008. Cutoff scores are the lowest placement test score a student could obtain and be assigned, and are set to 31 (the minimum to be eligible for assignment) for programs that are not oversubscribed in that year. Opacity is determined by 2007 enrollment counts, such that darker points indicate higher enrollment. Dashed line is a 45-degree line. The raw correlation is 0.95 and enrollment-weighted correlation is 0.98.

Figure B.2: Male-female difference in willingness-to-travel for program characteristics, by standardized test subscore)



Note: Solid lines are estimated gender differences in marginal willingness-to-travel (WTT), in kilometers, for the indicated program characteristic. These are obtained by transforming the appropriate linear combination of *male* × *characteristic* and *male* × *subscore* × *characteristic* coefficients to a WTT estimate in each region-year estimation cell, then computing the mean of these estimates, weighting by the number of students in the cell. Gray regions are 95% confidence intervals.

Figure B.3: Comparison of point-based STEM affirmative action and gender quota effects on STEM gap and welfare



Note: Lines plot the simulated change in welfare corresponding to the simulated reduction in the STEM gap, obtained by simulating varying levels of STEM AA or proportions of total seats reserved in STEM programs under female quotas. Simulations and welfare computations are as described in Section 3.3. Changes are means over 20 independent simulations of the assignment process accounting for uncertainty in student preference parameters, idiosyncratic student preferences, and random tie-breakers in assignment. Dashed vertical lines indicate the percentiles corresponding to the lowest and highest elite program cutoff scores.

Online Appendix C: STEM gap reductions from demand-side interventions

In the Scaling Stem AA section, we contextualize various levels of STEM AA using the results of several demand-side interventions that have reduced the STEM gender gap. We provide details on the specific outcomes and calculations of the gap reductions here.

Burgess, et al., (2022) examine the gender gap in the likelihood of graduating with a STEM degree in post-secondary education. They show that random assignment to a semi-external assessment in math (SEAM) reduces the gap in a math-requiring degree by 1.2 percentage points relative to the baseline gap of 5.5 percentage points. They show a larger decline in the gap for math-demanding STEM degrees, with random assignment to SEAM decreasing the baseline gap from 1.1 percentage points to 0.4 percentage points. See the introduction of the paper for details.

In a field experiment, Porter and Serra (2019) study the outcome of women majoring in economics. Looking specifically at women in introductory economics classes, they show that short exposures to successful women in economics substantially increase the likelihood that women major in economics. Using their reported outcomes for treatment classes, we calculate the baseline gender gap to be approximately 21 percentage points, with 8% of females majoring in economics (Table II) compared to 29% of males (Figure V). Following the intervention, 15% of females majored in economics (Table II) compared to 26% of males (Figure V), reducing the gap by approximately 10 percentage points or 48%.

Bottia, et al., (2015) examine students in North Carolina and study the effects of female math and science high school teachers on majoring in STEM in college. Across the college system, they report a 14 percentage point baseline gender gap in declaring a STEM major (Data section). From their model, they report that students who move from attending a school that is one standard deviation below the mean in their proportion of math and science teachers who are female to a school that is one standard deviation above the mean increases the chances of declaring a PSEM major (physical sciences, engineering, and mathematics) by 14%, where the results are only significant within female samples (Section 4.2 Long-term effects). Assuming a 14% increase for females from a baseline of 43% (Section 3.1 Data) to 49% and holding males constant at their baseline of 57% suggests a new STEM gap of 8 percentage points. This corresponds to a gap reduction of 6 percentage points, or 45% of the baseline gap.

Canaan and Mouganie (2020) study the impact of randomly assigned faculty advisors on student major choice in the American University of Beirut. They report a baseline STEM gender gap of 16.6 percentage points, with 9.3% of females and 25.9% of males enrolling in a STEM major, among students who declare their major after their first year. They find that exposure to a female science advisor reduces the STEM gender gap by approximately 8.6 percentage points, due to a 5.4 percentage point increase in females majoring in STEM and a statistically insignificant 3.2 percentage point reduction for males. This corresponds to a 52% decrease in the baseline STEM gender gap.

Joensen and Nielsen (2016) study a curriculum reform in Denmark that expanded the range of courses paired with advanced maths in ninth grade. We calculate a baseline gender gap in STEM majors of 12.7 percentage points, with 18.1% of females and 30.8% of males reporting a field of major within health science, natural science, technical science, or social science (adding the fields in Appendix Table A.3). Totalling the effects from Table 7 (instrumental variables estimates), we calculate that taking advanced maths in high school results in an 25.1 percentage point increase in female science majors (statistically significant for health sciences and technical sciences) and a 12.1 percentage point increase in male science majors (statistically insignificant for each category). This effectively closes the gap in science majors fully.

Carrell, Page, and West (2010) study the effects of random assignment to professors in the U.S. Air Force Academy, where 25% of women and 41% of men major in STEM. They find that high-performing female students are more likely to take future math and science courses and to graduate with a STEM degree when they are assigned to a female professor in mandatory introductory math and science courses. From their regression results focused on the top quartile of female students (approximated by SAT math scores), they estimate that “increasing the fraction of female professors from 0% to 100% would completely eliminate the gender gap in math and science majors."