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

October 2021

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 exam-based assignment also restricts access to selective STEM schools. Simulations show that changes to the assignment priority structure and preference-altering interventions have the potential to decrease the STEM gap, with differential effects by student achievement. Targeted affirmative action policies would eliminate the STEM gap with small impacts on welfare.

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, Universidad del Rosario, University of Hawaii, and Texas Tech University for helpful feedback. Aleeza Toribio provided excellent research assistance. Declarations of interest: none.

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. Researchers are actively trying to identify factors that contribute to the gap and to examine policies that decrease it. Important open questions include how much of the gap is determined by females selecting into STEM programs of study less often, how much is a result of differential access to STEM education, and how much is generated by more complex interactions between these two factors.[1] Differential access may be a particular concern in contexts with competitive school admissions, which often arise in school choice systems with test-based admissions criteria.[2] Understanding the relative roles of choice and access in generating the STEM gap is crucial for formulating policy responses that ameliorate it.

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 from 2004 through 2009, 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 exam. We use a ranked conditional logit model that allows us to estimate willingness to travel for different program types by gender, geographic location, and other covariates including academic achievement, STEM-specific achievement, and parental education. We 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, general and STEM-specific affirmative action, and general equilibrium responses in which students' preferences respond to priority structure changes. We also simulate choice nudges by increasing female willingness to travel for STEM programs.

Several key findings emerge from the model. First, preferences for STEM programs are highly gendered, with females expressing a lower willingness to travel to STEM programs after controlling for a variety of other factors. These gendered preferences contribute sub-

[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).

stantially to the STEM gender gap, with the overall gap disappearing under simulations where observably similar males and females have the same preferences. Second, access constraints are important for assignment to the highly-demanded STEM programs belonging to historically elite public institutions. Approximately 36% of the gender gap in elite STEM programs is due to test-based admissions policies, with females scoring an average 0.2 standard deviations lower on the placement test. Together, the model highlights a nuanced interaction between preferences and access across the achievement distribution and by school competitiveness. The results suggest that the effects of any proposed policies may be complex and heterogeneous.

The simulations confirm this, showing that policy changes can reduce the STEM gender gap, but with differential effects on various groups and program types. Specifically, policies that increase female admission priorities for STEM programs (e.g., STEM affirmative action) or policies that increase female demand for STEM programs (e.g., choice nudges) can reduce the gap substantially. However, these policies have heterogeneous effects, with a greater gap reduction for elite STEM programs and among higher-scoring females using STEM affirmative action but a greater gap reduction for non-elite STEM programs and lower-scoring females using choice nudges.

On the other hand, policies that increase female admission priorities to all programs (e.g., using lotteries or prior grades) actually increase the overall size of the gap. This is consistent with females listing more non-STEM programs and being given higher chances of admission into those schools. Again, the overall effects obscure more heterogeneous impacts—the overall increase in the STEM gender gap comes from a reduction in the elite STEM gap but a larger countervailing increase in the non-elite STEM gap.

Finally, when females' preferences respond to priority structure changes, we find larger reductions in the elite STEM gap but smaller reductions (or larger increases) in the non-elite STEM gap. Together, these results again suggest that a comprehensive understanding of the interplay between gendered preferences and access across student and program types is fundamental for reducing the STEM gender gap.

With respect to welfare, all of the counterfactual priority structures decrease average welfare for males and increase it for females. However, STEM affirmative action closes the STEM gap with minimal impact on aggregate student welfare and comparatively small gender-specific effects. Applying STEM affirmative action to the extent that it closes the gap entirely reduces average welfare equivalent to 0.03 kilometers (km) of one-way commuting distance, compared to an average commute of about 7 km. Males lose welfare equivalent to 0.5 km, while females gain by 0.4 km. These changes are small compared to those generated by general affirmative action and other policies that favor females in both STEM and non-STEM schools.

Our work leverages Mexico City's unique institutional setting to examine two important aspects of the STEM gender gap that have received limited attention. First, the gender gap begins early on in education, but much prior work has focused on the gap at the university level (Card and Payne 2021; Lichtenberger and George-Jackson 2013). Here, our data allow us to examine STEM high school programs, an earlier component of the STEM pipeline. Second, there is important heterogeneity in STEM, with large differences between more academically-focused and more technical programs. These more technical programs are less studied, but they are important labor market entry points for many economies (Rothwell 2013). In Mexico City, the well-defined program categories allow us to separately estimate preferences for the highly-demanded academic STEM programs and preferences for non-elite technical STEM programs.

Our finding of highly gendered preferences emphasizes the need for research that explores why females choose STEM programs less than males. Existing studies highlight the importance of environmental factors including culture, role models, family expectations, and chilly climates (Kahn and Ginther 2017). Others find gendered preferences for different types of coursework (Zafar 2013), with females preferring work that is perceived as more pro-social and cooperative (Y. Shi 2018).

These studies suggest that choices are highly mutable, and our choice nudge simulations demonstrate the large potential effects of changing preferences. This finding aligns with the effectiveness of various demand-side interventions aimed at changing choices. Interventions providing information about school quality (J. S. 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 (Breda et al. 2018; Carrell, Page, and West 2010; Lim and Meer 2019; Porter and Serra 2019), female peers (Schneeweis and Zweimüller 2012), and positive feedback with respect to course grades (Owen 2010).

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).

Here, 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 (Colleges 2018), lotteries (Shapiro 2020), and various combinations of the above (Karp 2021). Furthermore, there are ongoing debates around affirmative action (Bertrand, Hanna, and Mullainathan 2010; Francis and Tannuri-Pianto 2012; Frisancho and Krishna 2016; Fryer Jr and Loury 2005) 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 results suggest that closing the STEM gap may require targeted changes to the priority structure (e.g., STEM affirmative action) and that such policies may come at little welfare cost.

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). While our approach follows Abdulkadiroğlu et al. (2020) in allowing for rich preference heterogeneity with respect to observable student characteristics, we confront practical issues related to choice portfolio construction in an exam-based mechanism where scores are unknown at the time of portfolio submission. We thus differ significantly from the approach taken by Fack, Grenet, and He (2019), who give an empirical approach when students know their scores and the approximate cutoff scores of programs in their choice sets.

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

What is STEM?

Before proceeding, we present our definition of STEM for this analysis. Various STEM classifications exist, differing in their approaches to categorizing certain groups of oc-

[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), J. Hastings, Kane, and Staiger (2009), Kapor, Neilson, and Zimmerman (2020), and Pathak and P. Shi (2021).

cupations 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 underrepresentation 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 (Ceci et al. 2014; George-Jackson 2011).

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 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 (S. Dougherty and Harbaugh Macdonald 2019; Gewertz 2018).

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, Stiefel, et al. 2014), while others find no positive impacts (Bottia et al. 2018; Eisenhart et al. 2015) 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 the gender gap in STEM in Mexico City is likely associated with some negative labor market consequences for females.

2.2 School choice in Mexico City

Here, we describe more specifically the types of STEM programs in Mexico City high schools and the choice mechanism. 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).

There are two subsystems of programs that 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 elite STEM programs focus on science and engineering and are affiliated with the Instituto Politécnico Nacional (IPN), a prestigious national polytechnic university. The elite non-STEM programs offer broader, liberal arts-style curricula and are affiliated with the Universidad Nacional Autónoma de México

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

(UNAM). There are 16 and 14 elite STEM and elite non-STEM programs, respectively.[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 non-elite 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 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 (elite non-STEM) 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 schools, 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] Finally, students are assigned to schools using a serial dictatorship algorithm that is a special case of the student-proposing deferred acceptance algorithm characterized by Gale and Shapley 1962. Specifically, a computer orders all

[5]See Dustan, de Janvry, and Sadoulet (2017) for a more detailed description of the elite programs as well as their associated returns. One exception to the IPN/UNAM STEM classification is a small UNAM-affiliated nursing program that was offered until 2004, the first year of our sample.

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

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

students according to their placement test scores, and moving from highest scoring to lowest scoring, assigns each student to her highest-ranked school with a remaining seat when her turn arrives.

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 (84%) are assigned through this algorithm, while the remaining 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 A.4 that accounting for assignments in this later round does not substantively affect the size of the STEM gender gap using data from 2005, the one cohort for which we have full assignment data.

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

Let U_{ij} be student i's utility from being assigned to program j. The student's problem is to choose the rank-ordered list (portfolio) of programs that maximizes her expected utility, subject to the portfolio size constraint of 20 programs. Because COMIPEMS uses a student-proposing deferred acceptance (SPDA) mechanism, if there were no portfolio size constraint, a dominant strategy would be to rank programs in order of U_{ij} (Gale and Shapley 1962). This ranking is independent of the student's beliefs about her assignment probabilities as determined by program cutoffs and her own score, since the SPDA does not punish students for ranking programs from which they are rejected. Because only 2.4% of students in the data exhaust all 20 program choices, we treat the length constraint as non-binding. Following the notation in Abdulkadiroğlu et al. (2020), given a set of available programs $\mathcal{J} = \{1, 2, ..., J\}$, the student's first choice is:

$$R_{i1} = \underset{j \in \mathcal{J}}{\operatorname{argmax}} \, U_{ij},$$

while the following choices maximize over the set of programs excluding those which have already been chosen:

$$R_{ik} = \underset{j \in \mathcal{J} \setminus R_{im}: m<k}{\operatorname{argmax}} U_{ij}, k > 1.$$

The resulting portfolio $R_i = \left(R_{i1}, ..., R_{i\ell(i)}\right)'$ is an ordered list of programs of length $\ell(i)$, the number of programs chosen by student i.

Estimating student preferences for programs requires us to parameterize the utility function and make an assumption about the distribution of the unobserved factors affecting utility. We make modeling decisions informed by the empirical context and feasibility of estimation. Consider as a starting point the model in Abdulkadiroğlu et al. (2020):

$$U_{ij} = \delta_{c(X_i)j} + \tau_{c(X_i)} D_{ij} + \eta_{ij}. \tag{1}$$

This model permits rich heterogeneity with respect to student observables X_i by allowing separate program-specific utilities δ in each covariate cell $c(X_i)$, as well as allowing for cell-specific disutilities τ from distance from home D_{ij}. Including geographical region fixed effects in X_i and aggregating programs outside of the region into a single outside option limits the number of alternative-specific utilities that need to be estimated in each cell.

We deviate from this baseline model in two ways. The first relates to the way that student covariates enter the model. The covariate vector X_i consists of gender, COMIPEMS exam year, region fixed effects, exam score quintile, a parental high school education indicator, tercile in a measure of STEM comparative advantage on the exam, and an indicator for whether the applicant has already graduated from middle school.[8] Some of the cells are small and result in large numbers of program-specific mean utilities being unidentified when estimating Equation 1 because they are never ranked by students in the cell. We estimate a slightly less flexible model that collapses some dimensions of the cells so that they are defined only by gender, COMIPEMS exam year, region fixed effects, and exam score quintile, which are collected in X_{1i}. We refer to the cells defined by X_{1i}, $c(X_{1i})$, as *estimation cells*. The remaining covariates (X_{2i}, containing parental education, STEM comparative advantage, and graduation indicator) enter into U_{ij} directly and through interactions with program type indicators Z_j. The cells defined by the full set of covariates, $c(X_i)$, are called *covariate cells*.

[8]This results in 5,760 covariate cells (960 per year). STEM comparative advantage will be defined in the data description.

Second, we account for choice behavior induced by the exam-based assignment mechanism. Because assignment priority is determined solely by exam scores, rejection from a program with cutoff s_j implies rejection from any other program with cutoff $s_k > s_j$. Given that program cutoff scores are quite stable over time, students who understand this aspect of the mechanism may leave some programs unranked even if they are preferred to some of their ranked programs.[9] In practice, many students do rank programs with cutoff scores exceeding their higher-ranked programs despite their infeasibility. To avoid downward bias in estimated preferences for high-cutoff programs and to eventually simulate portfolios that more closely match those in the data, we account for this behavior in a parsimonious way by including a linear "cost" function that penalizes programs when their cutoff scores exceed the minimum cutoff among already-ranked programs.

The resulting parameterization of student utility for a program ranked k is:

$$\begin{aligned} U_{ijk} = \delta_{c(X_{1i})j} + \sum_{q=1}^{Q} \sum_{p=1}^{P} \beta_{c(X_{1i}),qp} X_{2i,q} Z_{j,p} + \tau_{c(X_{1i}),0} D_{ij} \\ + \sum_{q=1}^{Q} \tau_{c(X_{1i}),q} X_{2i,q} D_{ij} + \psi_{c(X_{1i}),0} S_{ijk} + \sum_{q=1}^{Q} \psi_{c(X_{1i}),q} S_{ijk} + \eta_{ij} \end{aligned} \tag{2}$$

where Z_j is a vector of program characteristics containing a constant and indicators for program type and $S_{ijk} = max\left\{0, s_j - min\left\{s_m : m < k\right\}\right\}$ is the amount by which program j's cutoff score exceeds the lowest cutoff already in the portfolio. Assuming that η_{ij} is independent of the observables in the model and follows an i.i.d. extreme value type I distribution leads to a conditional logit model, allowing Equation 2 to be estimated via maximum likelihood, separately by estimation cell. The choice set varies by region and year. All elite programs are included in the choice set for every region-year. Non-elite programs are aggregated into a region-year-specific outside option with utility normalized to zero if they are outside of a buffer surrounding the region.[10] Once a program is chosen, it is dropped from the choice set for subsequent choices, except for the outside option.

[9] In particular, it is common for students to list several elite programs at the top of their list, followed by non-elite programs, with no elite programs following the first non-elite one.

[10] The width of the buffer varies by region-year and is chosen such that 90% of non-elite choices, regardless of rank, are within the buffer. Buffer widths range from 2.7 km to 10.1 km.

3.2 Willingness to travel

The estimated student preference parameters need to be transformed in order to summarize STEM preference in the population and compare it between genders. For the population of students in covariate cell $c(X_i)$, the estimated average utility from program j (disregarding disutility from distance) is the sum of the $c(X_{i,1})$-specific program fixed effect (which varies by gender-year-region-quintile) and the within-$c(X_{i,1})$ deviation in utility from program characteristics due to the other covariates X_{2i} (parental education, STEM comparative advantage, and middle school graduate status):

$$\bar{V}_{c(X_i)j} = \widehat{\delta}_{c(X_{1i})j} + \sum_{q=1}^{Q} \sum_{p=1}^{P} \widehat{\beta}_{c(X_{1i}),qp} X_{2i,q} Z_{j,p}.$$

Similarly, the average utility of distance for a covariate cell is:

$$\bar{\tau}_{c(X_i)} = \widehat{\tau}_{c(X_{1i}),0} + \sum_{q=1}^{Q} \widehat{\tau}_{c(X_{1i}),q}.$$

The estimated average willingness to travel (WTT) for students in covariate cell $c(X_i)$ to program j thus:

$$\text{WTT}_{c(X_i)j} = -\frac{\bar{V}_{c(X_i)j}}{\bar{\tau}_{c(X_i)}}.$$

This WTT is in comparison to the outside option, which varies by gender-year-region-quintile.

In order to estimate average marginal WTT for program characteristics, we project program-cell WTT estimates on covariate cell fixed effects ($\mu_{c(X_i)}$, below), which account for differences in preference for the outside option, and program type indicators Z_j:

$$\text{WTT}_{c(X_i)j} = \mu_{c(X_i)} + \pi Z_j + \varepsilon_{c(X_i)j}. \tag{3}$$

The coefficients of interest are π, which give the average marginal WTT for different program types compared to the reference category, traditional non-elite programs. Observations are weighted by the number of students in the estimation cell. Standard errors are clustered at the estimation cell level.

To compare marginal WTT for program types between genders, we estimate:

$$\text{WTT}_{c(X_i)j} = \mu_{c(X_i)} + \pi Z_j + \gamma(\text{male}_i \times Z_j) + \varepsilon_{c(X_i)j}. \tag{4}$$

Finally, adding program-by-$c(X\backslash\text{male}_i)$ fixed effects, which correspond to cells generated by all student covariates except gender, identifies average gender differences in WTT using only variation between males and females with the same values of all other covariates.

3.3 Counterfactual simulations

Simulating counterfactual STEM assignment gaps under alternative priority structures and preferences requires two steps. First, the estimated (or assumed) preference parameters are used to simulate choice portfolios for the student population. Second, these portfolios are submitted to the assignment mechanism using the actual or counterfactual priority structure to obtain simulated program assignments. This section details both steps.

Generating portfolios using the estimated preference parameters is straightforward. To simulate the first choice we compute, for each student i in the choice estimation sample and each program j in their choice set $\mathcal{J}_{c(X_i)}$, the observable part of the utility:

$$\bar{U}_{ij1} = \bar{V}_{c(X_i)j} + \bar{\tau}_{c(X_i)} D_{ij},$$

again with the utility of the outside option normalized to zero. The conditional logit model gives a simple form for the probability of choosing each program:

$$P_{ij1} = \frac{\exp(\bar{U}_{ij1})}{\sum_{m \in \mathcal{J}_{c(X_i)}} \exp(\bar{U}_{ijm})}.$$

The first choice program is selected randomly according to these probabilities. The chosen program is then removed from the portfolio for subsequent choices (except for the outside option). For the second and later choices, $\bar{U}_{ijk}$ also includes the linear cost term from Equation 2. Because we do not model the student's stopping decision (i.e. when to stop listing choices), we take portfolio length $\ell(i)$ as exogenously given and use each student i's observed portfolio length from the data.[11]

For each year, these portfolios are submitted to the SPDA mechanism to generate simulated assignments. Program capacities (seat counts) match those used in the respective year's actual assignment process.[12] While in the true COMIPEMS assignment process,

[11]Once the portfolios 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.

[12]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.

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 quota fixed and use a single random tiebreaker. Once the assignment algorithm terminates, we summarize the simulated assignment outcomes, in particular the STEM gender gap.

After simulating the assignment gap under the estimated preferences and the COMIPEMS SPDA mechanism, we explore several counterfactuals to understand how sensitive the STEM gender gap is to hypothetical policies. These fall into two broad groups. The first is counterfactual priority structures, such as lottery assignment and point-based affirmative action. In all cases, the mechanism remains SPDA, so that the assignment simulation is unchanged except that the assignment priorities that students have at programs are different. We also consider the second-order effects that arise if priority structure changes affect students' stated preferences, i.e., girls receiving point-based affirmative action choose programs as if they were actually higher-scoring girls. The second set of simulations consists of imposing counterfactual preferences, leaving the priority structure unchanged but assuming that policies have increased girls' WTT for STEM programs. This is done by increasing, for girls only, the fixed effects associated with STEM programs.

Finally, we simulate distance-denominated changes in population and gender-specific average consumer surplus to show how counterfactual priority structures affect the level and distribution of welfare.[13] 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. For example, in the simulation with the status quo COMIPEMS priority structure, a student with a score of 50 has a feasible set composed of all programs for which the simulated cutoff score was 50 or below. Given preferences and a feasible choice set $\mathcal{J}_i^f$, Small and Rosen (1981) and Williams (1977) show that expected consumer surplus can be computed up to a constant:[14]

$$\mathrm{E}\left(CS_i\right) = \frac{1}{-\bar{\tau}_{c(X_i)}} \ln\left(\sum_{j \in \mathcal{J}_i^f} \exp(\bar{U}_{ij})\right) + C.$$

Differences in $\mathrm{E}\left(CS_i\right)$ between two priority structures result from differences in the feasi-

[13]The general approach and exposition here follow Train (2009).

[14]This assumes that students construct choice portfolios such that for any realized assignment priorities, they will be assigned to their most-preferred feasible program. This is consistent with our model of portfolio construction and the empirical fact that few students are constrained by the portfolio length limit.

ble sets they generate, f_i^0 and f_i^1:

$$\Delta \mathrm{E}\left(CS_i\right) = \frac{1}{-\bar{\tau}_{c(X_i)}}\left[\ln\left(\sum_{j\in\mathcal{J}^{f_i^1}} \exp(\bar{U}_{ij})\right) - \ln\left(\sum_{j\in\mathcal{J}^{f_i^0}} \exp(\bar{U}_{ij})\right)\right]. \tag{5}$$

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 population.

4 Data

The data used to analyze students' school choices and assignments are drawn from the COMIPEMS administrative databases. We include data from the years 2004 through 2009, years for which we can consistently classify STEM and non-STEM programs.[15] 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 placement exam scores, 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.

The analytical sample consists of students eligible for assignment and for whom we are able to estimate preferences.[16] 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 3.4% 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 analytical sample who were successfully assigned to a program (in actuality or according to the policy simulations).

[15]We exclude years prior to 2004 because we do not have program codes within technical schools. We exclude years after 2009 because of a change that allowed students to choose programs after enrolling in schools, making us unable to classify their choices as STEM or non-STEM.

[16]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 A.3. 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.[17]

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 exam on average, scoring 0.2 standard deviations higher than females, while females obtain 0.5 standard deviations higher middle school grades than males. Figure 1 displays the densities of middle school GPA and placement exam score and indicates that these average differences are not driven by outliers but rather by large differences throughout the distribution. The test score differential grows to a 0.3 standard deviation male advantage when looking specifically at the STEM subscore on the placement test (quantitative reasoning, math, physics, chemistry, and biology sections). Males also have a higher measure of STEM comparative advantage, defined as the STEM subscore divided by the non-STEM subscore.

With respect to choices, males choose STEM schools more than females. Specifically, 35.7% of male choice portfolios are STEM programs (14.1% elite STEM and 21.6% non-elite STEM), while only 25.7% of female choice portfolios are STEM programs (10.4% elite STEM and 15.3% non-elite STEM). In total, females list slightly more schools than males (9.2 versus 9.0), being more likely to list elite non-STEM and technical non-STEM schools on average. Males and females both list schools that are 8.7 km away on average. Males are slightly more likely to come from households with higher parental education.

Table 2 presents the summary statistics for program assignment by gender for students in the analysis sample assigned by the placement algorithm. 41.1% of males are assigned to STEM schools compared to 29.7% of females, generating a STEM gender gap of 11.4 p.p.. This comes from 5.4 and 6.0 p.p. gender gaps in elite STEM and non-elite STEM schools, respectively. Females are 0.9 p.p. more likely to be assigned to elite non-

[17] 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.

STEM programs, 3.4 p.p. more likely to be assigned to technical non-STEM programs, and assigned to schools 0.2 km closer than males.

Figure 2 examines heterogeneity in assignment by achievement, displaying the assignment gap by placement exam score quintile. The total gap is slightly higher in the highest quintiles (13.1 and 13.5 p.p.) compared with the lowest three quintiles (10.2 to 12.0 p.p.). The gap in non-elite STEM assignment dominates in the lower quintiles, and the gap in elite STEM assignment grows in higher quintiles, as assignment to more competitive schools grows with higher placement scores.

5 Results

This section presents the results from estimating the school choice model. First, we briefly assess model fit. Next, we present the WTT results, followed by the counterfactual simulations of the STEM gap and welfare under alternative policies.

5.1 Model fit

The estimated model parameters produce choice portfolios that match the data quite well. A summary of the model fit is presented in Table 3, comparing the simulated and actual proportions of males and females choosing different program types, as well as their differences. When all choices are pooled regardless of ordinal ranking, the simulated gender gaps are almost identical. In particular, the simulated 10.4 p.p. gap for STEM programs closely matches the 10.3 p.p. gap observed in the data. Restricting to students' first choice, the model's fit is less exact. The simulated first-choice STEM gap of 9.5 p.p. is similar to the actual 10.9 p.p. gap and the elite STEM gaps are nearly identical. The simulated proportions choosing non-elite STEM programs and the resulting gap do not match the data as closely: we simulate a 2.6 p.p gap, while the true difference is 4.2 p.p. The fit for the fifth choice is better, matching the STEM gap and its components closely.

Implementing the assignment mechanism on the simulated portfolios results in STEM assignment gaps that are similar to those in the data. The simulated overall STEM gap is 10.3 p.p., compared to the true 11.4 p.p. gap. The 1.1 p.p. discrepancy between the simulated and actual gaps comes from both the elite STEM gap (underestimated by 0.5 p.p.) and the non-elite STEM gap (underestimated by 0.6 p.p.). The elite and technical non-STEM gaps are also similar to those in the data.

5.2 Student preferences

We begin describing students' preferences by presenting the WTT estimates from equation 3, presented in Table 4. Columns 1 and 2 give similar results excluding and including covariate cell fixed effects, respectively. We focus on the covariate cell fixed effects results in column 2. On average, students have a high WTT for elite schools, both STEM and non-STEM. Compared to the base category of non-elite traditional high school programs, students are on average willing to travel 14.7 km farther to an elite STEM program. This is more than double the average distance between students and their assigned programs. The WTT for elite non-STEM programs is even higher, at 20.5 km. On the other hand, WTT for technical programs is negative compared to traditional academic options, with little difference in average preferences with respect to whether that technical program is STEM-oriented. WTT for non-elite STEM programs is -7.6 km compared to -7.8 km for technical non-STEM programs.

There is significant heterogeneity in preferences with respect to the student's exam score quintile, as summarized in Figure 3. WTT for elite STEM programs doubles between students in the first and fifth quintiles of the score distribution, a pattern that is similar for elite non-STEM programs. This stark pattern does not appear for non-elite preferences. The relationship between exam score and revealed preference for elite options may be due not only to heterogeneity in true preferences, but also due to students who anticipate low scores and choose not to list options they believe will be infeasible for them. This behavior may persist even though the SPDA mechanism does not punish students for listing options from which they are rejected.

Significant gender differences exist in preferences for STEM programs, as shown in columns 3 and 4 of Table 4. The results in column 3 do not include interactions of program fixed effects with covariate cell (excluding gender as a dimension), so that differences in preferences reflect both gender differences in covariates and gender differences in preferences conditional on other covariates. Males have a 3.0 km higher WTT to elite STEM programs than females, whose WTT is 13.2 km. WTT for non-elite STEM programs is 3.7 km higher for males than females, whose WTT is -9.3 km. Preferences for non-STEM options appear very similar in this specification.

The inclusion of program-by-covariate cell (excluding gender) fixed effects in column 4 ensures that comparisons are made between males and females who are otherwise observably similar. The estimated elite STEM preference gap falls from 3.0 to 2.3 km, indicating that about 25% of the apparent preference difference between males and females is attributable to differences in observable student characteristics such as exam score. For the same reason, females are now shown to have 0.42 km higher WTT for elite non-STEM programs than males. The finding that males have stronger preferences for non-elite STEM programs persists.

Gender differences in STEM preferences are persistent throughout the exam score distribution. Figure 4 shows the estimated gender differences from estimating the program-by-covariate cell fixed effect specification separately by quintile. Elite STEM WTT is at least 2 km higher for males in every quintile, and nearly 3 km higher in the top score quintile, where scores are almost always high enough for admission to students' preferred elite programs. Non-elite STEM preferences are also consistently higher for males.

In summary, the WTT estimates from the choice model indicate significant gender differences in STEM preferences. The preference gap is in part attributable to gender differences in observable characteristics that predict STEM preferences, but it persists even when comparing observably similar groups of males and females. The gap is not due to differences in the exam score distributions, as preference differences exist in each exam score quintile.

5.3 Counterfactual simulations

This section describes hypothetical changes to the priority structure and their simulated impacts on the STEM gender gap. 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). This analysis is followed by a similar exercise that holds the priority structure constant but simulates the impacts of various preference "nudges" that increase female WTT for STEM programs. For both priority structure changes and preference interventions, the impacts on the gender gap and its composition of elite and non-elite components vary significantly across exam score quintiles, requiring a nuanced evaluation of policy impacts on different types of students.

5.3.1 Alternative priority structures

Table 5 restates the simulated STEM gap and its elite and non-elite components under the status quo COMIPEMS priority structure, followed by the changes in this gap when this structure is altered. For these simulations, the choice portfolios are left unchanged from those simulated using the model parameters. The only change is the priority structure used by the SPDA mechanism to map these portfolios to program assignments. First, determining assignment priorities using a single lottery *increases* the gender gap in STEM assignment by 1.2 p.p., an increase of 11.9% over the baseline gap of 10.3 p.p. While the elite component of the STEM gap decreases by 36.7% because females' exam score disadvantage is removed, the non-elite gap increases by 55.7%. This increase is consistent with the preference parameter estimates, which show that females on average prefer non-STEM programs. Similarly, modifying the priority structure to prioritize students based on an “academic index” equally weighting normalized exam score and normalized middle school GPA, thus advantaging females, increases the STEM gap by 37.1%.

The effects of affirmative action policies depend on whether they advantage females at all programs (“AA”), or only at STEM programs (“STEM AA”). AA is implemented in the SPDA mechanism by adding, for each female applicant, a constant number of points to the exam score that determines assignment priority to all programs. STEM AA is similar, except that it gives students different priorities to STEM and non-STEM programs. Priority at non-STEM programs relies on the unmodified exam score, but priority at STEM programs is determined by the exam score plus the constant added to females’ scores.

A general AA program granting females 4 extra points (enough to equalize means between genders) would widen the STEM gap by 17.7%, with declines in the elite gap and 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 44.3%, reducing the elite gap by 38.0% 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.

STEM AA, in contrast, closes the STEM gap and its components. A 4-point STEM AA program would close the STEM gap by 41.5%, with similar reductions in elite and non-elite components. The larger 10-point STEM AA program is sufficient to eliminate the gap entirely, while actually reversing the elite STEM gap from 4.9 p.p. to -1.0 p.p.

Beyond the simulated changes in the overall gap, there is significant heterogeneity in policy impacts with respect to exam score. Panels A and B of Figure 5 show the quintile-specific effects of using the academic index and 10 point AA priority structures, respectively. Both academic index and general AA have minimal effects for the lowest scorers, but increase the non-elite STEM gap in higher-scoring quintiles as females’ higher priorities allow them to secure seats in their more-preferred programs, which are often non-STEM. The elite gap closes significantly in quintiles 3 and 4, where females who were previously below the cutoffs of elite STEM programs are now admitted. In Panel C, STEM AA reduces the non-elite gap in the first three quintiles by large margins. On the other hand, the non-elite gap increases slightly in quintiles 4 and 5 as females move from non-elite to elite STEM programs. The elite gap falls more in the upper quintiles than in the general AA case, because it advantages marginal females only at STEM elite programs rather than all elite programs.

5.3.2 Preference nudges

The final rows of Table 5 show the simulated effects of leaving the priority structure unchanged while implementing policy "nudges" that increase females' WTT to STEM programs. The "Nudge (all)" row corresponds to a 1 km increase in STEM WTT for both elite and non-elite programs for all females. This is an increase of about 11% over the 8.7 km average distance between students and their chosen programs. A preference change of this magnitude closes the STEM gap by 31.1%, but most of this decrease comes from a reduction in the non-elite gap rather than the 20.7% decrease in the elite gap. Panel D of Figure 5 shows that these reductions in the STEM gap come from all quintiles of the exam score distribution, with reductions in the elite gap necessarily driven by females in the upper two quintiles.

More targeted nudges, whether in terms of the programs affected or the females targeted, have smaller effects. When the nudge only increases STEM WTT for local schools ("Nudge (local)"), within 5 km of the student's home, the elite gap is nearly unchanged (-5.9%) because students live close to few, if any, elite STEM schools. The non-elite gap is more responsive because such options are more available locally—it closes by 19.1%. Student-level targeting, simulated as a nudge that only affects females in the top tercile of the STEM comparative advantage distribution ("Nudge (comp. adv.)"), closes the STEM gap by 8.6%, slightly less than one third of the effect when the nudge affects all terciles.

In order to understand the role of preferences and score distributions in explaining the STEM gap, the final row of Table 5 assigns to females the preference parameters of males who have the same values of all covariates (region, exam score quintile, etc.). When preference equality is enforced in this way so that the genders differ only in terms of exam score distributions, the STEM gap reverses, from 10.3 p.p to -1.0 p.p. Despite the overall reversal, the elite STEM gap is only closed by 64.3%, due to females' lower exam score distribution. In contrast, the non-elite gap's 151.5% reduction reflects the fact that non-elite STEM programs are generally less competitive than other options, so that when females' preferences shift toward such programs they are very likely to be assigned to them.

5.3.3 Welfare effects of priority structure changes

Table 6 summarizes the simulated changes in welfare under counterfactual priority structures, compared to the COMIPEMS status quo. Columns 1 and 2 show that each of the proposed priority structures reduces average welfare for males to varying degrees, while increasing welfare for females. Using a lottery or academic index to determine assignment priority reduces male welfare substantially, equivalent to commuting 1.2 or 1.3 km farther, respectively. Females gain slightly (0.1 km) from lottery assignment and greatly (1.0 km) from using the academic index for assignment. For perspective, males are on average 7.2 km from their assigned school, while females are on average 7.0 km away. Columns 3 and 4 show that both priority structures reduce overall welfare while widening the STEM gap. The decline in overall welfare is likely because these priority structures assign more students with low exam scores to high-cutoff programs. These students have lower WTT for such programs compared to their high-scoring counterparts. This is especially true for the lottery.

The welfare impacts of AA differ greatly based on whether or not it is STEM-specific. The general AA approach has a near-zero effect on overall welfare, but it substantially increases female welfare and decreases male welfare, while widening the STEM gap significantly. With 10 point AA, average male welfare is reduced by 1.5 km, while females gain by 1.5 km. In contrast, STEM AA closes the STEM gap (entirely, in the 10 point case) while having an approximately zero impact on overall welfare in the student population. The gendered impacts of STEM AA are more modest than other priority structure changes: in the 10 point case, male welfare declines by 0.5 km, while female welfare increases by 0.4 km. A likely source of this smaller welfare impact is that elite programs, particularly non-STEM ones, remain feasible for more males under STEM-specific AA.

The exam score quintile-specific welfare changes are illustrated in Figure 6 for selected priority structures. Panel A shows that using the academic index for assignment priority greatly increases the welfare of females with low exam scores, since they (on average) benefit from having high GPAs relative to males. High-scoring males experience large welfare losses due to their relatively low GPAs. Panel B shows that 10 point AA gives a more uniform pattern of welfare reallocation from males to females, with larger changes in the middle quintiles because these students are more likely to be close to an assignment cutoff, in particular for elite programs. Panel C illustrates the effects of 10 point STEM AA, which are uniformly small in comparison with the other policies. STEM AA, then, closes the STEM gap with relatively low welfare changes across genders and across the entire exam score distribution.

5.3.4 Simulating preference effects

Thus far, our simulated changes in the priority structure have not allowed student preferences to change endogenously in response. This is reasonable in the sense that all proposed priority structures continue to be part of a SPDA mechanism, such that no incentive to misreport preferences has been introduced. We now depart from this assumption, with the goal of understanding whether preference effects for females have the potential to amplify or dampen the effects of affirmative action. To allow for such effects of a general AA policy, we add the AA point "bonus" to females' scores and then use this score to determine their exam score quintile for the purpose of assigning them estimated preference parameters. For example, a student in the third exam score quintile may end up in the fourth quintile with the AA bonus, thus causing her to choose as if she were a higher scorer.

The first two rows of Table 7 compare the changes from AA without preference effects to the changes from AA with preference effects. Focusing on the 10-point AA policy, allowing preference effects dampens the increase in the STEM gap from 44.3% to 38.6%. This comes from a much larger reduction in the elite STEM gap, as higher-scoring girls now prefer elite STEM programs more. This is counterbalanced somewhat by an increase in the non-elite gap, since higher-scoring girls generally prefer non-elite STEM options less.

Extending this approach to STEM AA, we assign preferences for STEM programs using the exam score quintile in which a female student would reside when given the STEM-specific point bonus. Under this assumption, the 10-point STEM AA policy has a similar impact on the overall STEM gap, but the reversal in the elite gap is now even larger (-143.4% instead of -121.5%) and the decrease in the non-elite gap is smaller, again due to the fact that females in higher quintiles tend to prefer elite STEM programs more and non-elite STEM programs less.

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, provided that females actually list one or more STEM options in their choice portfolios (as is the case in our population). 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. 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 (even leaving aside the political feasibility of modifying any assignment process).

This paper also illustrates the utility and feasibility of using microsimulation to approximate the effects of various policy changes either changing a priority structure or affecting demand for different types of schools. 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 clear limitations to this approach, however. First, we model demand-side interventions that increase students' willingness to travel for STEM programs but do not specify how these would be implemented or whether our modelled "nudges" require large- or small-scale interventions. Here, we view our approach as a complement to the crucial and ongoing research focused on designing and evaluating interventions targeted at increasing females' likelihood of choosing STEM. Second, the model and simulations that we estimate are based on choices made under an existing system, but 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 composition of schools. Thus, even after modelling preference effects of priority structure changes, the long-run general equilibrium effects of the analyzed policies may differ markedly from the simulations. Furthermore, despite the fact that the SPDA mechanism in this paper is strategy-proof, a change in the priority structure that affects admission probabilities (such as affirmative action) may induce changes in revealed preferences beyond what we predict (Bobba and Frisancho 2020). Whether such effects are sufficient to substantively change the conclusions from policy simulations is unclear but is an interesting avenue for future research.

References

Abdulkadiroğlu, Atila et al. (2020). "Do parents value school effectiveness?" In: *American Economic Review* 110.5, pp. 1502–39.

Agarwal, Nikhil and Paulo Somaini (2018). "Demand analysis using strategic reports: An application to a school choice mechanism." In: *Econometrica* 86.2, pp. 391–444.

— (2020). "Revealed preference analysis of school choice models." In: *Annual Review of Economics* 12, pp. 471–501.

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

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

Beede, David N et al. (2011). "Women in STEM: A gender gap to innovation." In: *Economics and Statistics Administration Issue Brief* 04-11.

Bertrand, Marianne, Rema Hanna, and Sendhil Mullainathan (2010). "Affirmative action in education: Evidence from engineering college admissions in India." In: *Journal of Public Economics* 94.1-2, pp. 16–29.

Beuermann, Diether et al. (Dec. 2018). *What is a Good School, and Can Parents Tell? Evidence on the Multidimensionality of School Output*. Working Paper 25342. National Bureau of Economic Research. DOI: `10.3386/w25342`. URL: `http://www.nber.org/papers/w25342`.

Bobba, Matteo and Veronica Frisancho (2020). "Self-perceptions about academic achievement: Evidence from Mexico City." In: *Journal of Econometrics*. ISSN: 0304-4076. DOI: `https://doi.org/10.1016/j.jeconom.2020.06.009`. URL: `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." In: *Economics of Education Review* 77, p. 102011.

Bottia, Martha Cecilia et al. (2018). "Boosting the numbers of STEM majors? The role of high schools with a STEM program." In: *Science Education* 102.1, pp. 85–107.

Breda, Thomas et al. (2018). "Can female role models reduce the gender gap in science? Evidence from classroom interventions in French high schools." In:

Calsamiglia, Caterina, Chao Fu, and Maia Güell (2020). "Structural estimation of a model of school choices: The boston mechanism versus its alternatives." In: *Journal of Political Economy* 128.2, pp. 642–680.

Card, David and A Abigail Payne (2021). "High school choices and the gender gap in STEM." In: *Economic Inquiry* 59.1, pp. 9–28.

Carrell, Scott E, Marianne E Page, and James E West (2010). "Sex and science: How professor gender perpetuates the gender gap." In: *The Quarterly Journal of Economics* 125.3, pp. 1101–1144.

Ceci, Stephen J et al. (2014). "Women in academic science: A changing landscape." In: *Psychological Science in the Public Interest* 15.3, pp. 75–141.

Colleges, California Community (2018). *What is AB 705?* Tech. rep.

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." In: *Education Finance and Policy* 13.2, pp. 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." In: *Journal of Public Economics* 97, pp. 32–48.

Dougherty, Shaun and Isabel Harbaugh Macdonald (2019). "Can growth in the availability of STEM technical education improve equality in participation?: evidence from Massachusetts." In: *Journal of Vocational Education & Training*, pp. 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." In: *Journal of Human Resources* 52.3, pp. 756–799.

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

Eisenhart, Margaret et al. (2015). "High school opportunities for STEM: Comparing inclusive STEM-focused and comprehensive high schools in two US cities." In: *Journal of Research in Science Teaching* 52.6, pp. 763–789.

Fack, Gabrielle, Julien Grenet, and Yinghua He (2019). "Beyond truth-telling: Preference estimation with centralized school choice and college admissions." In: *American Economic Review* 109.4, pp. 1486–1529.

Francis, Andrew M and Maria Tannuri-Pianto (2012). "Using Brazil's racial continuum to examine the short-term effects of affirmative action in higher education." In: *Journal of Human Resources* 47.3, pp. 754–784.

Frisancho, Veronica and Kala Krishna (2016). "Affirmative action in higher education in India: targeting, catch up, and mismatch." In: *Higher Education* 71.5, pp. 611–649.

Fryer Jr, Roland G and Glenn C Loury (2005). "Affirmative action and its mythology." In: *Journal of Economic Perspectives* 19.3, pp. 147–162.

Gale, David and Lloyd S Shapley (1962). "College admissions and the stability of marriage." In: *The American Mathematical Monthly* 69.1, pp. 9–15.

George-Jackson, Casey (2011). “STEM switching: Examining departures of undergraduate women in STEM fields.” In: *Journal of Women and Minorities in Science and Engineering* 17.2.

Gewertz, Catherine (July 2018). “What is Career and Technical Education, Anyway?” In: *Education Week*. Accessed August 23, 2019. URL: 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.” In: *Educational Evaluation and Policy Analysis* 39.4, pp. 593–619.

Gnagey, Jennifer and Stéphane Lavertu (2016). “The impact of inclusive STEM high schools on student achievement.” In: *AERA Open* 2.2, p. 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.” In: *Education Finance and Policy* 11.2, pp. 177–202.

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

Hastings, Justine S and Jeffrey M Weinstein (2008). “Information, School Choice, and Academic Achievement: Evidence from Two Experiments.” In: *The Quarterly Journal of Economics* 123.4, pp. 1373–1414.

Hyde, Janet S et al. (2008). “Gender similarities characterize math performance.” In: *Science* 321.5888, pp. 494–495.

Jackson, C Kirabo (2010). “Do students benefit from attending better schools? Evidence from rule-based student assignments in Trinidad and Tobago.” In: *The Economic Journal* 120.549, pp. 1399–1429.

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

Kahn, Shulamit and Donna Ginther (2017). *Women and STEM*. Tech. rep. National Bureau of Economic Research.

Kapor, Adam J, Christopher A Neilson, and Seth D Zimmerman (2020). “Heterogeneous beliefs and school choice mechanisms.” In: *American Economic Review* 110.5, pp. 1274–1315.

Karp, Sarah (Oct. 13, 2021). “Big changes in how students are picked for CPS’ elite high schools start today.” In: *WBEZ Chicago National Public Radio*. URL: https://www.wbez.org/stories/big-changes-in-how-students-are-picked-for-

cps-elite-high-schools-start-today/27f4345f-4ed6-4f4e-a215-d577956504fd (visited on 10/18/2021).
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." In: *Journal of Career and Technical Education* 28.1, pp. 19–38.
Lim, Jaegeum and Jonathan Meer (2019). "Persistent effects of teacher-student gender matches." In: *Journal of Human Resources*, 0218–9314R4.
Means, Barbara et al. (2016). "STEM-focused high schools as a strategy for enhancing readiness for postsecondary STEM programs." In: *Journal of Research in Science Teaching* 53.5, pp. 709–736.
Ortega Hesles, Maria Elena and Shaun M. Dougherty (2017). *Academic Program Choice in Secondary Education: Regression Discontinuity Evidence from Mexico City*. Tech. rep. American Education Finance and Policy Conference Paper.
Owen, Ann L (2010). "Grades, gender, and encouragement: A regression discontinuity analysis." In: *The Journal of Economic Education* 41.3, pp. 217–234.
Pathak, Parag A and Peng Shi (2021). "How well do structural demand models work? Counterfactual predictions in school choice." In: *Journal of Econometrics* 222.1, pp. 161–195.
Polikoff, Morgan et al. (2018). "The impact of speedometry on student knowledge, interest, and emotions." In: *Journal of Research on Educational Effectiveness* 11.2, pp. 217–239.
Pop-Eleches, Cristian and Miguel Urquiola (2013). "Going to a better school: Effects and behavioral responses." In: *American Economic Review* 103.4, pp. 1289–1324.
Porter, Catherine and Danila Serra (2019). "Gender differences in the choice of major: The importance of female role models." In: *American Economic Journal: Applied Economics*.
Rothwell, Jonathan (2013). *The hidden STEM economy*. Metropolitan Policy Program at Brookings.
Schneeweis, Nicole and Martina Zweimüller (2012). "Girls, girls, girls: Gender composition and female school choice." In: *Economics of Education review* 31.4, pp. 482–500.
Shapiro, Eliza (Dec. 18, 2020). "New York City Will Change Many Selective Schools to Address Segregation." In: *The New York Times*. URL: https://www.nytimes.com/2020/12/18/nyregion/nyc-schools-admissions-segregation.html (visited on 10/18/2021).
Shi, Ying (2018). "The puzzle of missing female engineers: Academic preparation, ability beliefs, and preferences." In: *Economics of Education Review* 64, pp. 129–143.

Small, Kenneth A and Harvey S Rosen (1981). "Applied welfare economics with discrete choice models." In: *Econometrica: Journal of the Econometric Society*, pp. 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." In: *PloS one* 8.3, e57988.
Train, Kenneth E (2009). *Discrete choice methods with simulation*. Cambridge university press.
Treschan, Lazar et al. (2013). *The Meaning of Merit: Alternatives to Determining Admission to NYC's Specialized High Schools*. Tech. rep. Community Service Society.
Williams, Huw CWL (1977). "On the formation of travel demand models and economic evaluation measures of user benefit." In: *Environment and planning A* 9.3, pp. 285–344.
Wiswall, Matthew, Leanna Stiefel, et al. (2014). "Does attending a STEM high school improve student performance? Evidence from New York City." In: *Economics of Education Review* 40, pp. 93–105.
Wiswall, Matthew and Basit Zafar (2014). "Determinants of college major choice: Identification using an information experiment." In: *The Review of Economic Studies* 82.2, pp. 791–824.
Zafar, Basit (2013). "College major choice and the gender gap." In: *Journal of Human Resources* 48.3, pp. 545–595.

Tables

Table 1: Academic, choice, and demographic by gender

	(1) Male	(2) Female	(3) Difference
Academic outcomes			
Total placement test score (normalized within year)	0.11	-0.11	0.22***
	(1.02)	(0.97)	(0.00)
STEM subscore on placement test (normalized within year)	0.14	-0.14	0.28***
	(1.02)	(0.96)	(0.00)
STEM comparative advantage (STEM / non-STEM subscore)	1.05	1.00	0.05***
	(0.25)	(0.24)	(0.00)
Middle school GPA (normalized within year)	-0.28	0.27	-0.55***
	(0.95)	(0.97)	(0.00)
School choices			
Total number of choices listed	9.01	9.19	-0.18***
	(3.72)	(3.72)	(0.01)
Percent of choices that are elite STEM	14.12	10.37	3.76***
	(21.53)	(17.46)	(0.03)
Percent of choices that are non-elite STEM	21.64	15.33	6.31***
	(24.96)	(21.32)	(0.04)
Percent of choices that are elite non-STEM	26.24	31.28	-5.04***
	(28.77)	(29.84)	(0.05)
Percent of choices that are technical non-STEM	8.91	10.07	-1.16***
	(15.10)	(16.28)	(0.03)
Average distance to all choices (km)	8.68	8.66	0.03***
	(4.67)	(4.55)	(0.01)
Demographics			
High parental education (high school or more)	57.39	54.38	3.01***
	(49.45)	(49.81)	(0.08)
Observations	717935	761897	1479832

Note: Calculations are for all students in the analysis sample for the the 2004 through 2009 COMIPEMS cycles. The subject modules included in the STEM subscore are quantitative reasoning, math, physics, chemistry, and biology. 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 2: School assignment summary statistics, 2004-2009

	(1) Male	(2) Female	(3) Difference
STEM assigned program (elite or non-elite)	41.1	29.7	11.4***
	(49.2)	(45.7)	(0.1)
Elite STEM assigned program	12.0	6.6	5.4***
	(32.5)	(24.7)	(0.1)
Non-elite STEM assigned program	29.1	23.1	6.0***
	(45.4)	(42.2)	(0.1)
Elite non-STEM assigned program	15.6	16.6	-0.9***
	(36.3)	(37.2)	(0.1)
Technical non-STEM assigned program	12.5	15.9	-3.4***
	(33.1)	(36.6)	(0.1)
Distance to assigned program	7.2	7.0	0.2***
	(6.2)	(6.1)	(0.0)
Observations	626037	615879	1241916

Note: Calculations are for all students in the analysis who were sample assigned to a school by the placement algorithm in the 2004 through 2009 COMIPEMS cycles. 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
All choices						
STEM	36.6	26.2	10.4	36.9	26.6	10.3
Elite STEM	14.0	10.1	3.9	14.9	10.9	4.0
Non-elite STEM	22.6	16.2	6.4	21.9	15.7	6.3
Elite non-STEM	25.9	30.3	−4.4	25.8	30.1	−4.3
Technical non-STEM	8.9	10.6	−1.7	8.8	10.4	−1.6
First choice						
STEM	27.2	17.7	9.5	31.5	20.6	10.9
Elite STEM	19.2	12.4	6.9	18.0	11.3	6.7
Non-elite STEM	8.0	5.4	2.6	13.6	9.3	4.2
Elite non-STEM	49.8	57.7	−7.9	43.5	52.8	−9.3
Technical non-STEM	3.3	3.2	0.1	6.8	6.7	0.1
Fifth choice						
STEM	36.8	26.4	10.4	37.5	27.0	10.6
Elite STEM	14.1	10.4	3.7	15.7	11.9	3.7
Non-elite STEM	22.7	16.0	6.7	21.9	15.0	6.8
Elite non-STEM	23.2	27.9	−4.7	23.7	27.8	−4.1
Technical non-STEM	9.2	10.3	−1.1	8.5	9.6	−1.1
Assignment						
STEM	40.8	30.5	10.3	41.1	29.7	11.4
Elite STEM	11.6	6.7	4.9	12.0	6.6	5.4
Non-elite STEM	29.3	23.9	5.4	29.1	23.1	6.0
Elite non-STEM	15.2	16.5	−1.3	15.6	16.6	−0.9
Technical non-STEM	12.2	16.2	−4.0	12.5	15.9	−3.4

Note. Columns 1 through 3 report the simulated proportion of choices of, or assignments to, the indicated program type, and their difference. Columns 4 through 6 show the actual proportions in the data. Proportions are reported in percentages. Sample is 2004 through 2009 COMIPEMS cycles. "All choices" percentages differ slightly from those reported in Table 1 because Table 1 reports means of individual-level portfolio compositions, while this table is the mean of all (pooled) choices in the sample.

Table 4: Willingness to travel to different program types

	(1)	(2)	(3)	(4)
Elite STEM	14.88***	14.70***	13.22***	
	(0.260)	(0.256)	(0.304)	
Non-elite STEM	-7.52***	-7.56***	-9.34***	
	(0.141)	(0.140)	(0.135)	
Elite non-STEM	20.70***	20.51***	20.50***	
	(0.276)	(0.271)	(0.369)	
Technical non-STEM	-7.70***	-7.82***	-7.83***	
	(0.103)	(0.106)	(0.169)	
Male × Elite STEM			3.04***	2.30***
			(0.473)	(0.068)
Male × Non-elite STEM			3.68***	3.66***
			(0.184)	(0.049)
Male × Elite non-STEM			0.02	-0.42***
			(0.542)	(0.060)
Male × Technical non-STEM			0.02	0.10
			(0.205)	(0.108)
Observations	1040256	1040256	1040256	1040256
Adjusted R^2	0.580	0.625	0.630	0.888
Fixed effects	None	Covariate cell	Covariate cell	Covariate cell, (Covariate cell \ gender) × program

Note: Models listing fixed effects have one fixed effect per distinct combination of covariates: gender, COMIPEMS year, region, exam quintile, applicant type, parental education indicator, and STEM comparative advantage tercile. These fixed effects are fully interacted with gender or school fixed effects, as indicated. Standard errors clustered at the estimation cell level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: Effects of simulated priority structure changes and choice nudges on STEM gender gap and its components

	(1) Total	(2) Elite	(3) Non-elite
Baseline gap	10.3	4.9	5.4
Simulated changes in gap			
Lottery	1.2	−1.8	3.0
	[11.9%]	[−36.7%]	[55.7%]
Academic index	3.8	−1.7	5.5
	[37.1%]	[−34.1%]	[101.2%]
AA (4 pt)	1.8	−0.7	2.6
	[17.7%]	[−14.9%]	[47.1%]
AA (10 pt)	4.6	−1.8	6.4
	[44.3%]	[−38.0%]	[118.4%]
STEM AA (4 pt)	−4.3	−2.4	−1.9
	[−41.5%]	[−48.3%]	[−35.3%]
STEM AA (10 pt)	−10.6	−5.9	−4.7
	[−103.2%]	[−121.5%]	[−86.7%]
Nudge (all)	−3.2	−1.0	−2.2
	[−31.1%]	[−20.7%]	[−40.5%]
Nudge (local)	−1.3	−0.3	−1.0
	[−12.9%]	[−5.9%]	[−19.1%]
Nudge (comp. adv.)	−0.9	−0.4	−0.5
	[−8.6%]	[−7.7%]	[−9.4%]
Male preferences	−11.3	−3.1	−8.2
	[−110.2%]	[−64.3%]	[−151.5%]

Note. Levels and changes are in percentage points. Percent changes compared to baseline are in brackets. Sample is 2004 through 2009 COMIPEMS cycles.

Table 6: Effects of simulated priority structure changes on welfare

	(1) Male	(2) Female	(3) Overall	(4) Δ STEM gap (%)
Lottery	-1.23	0.12	-0.53	11.9
Academic index	-1.34	0.99	-0.14	37.1
AA (4 pt)	-0.61	0.59	0.01	17.7
AA (10 pt)	-1.54	1.46	0.01	44.3
STEM AA (4 pt)	-0.18	0.16	-0.00	−41.5
STEM AA (10 pt)	-0.45	0.37	-0.03	−103.2

Note: Values are simulated changes in mean expected consumer surplus from implementing the SPDA mechanism using the indicated priority structure instead of the status quo COMIPEMS structure. Surplus changes are expressed in kilometers of one-way commuting distance and are computed using Equation 5.

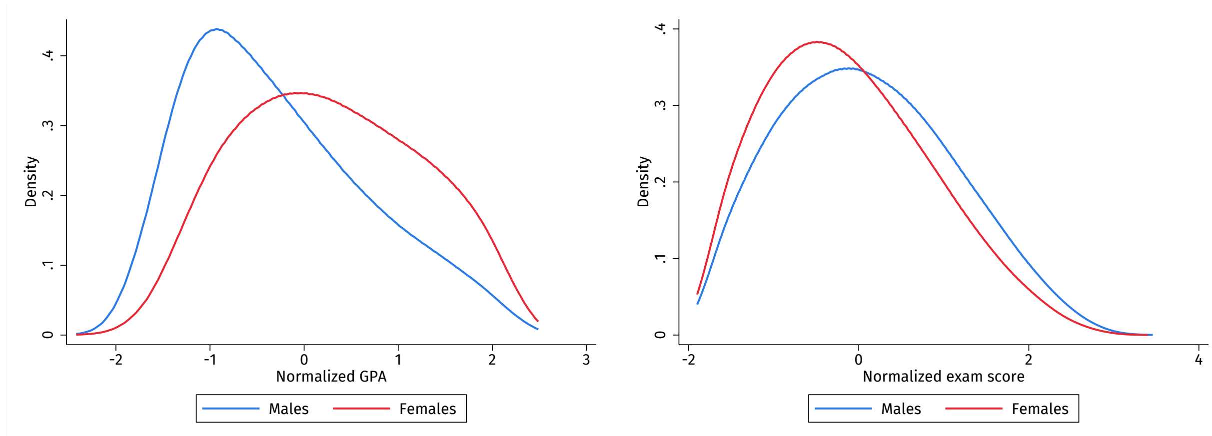
Table 7: Effects of simulated priority structure changes with and without preference effects

	With preference effects			Without preference effects		
	(1) Total	(2) Elite	(3) Non-elite	(4) Total	(5) Elite	(6) Non-elite
Baseline gap	10.3	4.9	5.4	10.3	4.9	5.4
Simulated changes in gap						
AA (4 pt)	1.6	−1.3	2.9	1.8	−0.7	2.6
	[15.4%]	[−27.3%]	[53.9%]	[17.7%]	[−14.9%]	[47.1%]
AA (10 pt)	4.0	−3.3	7.3	4.6	−1.8	6.4
	[38.6%]	[−68.6%]	[135.1%]	[44.3%]	[−38.0%]	[118.4%]
STEM AA (4 pt)	−4.4	−2.9	−1.5	−4.3	−2.4	−1.9
	[−42.6%]	[−60.1%]	[−26.8%]	[−41.5%]	[−48.3%]	[−35.3%]
STEM AA (10 pt)	−10.8	−7.0	−3.8	−10.6	−5.9	−4.7
	[−105.1%]	[−143.4%]	[−70.7%]	[−103.2%]	[−121.5%]	[−86.7%]

Note. Levels and changes are in percentage points. Percent changes compared to baseline are in brackets. Sample is 2004 through 2009 COMIPEMS cycles.

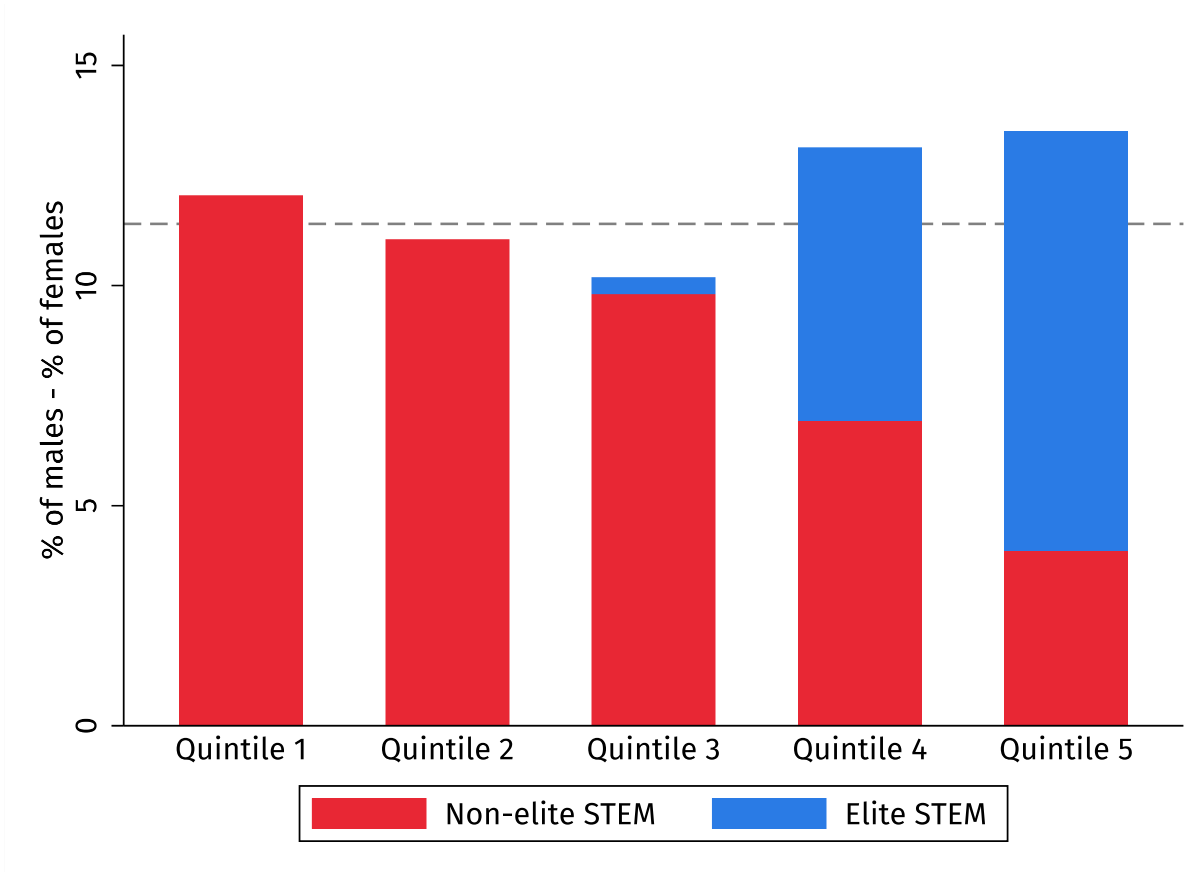
Figures

Figure 1: Achievement distributions



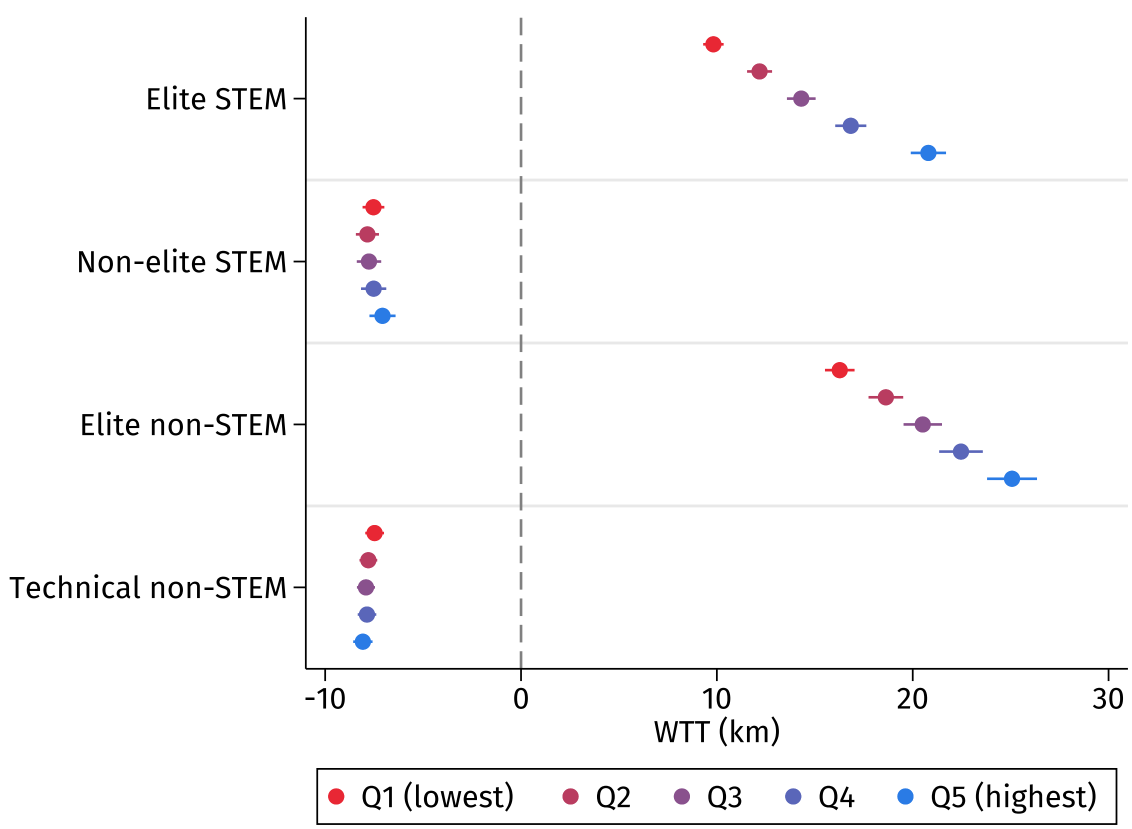
Note: Calculations are for all students in analysis sample assigned to a school by the placement algorithm in the 2004 through 2009 COMIPEMS cycles.

Figure 2: STEM gender gap in school assignment by placement test quintiles



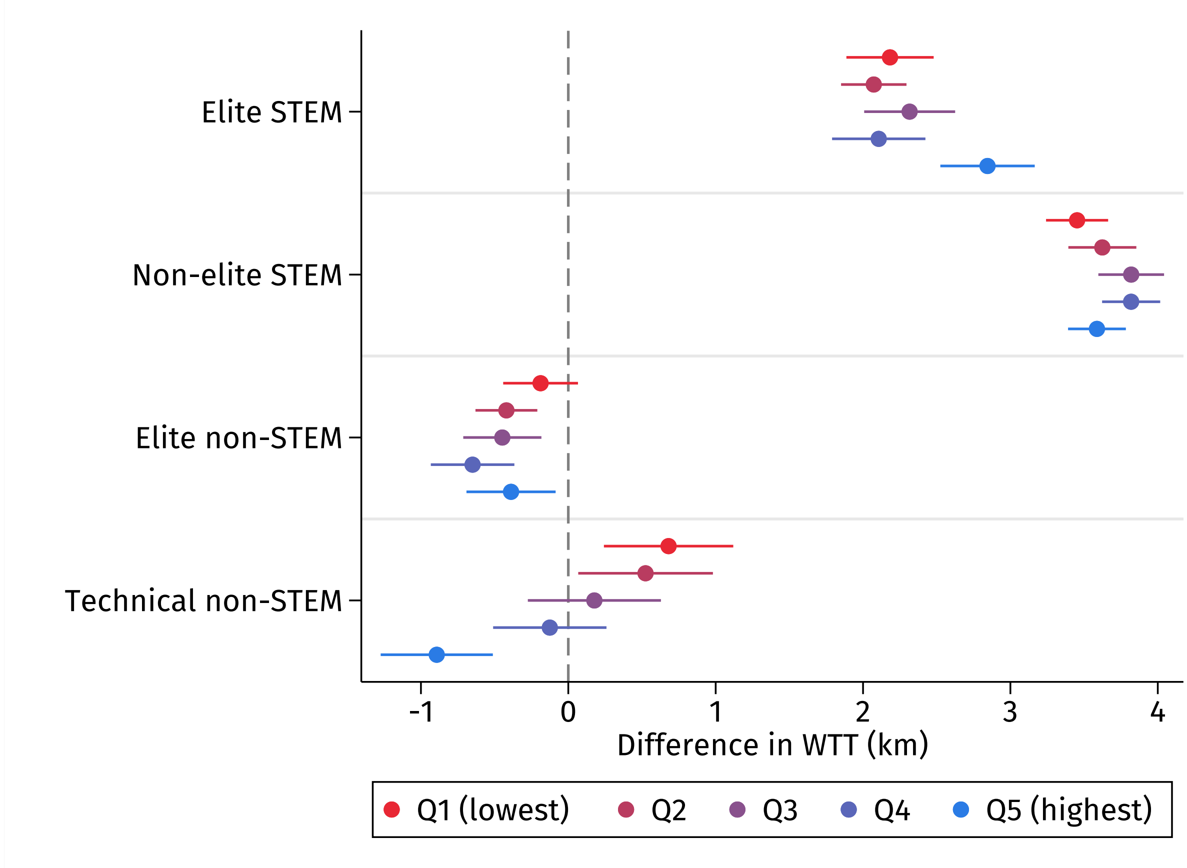
Note: Calculated percentages are for all students in analysis sample assigned to a school by the placement algorithm in the 2004 through 2009 COMIPEMS cycles. Test score quintiles are from the pooled (male and female) distribution of scores within each year.

Figure 3: Willingness to travel for program type, by score quintile



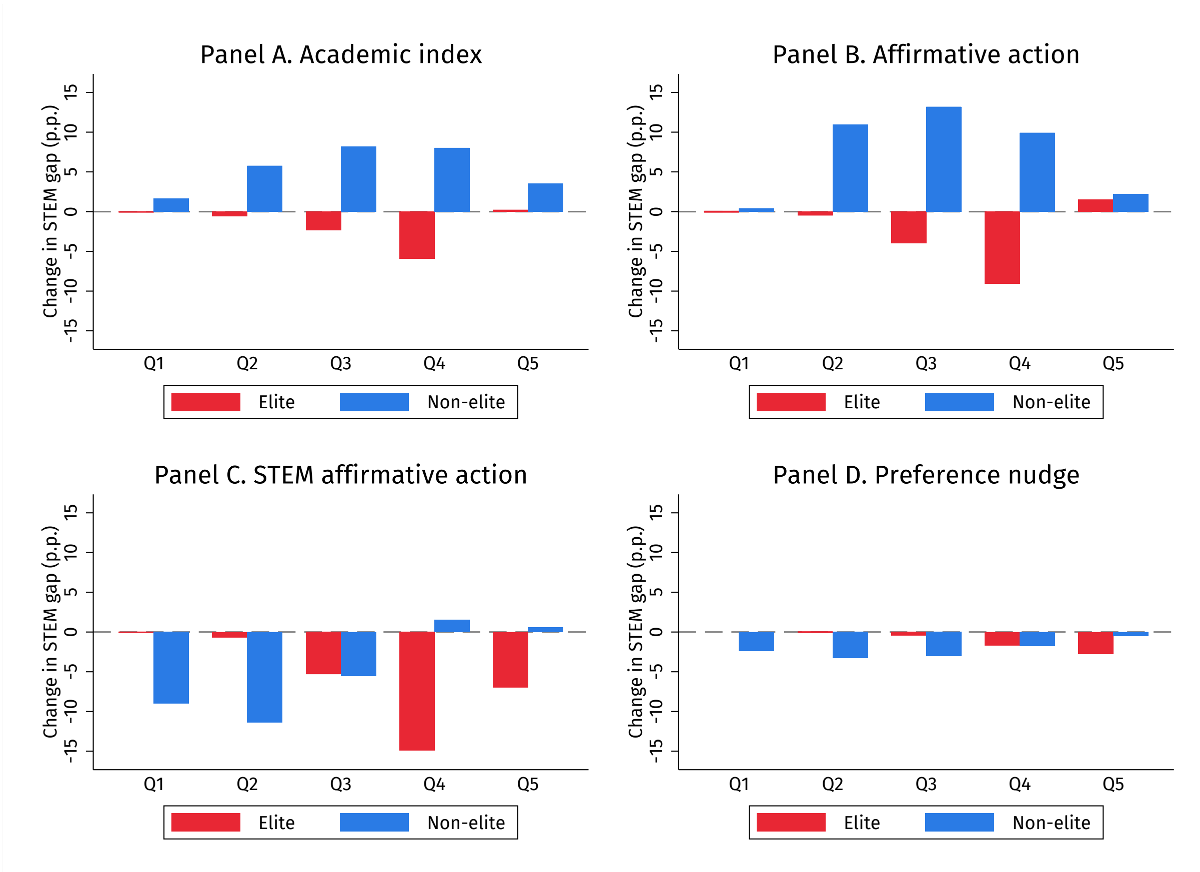
Note: Circles are estimated WTT coefficients for the specified program type from Equation 3. Lines are 95% confidence intervals. Base category is traditional academic program (non-elite, non-STEM).

Figure 4: Gender differences in willingness to travel for program type, by score quintile



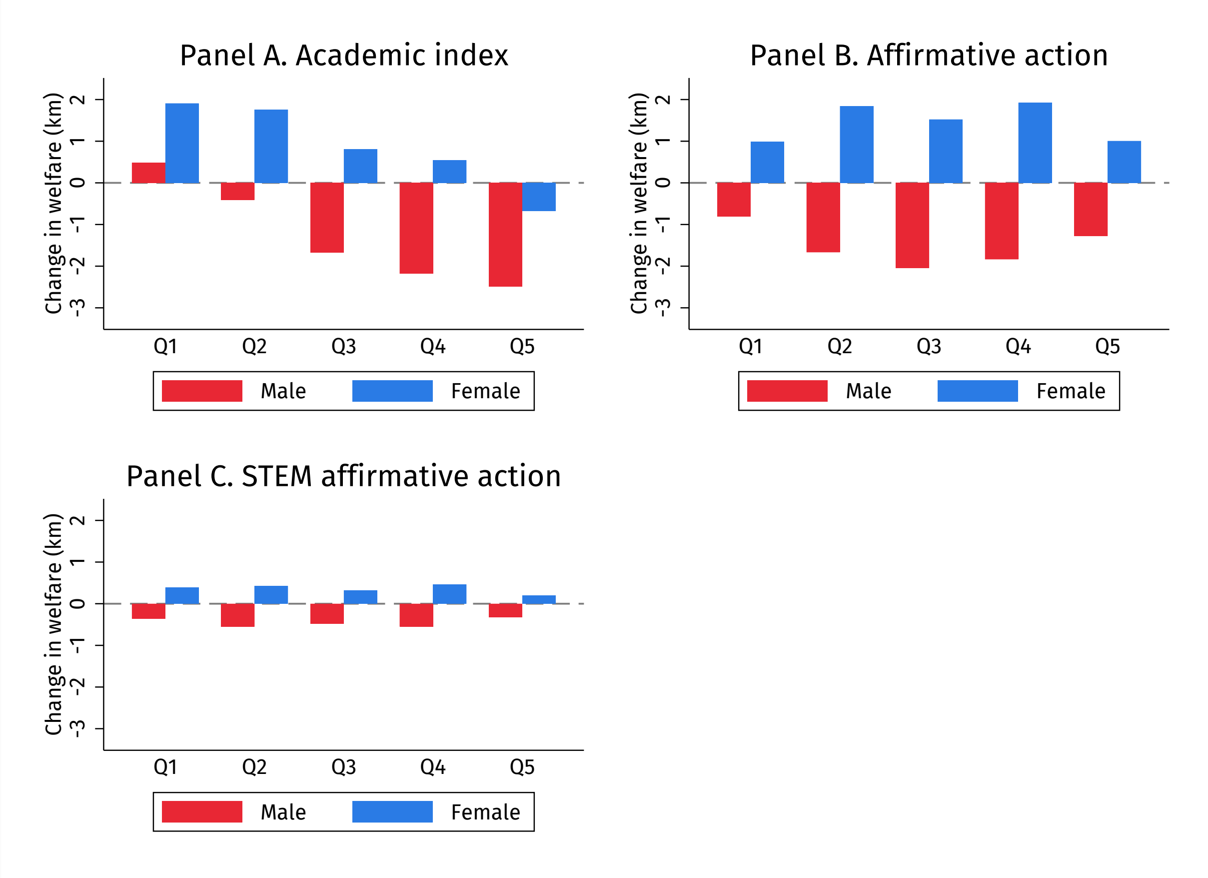
Note: Circles are estimated gender differences in WTT for the specified program type from Equation 4, including (Covariate cell $\setminus$ gender) $\times$ program fixed effects. Lines are 95% confidence intervals. Base category is traditional academic program (non-elite, non-STEM).

Figure 5: Effects of simulated priority structure changes and choice nudges on STEM gender gaps, by exam score quintile



Note: Each panel shows the simulated impact of the indicated priority structure change or preference nudge on the elite and non-elite STEM gender gaps, compared to the gaps in the baseline simulation. Magnitudes are specific to the labeled exam score quintile. Panels B and C refer to 10 point policies.

Figure 6: Effects of simulated priority structure changes on welfare, by exam score quintile



Note: Each panel shows the simulated impact of the indicated priority structure change on mean expected consumer surplus for males and females, compared to the status quo COMIPEMS structure. Magnitudes are specific to the labeled exam score quintile. Panels B and C refer to 10 point policies.

Appendix (not for publication)

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.[18]

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

[18]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.

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 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 mean/(sd)	(2) STEM occupation mean/(sd)	(3) Difference STEM – non-STEM b/(se)
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).

Table A.3: 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 A.4: STEM school assignment summary statistics for 2005, before and after post-computerized assignment phase

	(1) Male	(2) Female	(3) Difference
Computer-based assignment			
STEM assigned school (elite or non-elite)	41.2	29.5	11.7***
	(49.2)	(45.6)	(0.2)
Elite STEM assigned school	12.0	6.2	5.8***
	(32.5)	(24.2)	(0.1)
Non-elite STEM assigned school	29.2	23.3	5.9***
	(45.5)	(42.3)	(0.2)
Elite non-STEM assigned school	16.2	17.2	-1.1***
	(36.8)	(37.8)	(0.2)
Technical non-STEM assigned school	12.4	16.3	-3.9***
	(32.9)	(36.9)	(0.2)
Assigned to any school	86.5	79.7	6.8***
	(34.2)	(40.2)	(0.1)
Finalized assignment			
STEM assigned school (elite or non-elite)	41.4	29.9	11.5***
	(49.2)	(45.8)	(0.2)
Elite STEM assigned school	11.0	5.5	5.5***
	(31.3)	(22.8)	(0.1)
Non-elite STEM assigned school	30.3	24.4	5.9***
	(46.0)	(42.9)	(0.2)
Elite non-STEM assigned school	14.9	15.3	-0.4***
	(35.6)	(36.0)	(0.1)
Technical non-STEM assigned school	12.7	17.4	-4.7***
	(33.3)	(37.9)	(0.1)
Assigned to any school	94.2	90.0	4.2***
	(23.4)	(30.0)	(0.1)
Observations	121805	127694	249499

Note: All variables are indicators, and means are multiplied by 100 to display percentages. Standard errors are in parentheses. STEM assignment indicators are conditional upon assignment to any school. Computerized assignment is student assignment in the serial dictatorship mechanism. Finalized assignment is student assignment after students left unassigned in the computerized assignment phase were able to choose a school that had not filled its quota (or remain unassigned).

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