WIFO WORKING PAPERS
691/2024

Targeting the Gender Placement Gap: Marks versus Money

Pelin Akyol
Kala Krishna
Sergey Lychagin

Targeting the Gender Placement Gap: Marks versus Money

Pelin Akyol (Bilkent University), Kala Krishna (Penn State),
Sergey Lychagin (WIFO)

WIFO Working Papers 691/2024
December 2024

Abstract
Using data from the Turkish University Entrance Exam, we investigate the extent of the gender gap in college placement, its underlying causes, and potential policy interventions. We estimate preferences using a novel approach which improves our ability to capture substitution patterns and find clear evidence that placement differences are primarily driven by preference differences across gender. We compare stipend subsidies to score bonuses that achieve gender parity. Score subsidies improve the welfare of women almost entirely at the cost of men with similar scores and favour high-income women. Stipend subsidies improve the welfare of women, but at little cost to men and favour low-income women. Our work is the first to show that how gender neutrality is achieved matters to society.

E-Mail: sergey.lychagin@wifo.ac.at

2024/1/W/0

Media owner (publisher), producer: Austrian Institute of Economic Research
1030 Vienna, Arsenal, Objekt 20 | Tel. (43 1) 798 26 01 0 | https://www.wifo.ac.at
Place of publishing and production: Vienna

Targeting the Gender Placement Gap: Marks versus Money*

Pelin Akyol † Kala Krishna‡ Sergey Lychagin§

December 13, 2024

Abstract

Using data from the Turkish University Entrance Exam, we investigate the extent of the gender gap in college placement, its underlying causes, and potential policy interventions. We estimate preferences using a novel approach which improves our ability to capture substitution patterns and find clear evidence that placement differences are primarily driven by preference differences across gender. We compare stipend subsidies to score bonuses that achieve gender parity. Score subsidies improve the welfare of women almost entirely at the cost of men with similar scores and favor high-income women. Stipend subsidies improve the welfare of women, but at little cost to men and favor low-income women. Our work is the first to show that how gender neutrality is achieved matters to society.

*Formerly called "Deconstructing the Gender Placement Gap: Performance Versus Preferences". We would like to thank Peter Arcidiacono, Dennis Epple, the seminar participants at Penn State, and the participants of ASSA Annual Meeting 2022, Annual Conference of the International Association for Applied Econometrics 2022, Workshop on Gender and Education at Collegio Carlo Alberto 2022, Human Capital Meeting 2022 and Econometric Society Australasian Meeting 2023 for their valuable comments and suggestions. All remaining errors are our own.

†e61 Institute, Sydney, Australia and Bilkent University, Ankara, Turkey, e-mail:pelin.akyol@e61.in

‡The Pennsylvania State University, NBER, IGC and CES-IFO, e-mail:kmk4@psu.edu

§Corresponding author. WIFO Institute, Vienna, Austria, e-mail:lychagin@gmail.com

1 Introduction

In the United States today, more than 50% of entering law school students are female. In 1958-1959 this number was about 3.1%.[1] Even Ruth Bader Ginsburg, after graduating first in her class from Columbia Law School in 1959, was turned down for a clerkship by Supreme Court Justice Felix Frankfurter because she was female. In economics, among the top 100 US universities, there are more than two men for every woman at the undergraduate level. This ratio is roughly the same at the Ph.D. level, and only 25% of assistant professors and only 13% of full professors are women (Lundberg and Stearns (2019)).[2]

Reducing the gender gap in majors is important, not just for equity reasons but also for efficiency. If intrinsic comparative advantages exist, and women face barriers to entering certain fields, large efficiency losses may ensue.[3] Hence, understanding the drivers of gender differences in the choice of college field choices is essential for designing effective policies, as using different policy instruments can lead to dramatically different consequences for the patterns of winners and losers. A common complaint about policies like affirmative action is that not only do they displace particular groups but they benefit the more advantaged of the disadvantaged rather than the truly deserving disadvantaged. We show that there is reason to expect that using money rather than marks reduces both the negative spillovers to the remaining students and benefits the more disadvantaged.

In this paper, we contribute to this line of inquiry by examining the extent to which gender differences in college placement are driven by performance versus preferences and the implications of this for the effectiveness of different policy interventions. We use data from the Turkish University Entrance Exam to do so. The advantage of the Turkish context is that the mechanism used to allocate students to university programs is extremely centralized and clear-cut. Students list their preferences once they know their scores and are allocated

[1] https://www.americanbar.org/content/dam/aba/administrative/legal_education_and_admissions_to_the_bar /statistics/jd_enrollment_1yr_total_gender.authcheckdam.pdf

[2] The representation of women across the subfields in economics also varies substantially, as measured by papers on the program in the NBER Summer Institute. In finance, the share of female authors is roughly 14.4 %; in macro & international, it is around 16.4 %; and in micro, the share is highest at 25.9 % (see Chari and Goldsmith-Pinkham (2017)).

[3] For instance, Hsieh et al. (2019) argue that between 20% and 40% of growth in aggregate market output per person from 1960 to 2010 can be explained by the improved allocation of talent.

to their most preferred choice with priority determined based on their score. Competition is fierce and applicants face considerable stress as a result. There is also significant gender bias in placement. For example, over 76% of students who major in engineering are male.

Our analysis is based on detailed administrative data from the exam, which provides rich student-level information on background characteristics, preferences, performance, and admission to college programs. This allows us to take a model-based data-driven approach needed to investigate the underlying factors contributing to these gender disparities.

We begin our analysis by exploring three potential drivers of the gender gap in placement: differences in entrance exam scores, differences in preferences, and more conservative application behavior among women.[4] Since college seats in Turkey are allocated based on placement scores, a gender gap in scores could explain the gender gap in placements, particularly in competitive majors. We do find evidence of such a gap. This gap may result from fewer resources being invested in girls' education,[5] or from female students underperforming in high-stakes environments.[6]

A gender gap in placement could also result from differences in preferences. We find that preferences vary considerably by gender: for instance, engineering and technical fields attract few female applicants, even after controlling for entrance exam scores. Women may avoid highly competitive STEM programs, or they may have different preferences for social or cultural reasons, including considerations related to the marriage market.[7] Certain fields may be viewed as culturally inappropriate for women (e.g., veterinary science) or as more attractive due to being low-pressure and family-friendly (e.g., education), even if they are lower-paying.[8]

[4]Another channel, documented by Saygin (2016), is the tendency of male students to retake the entrance exam, potentially improving their scores and increasing their chances of being placed in competitive programs. While we account for gender differences in retaking behavior in our preference estimation (see Section 4), this is not our primary focus here.

[5]While we observe gender gaps in performance, we do not find any evidence of under-investment by parents or schools in preparing female high school students for the college admissions process.

[6]Taylor (2019) argues that high-stakes testing in admissions to New York City's elite public high schools disadvantages women. Azmat et al. (2016) show that in high-stakes environments, women's performance relative to men's worsens. Arenas and Calsamiglia (2020) show that increased exam stakes in Spain negatively affected female performance, particularly in highest-stakes exams.

[7]For example, see Kirkebøen et al. (2021) and Arum et al. (2008).

[8]Table 12 in the Appendix shows that incomes and employment probabilities for women are lower in the fields they sort into. For example, teacher training and education pay about 1280 Turkish Lira (TL) for

Finally, gender gaps in placement may arise from women applying more conservatively than men, conditional on exam scores. It has been suggested that women are more risk-averse and less competitive in various contexts.[9] In Section 3.3, we show that the data do not support this hypothesis. While the difference between placement scores and the cutoff scores for their chosen programs tends to be higher for women, once we control for major, women do not appear to be aiming lower than men. This suggests that the observed gap is due to differences in preferences, not competition aversion, so we do not consider this channel in our analysis.

We then quantify the importance of preferences versus performance. At the core of this exercise is the estimation of a model that captures student preferences for college majors and how these preferences translate into actual placements. We propose an approach that balances the simplicity of using placement data with the recognition of unobserved preference heterogeneity among students. This allows us to identify systematic differences in preferences that would be difficult to detect using only realized placements. As detailed in Section 4, one could estimate preferences using the entire list of programs submitted by students (Larroucau and Rios (2020) is one recent example of this approach). We argue that economically relevant alternatives should be the focus, as not all choices are equally relevant when the list size and the number of alternatives are large.[10] Following Berry et al. (2004), we emphasize that since students are informed of last year's cutoffs, their rankings of college programs should reflect feasible program choices based on the previous year's or realized cutoffs. This approach prioritizes quality over quantity in the data. We show that it better captures substitution patterns in the data compared to methods that only use realized choices, but remains computationally inexpensive to model markets with thousands of available choices.

After estimating preference parameters, we compare policies aimed at reducing the gender gap and simulate their effects on college placement outcomes. First, we eliminate the gender score gap by awarding bonus points equal to the average gender score gap to all fe-

women aged 25-30, rising to 1570 at age 40-50, whereas engineering pays 1420 at a younger age and 2050 at older ages.

[9]Niederle and Vesterlund (2011), Niederle and Vesterlund (2007), and Eckel and Grossman (2008) provide examples, and Saygin (2016) discusses this in the Turkish context.

[10]Students can include up to 24 programs from a list of over 7,000 alternatives for students with high scores.

male students. Surprisingly, this does little to increase female representation in engineering, a traditionally male-dominated field. The reason is that women's preferences differ substantially from men's, so additional points raise cutoffs in fields favored by women without reallocating them to engineering. Next, we eliminate the preference gap by giving women the same preferences as men. This significantly increases female representation in engineering, though it does not completely close the gap. Our findings align with those of Arcidiacono (2004), who also find that preferences play a critical role in student major choices.

Finally, we perform a novel counterfactual experiment: we compare two policies that reduce gender bias in engineering placements. One policy uses stipends and the other provides score subsidies for women entering engineering. While both policies reduce gender bias, they have starkly different outcomes. Score subsidies primarily benefit high-income women and come at a cost to men with similar scores, while stipend subsidies favor low-income women and have little adverse impact on men. The absence of an adverse impact on men would likely make this policy more palatable. Overall, our message is that *how* gender neutrality is achieved matters for societal outcomes.

Unlike studies focused on North America, our research benefits from the transparent and rigid nature of the Turkish college admission system. The clear-cut allocation mechanism, program seat quotas, and well-documented student priorities enable us to simulate placement outcomes as a market equilibrium. This would be much harder using the U.S. data, where college admissions processes, especially at elite schools, are far from transparent.

Our work connects well with the literature, while having several novel features we highlight as we proceed. Ngo and Dustan (2021) look at the STEM gender gap in Mexican high schools and decompose the contribution of preference differences and marks differences across genders in accounting for this gap. This is perhaps closest to our work. They estimate preferences based on Fack et al. (2019), which sets the actual placement of a student to be the most preferred one in his feasible set and allows for observed heterogeneity. Our approach is more novel, as we not only incorporate unobserved heterogeneity into our estimation, but also build on Berry (1994), which significantly improves the ability of the estimated model to match data patterns. They do counterfactuals that show that preference differences drive the gender gap overall, but this is less so for elite schools. Giving women male preferences would reverse the gender gap while giving women male scores would reduce it only by about 25%. We also find that preferences are more important than performance, but far less so in our setting.

Our focus is not just on the role of preferences versus marks driving the gender gap as in the literature but also on the implementation of different policies for closing the gender gap. We look at the difference between offering point bonuses and offering money and show that their effects vary by ability and income. By doing so, our paper contributes to the affirmative action literature. In particular, it connects to the literature that compares race-based affirmative action to socioeconomic class-based affirmative action (see, for example, Cestau et al. (2017)). However, there is far less work comparing affirmative action using point bonuses to affirmative action using stipend bonuses. There is one exception we are aware of. Arcidiacono (2005) shows that removing advantages in admissions substantially decreases the number of black students at top-tier schools, while removing advantages in financial aid causes a decrease in the number of blacks attending college overall.

Understanding who benefits and who loses from the implementation of affirmation action policies has important policy implications. When considering affirmative action based on race, caste, or ethnicity, it is often argued that such programs benefit the more advantaged groups rather than the disadvantaged, which runs counter to the rationale behind the programs and creates opposition to them. Our results show that the stipend bonus results in significant welfare gains for women, particularly at the upper end of the score distribution, and while men experience some welfare losses under this policy, these losses are small. This makes the stipend bonus a "win-win" policy. In contrast, the score bonus increases welfare for women but reduces men's welfare by similar amounts, especially for higher-scoring individuals. Both policies have the greatest effect on high-scoring students, as they are more likely to apply to engineering programs, whereas low-scoring students remain less affected.

Additionally, we find that low-income women gain the most from the stipend policy, while losses for men are generally smaller and concentrated among high-income males. The score bonus, however, favors high-income women more and disproportionately harms high-income men. Overall, the stipend bonus is not only more effective in fostering gender equality, but also has a redistributive impact, benefiting lower-income women and imposing smaller costs on men.

A more tangentially related line of work focuses on specific mechanisms generating gender gaps. Wiswall and Zafar (2021), Stinebrickner and Stinebrickner (2014), and Arcidiacono et al. (2020) investigate the role of subjective expectations in preference for majors. Carrell et al. (2010) argue that hysteresis may play a role since women are more likely to take STEM courses if female professors teach their introductory courses in these areas. A hostile environment for females in the field could be another reason.[11] Several studies attribute gender gaps in placement to student performance in placement tests (Turner and Bowen (1999)), and to early tracking and the choice of advanced courses in high school (Card and Payne (2021)). See Kahn and Ginther (2018) for more on studies on this topic.

The remainder of the paper is organized as follows. The next section provides an overview of the data and the institutional background of the university entrance system in Turkey. In Section 3, we present reduced-form evidence on gender gaps in exam scores, college preferences, and competition aversion. Section 4 explains how our approach to estimating preferences fits into the literature and show that it indeed better captures substitution patterns in the data compared to alternative methods. In Section 5, we disentangle the impacts of preference and performance gaps on placements and use the model to evaluate counterfactual policies aimed at achieving gender balance. Finally, Section 6 concludes.

2 The Turkish Setting

In Turkey, a year after students start high school, they choose one of four academic tracks: Science, Turkish-Math, Social Studies, or Language.[12] In each track, students study a different curriculum. In their senior year, they take the centralized University Entrance Exam, which is conducted by the Student Selection and Placement Center (ÖSYM) once a year. Students' track, GPA, and score in the exam determine their placement scores. Both high school seniors and past high school graduates can take the exam, and almost every high school senior chooses to do so. Students are free to repeat the exam, but the score obtained

[11]See Wu (2018) and Wu (2020) for more on this.

[12]We only consider students from the first three tracks in this paper as students in the language track have to take additional exams and so can be considered a distinct market.

in a year can be used only in that year.

In 2002, this exam included tests in four subjects: Turkish, Social Science, Math, and Science. Students' scores are calculated as a weighted average of their standardized raw scores on each test. For each student, three different scores, Quantitative (ÖSS-SAY), Turkish-Math (ÖSS-EA) and Social Science (ÖSS-SÖZ) are calculated. Each score puts more weight on subjects considered relevant. The high school grade point average (GPA) is added to each ÖSS score with a weight to form the respective placement score (Y-ÖSS-SAY, Y-ÖSS-EA, Y-ÖSS-SÖZ); the placement scores are the *only* determinants of college admission. Each university program is associated with a specific major and uses the relevant placement score out of these three to rank students for admission. Thus, if a student from the Science track applies for engineering, their Y-ÖSS-SAY score would be used, while if they apply for economics, their Y-ÖSS-EA score would be used. Note also that the track chosen in high school matters for calculating the placement scores: two students with the same raw ÖSS scores and the same weighted GPA but in different tracks would get different placement scores as the weights are designed to keep students in their tracks in college.[13]

After the exam, students are informed of their scores. Students who get at least 120 points in a score type are eligible to submit preferences for all college programs that admit students based on that type of score. Students whose scores are between 105 and 120 are only allowed to submit preferences for 2-year college programs and distance education programs. Students can submit up to 24 preferences, and at most 18 of these can be for 4-year or 2-year programs. Upon submitting their preferences, students are ranked and placed following the multi-category serial dictatorship mechanism.[14] At the time of submitting preferences, students in the period of 1999-2003 had a fairly good idea of what their feasible sets were as the cutoff admission scores in most programs have been relatively stable[15] and each student received a booklet with every program's cutoff from the past year.[16]

[13]See Krishna et al. (2018) for details of this process.

[14]Balinski and Sönmez (1999) describe the mechanism in detail and show that it is equivalent to the Gale-Shapley college-optimal mechanism.

[15]These admission cutoffs for programs are depicted in Figure 9. On the vertical axis are the cutoffs in 2000 and 2001, while the cutoff in 2002, the year of our data, is on the horizontal axis. As is evident, the cutoffs tend to lie on the 45-degree line. The clustering around the 45 degree line is tighter for 2001 than for 2000. This would be expected: the farther back in time we go, the more things would have changed.

[16]Booklets for previous years are also easily available.

Students face fierce competition, especially at the top. For example, the highest-ranked engineering program had a cutoff of 223 (out of a maximum of 224), while the next highest one had 221 points. Consistent with this, Krishna et al. (2018) shows that utility increases steeply with scores at the top of the score distribution. Around 1.5 million students took the University Entrance Exam in 2002 and only one-third of these are offered a place in a university program. In Turkey, most universities are public as are many of the very best ones. Tuition fees in public universities tend to be very low, though private universities offer scholarships that reduce or remove fees. These scholarships are program-specific,[17] and are merit-based, rather than need-based in contrast to the norm in the U.S.

2.1 Data

The data used in this study come from multiple sources. The main source is administrative data on a random sample of 2002 University Entrance Exam participants. It includes data on performance and submitted preferences as well as background information, in particular, raw test scores in each test, weighted ÖSS test scores, high school ID, track, high school GPA, gender, family background, ranked preference list, and the program they are assigned, if any. The high school GPA (AOBP in Turkish) is scaled by the authorities to account for grade inflation; it is not directly available, but we can recover it from the data. Details of this process are explained in the Appendix E. We have a random sample of about 40,000 students from each track (Social Science, Turkish-Math, Science), including both first-time and repeat takers.

The second source of data is the booklet published by ÖSYM, which includes the minimum cutoff scores and the number of available seats in each college program for the years 2000, 2001, and 2002. We also observe the tuition cost of each department, and the amount of the scholarship, if provided.[18] In addition, we collected the distance between each pair of districts in Turkey from the General Directorate of Highways.

Summary statistics on first time exam takers are presented in Table 1 for each track and

[17] Admission is to a program in a university, as well as the scholarship offered, and not to the university more broadly. Consequently, cutoffs vary by scholarship level, even when the program and university are identical.

[18] Tuition costs in public universities do not vary across universities, but they vary according to the major.

gender. Columns 1 and 2 present the means and standard deviations of each variable. Column 3 presents the difference between females and males. The same statistics are presented in columns 4 to 6 for Turkish-Math track students and in columns 7 to 9 for Social Science track students. Note that the gender gap in the ÖSS-SAY score is the largest among Science track students. The ÖSS-SAY score of female students is 4.2 points lower than that of male students. However, female students' normalized high school GPA is 3.2 points higher than that of males, which closes the part of the Y-ÖSS (placement score) gap between males and females.

The second group of variables presented have to do with prep school expenditures. These expenditures can be missing, zero, low (less than one billion TL), medium (one to two billion TL), and high (more than two billion TL).[19] For each level of expenditure, the table gives the fraction of that gender in this expenditure group. It is evident that women are less present in the low-expenditure groups and more present in the higher-expenditure groups, especially when they are in the Science track. Thus, gender bias in terms of prep school expenditure is unlikely to be the reason behind the worse performance of women in the university entrance exam. The next row gives the proportion by gender that obtained a scholarship for prep school.[20] Somewhat surprisingly, males are significantly more likely to obtain scholarships in the Science track. The difference is there, but small and not significant in other tracks.[21]

The third group of variables deals with parental education. Again, the numbers give the proportion by gender by parental education. The numbers suggest that women whose parents are more literate are more likely to apply to college as expected. The fourth group of variables deals with parental income. The numbers suggest that women taking the university entrance exam are less likely to come from poorer families. This reflects the fact that women from poorer and more conservative households do not end up finishing high school. The next group of variables deals with the type of school the students go to. Note that women are

[19] Turkey had hyperinflation up till 2004, after which the old TL was replaced with the new TL where 1 million old TL were converted to one new TL. In 2004, two billion TL would have been about 1500 US dollars.

[20] Each prep school in Turkey has an exam taken in the 11th grade to obtain a merit-based scholarship. This serves the prep schools as they advertise the performance of their students to attract customers.

[21] This is probably because non-Science track students are very unlikely to get scholarships to begin with.

not less likely to go to science high schools,[22] conditional on finishing high school, but are less likely to go to private schools if they are in the Science and Turkish Math Tracks. This might be because science high schools are free, even though they are fiercely competitive. Fellowships to cover expenses are also available on a competitive basis. Private high schools are expensive, and there are very few scholarships offered. The last variable is the fraction that comes from the east of Turkey which is seen as being poorer and more conservative than the western part. As expected, the fraction of female from the east is significantly less than the fraction of male in all tracks. The difference is the smallest (5.4%) for the Science track and largest for the Social Science track (10.4%).

Table 1: Descriptive Statistics

	Science Track			Turkish-Math Track			Social Science Track		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Female	Male		Female	Male		Female	Male	
	Mean	Mean	(1)-(2)	Mean	Mean	(4)-(5)	Mean	Mean	(7)-(8)
VARIABLES	(SD)	(SD)	Diff.	(SD)	(SD)	Diff.	(SD)	(SD)	Diff.
ÖSS-SAY	134.379	138.586	-4.206***	111.794	112.842	-1.049***	102.332	102.879	-0.547***
	(20.493)	(21.216)		(12.138)	(11.951)		(4.981)	(5.003)	
ÖSS-EA	127.296	126.987	0.309	119.703	119.479	0.224	110.508	110.405	0.104
	(16.903)	(18.656)		(12.536)	(12.538)		(7.474)	(7.628)	
ÖSS-SÖZ	118.607	117.353	1.254***	126.306	125.681	0.625*	119.845	121.051	-1.205***
	(18.404)	(21.535)		(13.572)	(14.077)		(10.936)	(11.537)	
GPA (OBP)	55.211	52.005	3.206***	52.758	48.518	4.240***	50.809	48.497	2.312***
	(9.329)	(10.115)		(8.784)	(9.113)		(8.061)	(8.003)	
Prep School Expenditure:									
Missing	0.068	0.078	-0.010*	0.142	0.144	-0.002	0.286	0.276	0.010
	(0.251)	(0.268)		(0.349)	(0.351)		(0.452)	(0.447)	
No prep school	0.075	0.089	-0.014**	0.169	0.159	0.010	0.296	0.292	0.004
	(0.263)	(0.285)		(0.375)	(0.366)		(0.456)	(0.455)	
Low	0.419	0.439	-0.021*	0.375	0.425	-0.050***	0.275	0.307	-0.032*
	(0.493)	(0.496)		(0.484)	(0.494)		(0.446)	(0.461)	
Medium	0.279	0.235	0.044***	0.210	0.180	0.030***	0.106	0.089	0.017
	(0.448)	(0.424)		(0.407)	(0.384)		(0.308)	(0.285)	
High	0.116	0.102	0.014**	0.081	0.074	0.007	0.024	0.021	0.003
	(0.320)	(0.302)		(0.273)	(0.261)		(0.152)	(0.142)	
Scholarship	0.044	0.057	-0.013***	0.023	0.019	0.005	0.014	0.015	-0.001
	(0.205)	(0.232)		(0.151)	(0.135)		(0.118)	(0.122)	
Highest Parental Education:									
Missing	0.072	0.052	0.020***	0.069	0.050	0.019***	0.065	0.040	0.025***
	(0.259)	(0.222)		(0.254)	(0.218)		(0.246)	(0.195)	
Literate	0.034	0.062	-0.027***	0.042	0.090	-0.048***	0.058	0.112	-0.054***
	(0.182)	(0.241)		(0.200)	(0.286)		(0.234)	(0.316)	
Primary School	0.237	0.256	-0.019*	0.317	0.330	-0.013	0.437	0.445	-0.008
	(0.425)	(0.437)		(0.466)	(0.470)		(0.496)	(0.497)	

(continued on next page)

[22] All students in science high schools are from the Science track, therefore the entries are blank in other tracks.

VARIABLES	Science Track (1) Female Mean (SD)	(2) Male Mean (SD)	(3) (1)-(2) Diff.	Turkish-Math Track (4) Female Mean (SD)	(5) Male Mean (SD)	(6) (4)-(5) Diff.	Social Science Track (7) Female Mean (SD)	(8) Male Mean (SD)	(9) (7)-(8) Diff.
Middle/high School	0.333	0.315	0.018*	0.367	0.342	0.025**	0.345	0.313	0.032*
	(0.471)	(0.465)		(0.482)	(0.475)		(0.475)	(0.464)	
College	0.324	0.315	0.009	0.205	0.188	0.017*	0.095	0.090	0.005
	(0.468)	(0.465)		(0.403)	(0.391)		(0.293)	(0.286)	
Income:									
Less than 250 TL	0.260	0.283	-0.023**	0.328	0.362	-0.034***	0.407	0.464	-0.058***
	(0.439)	(0.451)		(0.470)	(0.481)		(0.491)	(0.499)	
250-500 TL	0.422	0.414	0.008	0.427	0.396	0.030***	0.426	0.375	0.051***
	(0.494)	(0.493)		(0.495)	(0.489)		(0.495)	(0.484)	
More than 500 TL	0.318	0.303	0.015	0.245	0.241	0.004	0.167	0.161	0.006
	(0.466)	(0.460)		(0.430)	(0.428)		(0.373)	(0.367)	
Type of the High School:									
Science school	0.024	0.026	-0.002	0.000	0.000	0.000	0.000	0.000	0.000
	(0.155)	(0.160)		(0.000)	(0.000)		(0.000)	(0.000)	
Anatolian school	0.338	0.339	-0.001	0.196	0.235	-0.039***	0.052	0.054	-0.002
	(0.473)	(0.474)		(0.397)	(0.424)		(0.223)	(0.227)	
Private school	0.052	0.069	-0.017***	0.040	0.052	-0.012**	0.020	0.019	0.001
	(0.222)	(0.253)		(0.196)	(0.222)		(0.140)	(0.137)	
Home region:									
Eastern region	0.212	0.266	-0.054***	0.238	0.307	-0.069***	0.221	0.325	-0.104***
	(0.409)	(0.442)		(0.426)	(0.461)		(0.415)	(0.468)	
Observations	5720	7785		6681	5983		2196	2569	

* p<0.1 ** p<0.05 *** p<0.01.

3 Direct Evidence on Gender Gaps

In this section, we examine different channels that can lead to gender gap in college major choice. We start by focusing on the gender gap in performance and preferences and present direct evidence of this disparity. It has also been suggested that women tend to do worse in placements because they are less aggressive in applying (see Saygin (2016)). We examine this channel and find no such evidence once we control for the broad field of application.

3.1 Do Women Do Worse in the Entrance Exam?

As we presented in the previous section, women perform worse on the exam in the science subjects than men. This is more so in the Science track. Note, however, that women do better in high school than men: the mean GPA for women is significantly higher than that for men as reported in Table 1.[23]

[23] We present the full distributions of scores and GPA for men and women in Appendix F.

There are many explanations for the gender gap in such high-stakes exams. The primary one seems to be that women perform worse under pressure than men, and/or that women do worse in high-stakes multiple-choice exams because they tend to not guess when it would be optimal for them to guess. Akyol et al. (2022), using the same data we use, show that women do seem to be more risk averse than men. Ors et al. (2013) show that men outperform women in a high-stakes exam for admission to an elite MBA. Gneezy et al. (2003) show in an experimental setting that women seem to perform worse than men in competitive environments, and more so as competition rises, especially when competing with men. Niederle and Vesterlund (2007) in addition show that in experiments, men choose a tournament compensation system over a non-competitive piece rate system much more often than women. They argue that this difference is driven by men being more overconfident, so "women shy away from competition, while men embrace it".

The raw difference in scores suggests a significant gender gap; however, this disparity may arise from various factors. One factor is the documented trend in the period of interest, where females were less likely to enroll in high schools compared to males. This trend could cause selection issues and elevate the average performance of female students relative to males. Therefore, it is crucial to control for background variables and proxies for ability. To address this, we run the following regression

$$OSS_{ij} = \alpha_j MALE_i + \beta_j X_{ij} + e_{ij} \tag{1}$$

where i indexes the student and j indexes the track of the student. For each student, we only use the track-specific aggregate score (ÖSS-SAY for the Science track, ÖSS-EA for the Turkish-Math track, and ÖSS-SÖZ for those in the Social Science track).

The individual-level controls, represented by X_{ij}, include factors such as the parental income and education, the normalized high school GPA, school fixed effects, and expenses on preparatory courses. We also control for high school specialization by estimating the above regression independently for each high school track. By doing so, we can account for potential explanations related to parental investment and high school choice, preparation for the entrance exam, as well as for learning while in high school.[24] Overall, these controls help to ensure that our analysis accurately captures the relationship between gender and exam performance while accounting for various other factors that may affect the results.

The estimates are reported in Table 2. We progressively include controls to check if the gender gap is driven by parental under-investment or selection based on parental education and income. The gender gap estimates do not change by much, which suggests that the above channels are not driving the difference in scores. The size of the gap does vary by track when measured in points, but once the estimates are scaled by the standard deviation, the difference is much smaller.[25]

3.2 Are Women's Preferences Different?

In addition to the gender gap in scores, we find strong evidence that preferences differ by gender as well. Figure 1 presents the percentage of female and male students in each major according to placement.[26] As is evident, there are large differences in the share of women: at one extreme, 76.3% of students who are placed in an engineering program are male, on the other, 6.6% of students placed in a health service major (which includes nursing, midwifery, and health-related social work) are male. Social and behavioral science majors are female-dominated being 75.7% female, while technical science, technical services and veterinary medicine are male-dominated with a 60.9, 85.3 and 83.7 % male share.

These patterns in placements could arise from the difference in scores. For example, women may be underrepresented in engineering programs just because their scores are lower and engineering is a competitive field. For this reason, we look directly at the preference lists while controlling for scores and track. Figure 2 shows the fraction of students in the Science track who put the major as their first preference as a function of the relevant placement

[24] Since we have school-fixed effects, it will make no difference whether we use the normalized or plain high school GPA.

[25] The standard deviations are around 20 points in ÖSS-SAY for the Science track, 12 points in ÖSS-EA for the Turkish-Math track, and 11 points in ÖSS-SÖZ for the Social Studies track. See Table 1 for more details.

[26] Figures 10, 11, 12 in the Appendix present the percentage of male and female students in each college major for each of the three tracks separately.

Table 2: Gender Gap in ÖSS Score

	(1)	(2)	(3)	(4)
	Science Track			
VARIABLES	ÖSS-SAY Score			
Male	8.909***	9.077***	9.326***	9.188***
	(0.328)	(0.317)	(0.236)	(0.230)
Observations	13,505	13,505	13,505	13,505
	Turkish Math Track			
VARIABLES	ÖSS-EA Score			
Male	2.989***	3.159***	3.747***	3.571***
	(0.228)	(0.218)	(0.160)	(0.154)
Observations	12,664	12,664	12,664	12,664
	Social Science Track			
VARIABLES	ÖSS-SÖZ Score			
Male	2.612***	2.745***	4.060***	3.795***
	(0.343)	(0.334)	(0.318)	(0.308)
Observations	4,764	4,764	4,764	4,764
Controls:				
Prep School Expenses	No	No	No	Yes
High School Fixed Effects	No	No	Yes	Yes
Parental education	No	No	Yes	Yes
Income	No	Yes	Yes	Yes
High School GPA	Yes	Yes	Yes	Yes

Standard errors are clustered at the school level. *** p<0.01, ** p<0.05, * p<0.1

Figure 1: Gender Differences in Major Choice (All Tracks)

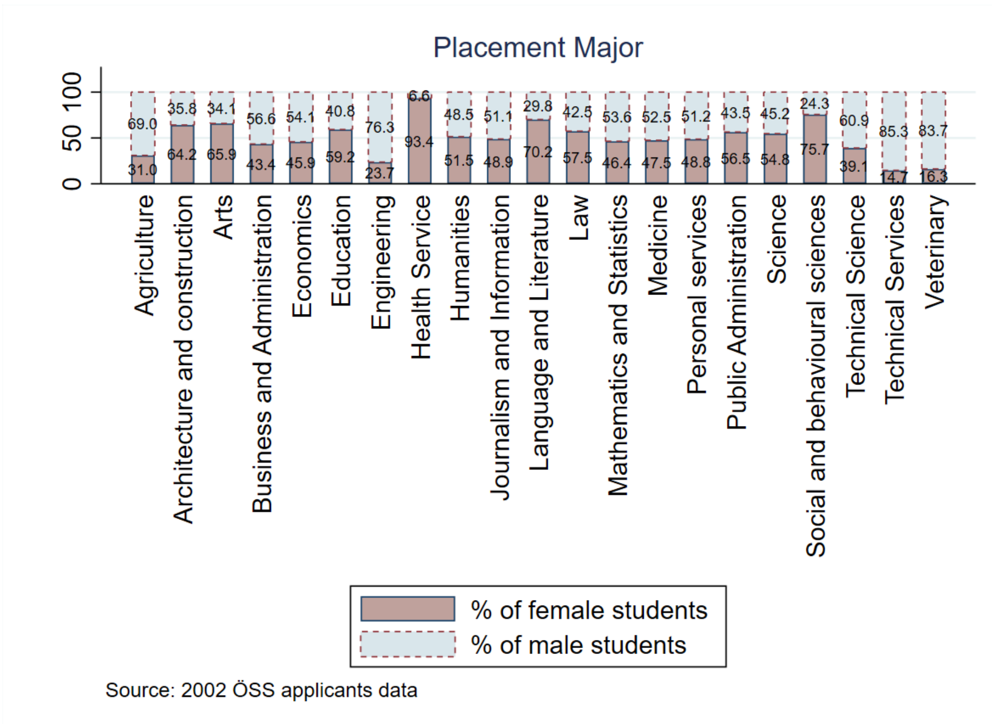
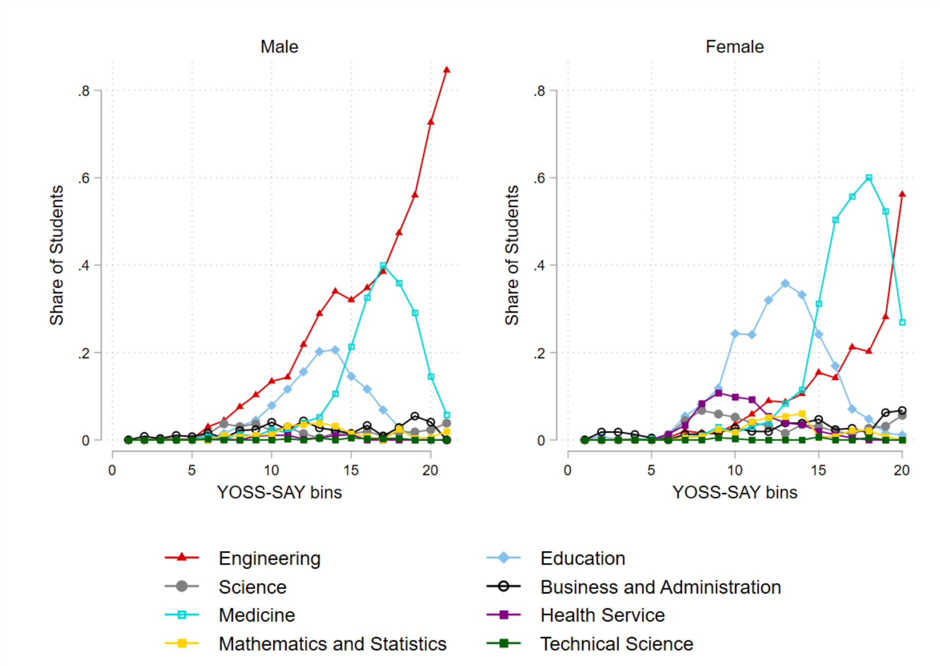


Figure 2: 1^{st} Preference Major (Science Track)



score, separately for each gender.[27] [28] In almost all score bins, male students are more likely than female ones to be placed in engineering programs and put engineering programs first on their list. Moreover, the preferences (and placement) of female students vary much more with their scores: while women with high scores are more likely to apply and be placed in engineering programs, those in the middle of the distribution seem to prefer education, while those with even lower scores seem to prefer health service. In addition, women are more likely than men to apply for medicine at all scores. In contrast, the preference for engineering falls much more slowly with rank for men. This pattern is the result of systematic differences in the preferences of female and male students.

3.3 Are Women Less Aggressive in Applying?

Work using similar data from Turkey suggests that women are less aggressive in applying than men (Saygin (2016)). We find reason to question this conclusion. If women aim lower than men, then the gap between the student's placement score and the cutoff for the program of placement should be larger for women. We show that while this difference is negative and significant, once we account for the majors students are placed in, women do *not* seem to aim lower than men. In other words, women tend to apply to majors where there is a larger dispersion of scores among students, rather than being less aggressive in their applications.

We run the difference in the student's placement score and the cutoff score in 2001 (for the student's program of placement) on the male dummy and the background controls and present the results in Table 3. The specification in Column 1 does not include any controls. This gives a negative and significant coefficient on the male dummy of -0.58, which suggests that males on average are more aggressive in their applications. In Column 2, we add province fixed effects based on the location of the high school the student attended. This makes the coefficient slightly more negative. In Column 3, we add more individual background controls including prep school expenses, income, and parental education level. This has almost no effect on the mean gap. Finally, we add controls for the major in which the student was

[27]We construct score bins of width 5 starting from 120.

[28]The same graphs for the Social Studies and Turkish-Math track students are presented in Figures 13 and 14 in the Appendix.

placed. The effect of this is startling. First, the male dummy that we have been focusing on becomes insignificant and, if anything, slightly positive, suggesting that males, on average, are *less* aggressive in their applications. Second, the major dummies for health service, technical science, science, and veterinary science are positive and significant, indicating that students applying to these majors tend to be less aggressive. In other words, these majors have a longer right tail in terms of the placement score distribution of applicants. Thus, the negative mean gap, we obtain in Columns 1-3, seems to be coming from a composition effect. If women apply to majors where the average difference in score and the cutoff score is large, it looks as if men are applying more aggressively than women if we do not control as we do in Column 4. This suggests that the difference in placement score and cutoff between men and women we thought we had identified in Columns 1-3 comes from a compositional effect.[29]

4 Modeling of College Preferences

To deconstruct the gap in placements into what comes from performance and preferences and to study policies aimed at closing this gap, we set up and estimate a model of demand for college seats. There has been considerable progress made in estimating preferences over schools in recent years and it is important to place our approach within this literature.

The identification challenge central to this literature is that student preference lists cannot be interpreted as their true preferences. This is obviously an issue in the settings where applicants can benefit from misrepresenting their preferences, for example, in the Boston mechanism (see He (2017) and Calsamiglia et al. (2020)). Even for strategy-proof mechanisms, there is ample evidence that agents routinely make mistakes and omissions in their submitted lists.[30]

The literature offers two ways to accommodate misreporting. One is to predict and match preference lists in the data assuming that students perfectly maximize their payoffs;

[29] We also ran the regression including interactions of the male dummy and the major dummies. This did not affect our conclusion, nor were any of these interactions significant.

[30] Prevalence of reporting mistakes is demonstrated, in particular, in Hassidim et al. (2017), Hassidim et al. (2021), Rees-Jones (2018), Shorrer and Sóvágó (2023) and Artemov et al. (2017).

Table 3: Factors Affecting the Difference Between Y-ÖSS Score and Admission Cutoff (Science Track)

VARIABLES	(1)	(2)	(3)	(4)
Male	-0.580***	-0.684***	-0.676***	0.017
	(0.167)	(0.172)	(0.168)	(0.220)
Subject of Major				
Architecture and construction				-0.606
				(1.093)
Education				0.360
				(0.943)
Engineering				-0.822
				(0.926)
Health Service				2.816**
				(0.838)
Mathematics and Statistics				-0.759
				(0.987)
Medicine				0.855
				(0.864)
Science				2.572*
				(1.107)
Technical Science				12.461***
				(2.426)
Technical Services				1.514
				(0.902)
Veterinary				3.602*
				(1.409)
Observations	3,878	3,878	3,878	3,878
High School City FE	NO	YES	YES	YES
Income, prep school expenses, parental education FE	NO	NO	YES	YES

Standard errors are clustered at the school level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Agarwal and Somaini (2018), Calsamiglia et al. (2020) and Larroucau and Rios (2020) follow this route. Another approach is to only trust data on student placements, as some parts of preference lists may be prone to optimization mistakes, especially those containing programs with low probability of placement (Fack et al. (2019); Artemov et al. (2017)). Naturally, choosing between these two approaches involves trade-offs. Using full preference lists rather than just the program of placement extracts more information from the data. However, the latter approach is more robust to reporting mistakes. Using full lists may also run into prohibitive computational costs if the number of choices is high, although recent papers made progress in easing this constraint (Larroucau and Rios (2020)).

In this paper, we propose a middle-ground approach. On the one hand, the placement-only approach is attractive, since we deal with a large choice set[31] and preference lists in our data suggest manipulation.[32] On the other hand, students in Turkey are likely to have rich unobserved preference heterogeneity, which is evident from their preference lists but is hard to pin down if one relies only on the program of placement. For example, some students may only apply to programs in engineering and mathematics, while others apply only to medical programs, suggesting that students have systematic differences in their preferred majors. It is well-documented in the empirical industrial organization literature that identifying complex substitution patterns is extremely difficult if one only uses realized choices (Berry et al. (2004)).

Our estimation method is based on two sources of identification. We extend the asymptotic ex-post stability approach of Fack et al. (2019) who show that for economies with many students the uncertainty in admission cutoffs is negligible, which means that students can predict their feasible sets of programs with near certainty when preferences are stated.[33] As a result, in the limit, students are placed in their most preferred feasible option given the realized cutoffs. Just as in Fack et al. (2019), we use the program of actual placement to identify

[31] In our data, students choose 24 programs out of roughly 3,500. For comparison, the numbers are 3 out of 13 in Agarwal and Somaini (2018), 10 out of more than 300 in Calsamiglia et al. (2020), and 10 out of 1,400 in Larroucau and Rios (2020).

[32] For instance, students rarely list programs with cutoffs far from their own placement scores, even when doing so is costless.

[33] This assumption makes sense in Turkey as there are hundreds of thousands of students taking the university entrance exam, the system was stable in the period we consider and preferences are stated after the student knows his score.

student preferences. The second source of identification aims to give us a better handle on substitution patterns in students' true preferences if the cutoffs slightly change relative to their 2002 values.[34] We posit a question: what program would each student choose, had the cutoffs in 2002 been the same as in 2001? When each student strategizes over his preference list, the cutoffs for 2001 are available to him in the application materials. It is only natural that these cutoffs are used to evaluate any candidate preference list to make sure the list selects the most preferred program among those feasible. The hypothetical program of placement under the cutoffs from 2001 and the student's preference list from 2002 serves as our second source of identification. We show that our approach captures substitution patterns in the data better than other feasible approaches, including that of Fack et al. (2019).

Our contribution on the methodological front is thus a strategy to identify demand for college seats using information on substitution patterns from student preference lists in a simple and easily implementable manner. We show that our approach does much better at reproducing the substitution patterns found in the data than the placement-based estimator in Fack et al. (2019) and thus yields considerable rewards at low computational cost. For example, when we use placement outcomes only, the model predicts excessive switching of majors relative to what we would expect based on preference lists. This suggests that the additional information on preferences we use is instrumental in capturing the substitution patterns correctly.[35]

4.1 Notation and Identifying Assumptions

Our method is based on three identifying assumptions. First, we follow Fack et al. (2019) and assume that observed placements are asymptotically ex-post stable. This means that the placements observed in 2002 are optimal under the realized admission cutoffs for all students except for a vanishingly small share. Second, we assume that the placement generated by

[34]This idea is inspired by Berry et al. (2004) who used hypothetical second choices to infer substitution patterns in the consumer preferences for cars.

[35]It is vital to show that our approach accurately captures the substitution patterns; otherwise, counterfactual analyses could be entirely wrong. When the specification does not capture the substitution patterns well, the random shocks in the preference model tend to be blown up in an attempt to explain variation of choice outcomes in the data. If the model overestimates the contribution of idiosyncratic shocks, it would also overshoot in it's predictions of major switching. For example, see Houde (2012).

each applicant's submitted list is the most preferred choice among the programs that would have been feasible under the 2001 cutoffs. The students are nudged by the system to assign special importance to past year's cutoffs: at the time they are asked to rank their preferences, they are provided the past minimum admission scores for each program. Minimum scores from 2001 are included in the same application package that contains forms for preference list submission. As a result, students are likely to use the 2001 cutoffs as an important benchmark for the lists they submit in 2002. Finally, we assume that programs are listed in the order of true preferences.[36]

Applicants $i = 1, .., I$ choose between programs $j = 1, .., J$. By choosing program j, the student obtains utility

$$u_{ijt} = \underbrace{X_{ij}\beta + Z_{ij}\gamma_t}_{\delta_{ijt}} + \varepsilon_{ijt}$$

Each student may belong to one of T unobservable types: $t = 1, .., T$. Programs are characterized by observables which are contained in X_{ij} and include the level of tuition and its interaction with the income group, the distance to the applicant's high school, whether the program is in the same province or not, the rank of the department, whether it is an evening program, the full set of university dummies. The coefficients on X_{ij} vary by track and gender but not by the unobservable type of the student. For a complete list of the variables in X_{ij} and the reasons for including them see Appendix C.

Types may differ in their preferences for a subset of program characteristics, Z_{ij}. In particular, Z_{ij} includes j's major of study. This is motivated by the data: the choice of major for the top program in a student's list strongly correlates with the major of the second choice.[37] A complete list of the variables in Z_{ij} is in Appendix C. The shares of types in the population are denoted as σ_t.

The term ε_{ijt} captures idiosyncratic preferences and is drawn from the standard Gumbel

[36]This assumption is quite innocuous as a rank order list that does not respect their true order is weakly dominated (Haeringer and Klijn (2009)).

[37]Figure 15 shows the density of students according to the share of the dominant major in their ranked preference list. It is clear from the figure that students fill their preference list with certain types of majors. This suggests that there are different types of students: some who, for example, strongly prefer medicine or engineering and only include such programs on their lists, and others who are more flexible and willing to substitute between different subsets of programs. This motivates us to allow for unobserved heterogeneity in our estimation.

distribution independently across agents, programs, and unobservable types. The well-known property of i.i.d. Gumbel shocks to produce unrealistic substitution patterns is addressed by allowing the coefficients γ_t to vary across the unobservable types. The non-idiosyncratic part of the utility function is denoted as δ_{ijt}.

Each applicant has a set of exam scores, s_i, which determines i's priority in the allocation mechanism. Let C_{i1} denote the set of programs whose minimum admission scores in 2001 are below student i's exam score in 2002. Similarly, C_{i2} is the set of programs ex-post feasible for i in 2002. Finally, for any set of ex-post feasible programs C let $c_{it}(C) = \arg\max_{j \in C} u_{ijt}$ be the most preferred program and $p(C, L_i)$ be the placement outcome given i's submitted preference list, L_i.

Our identification strategy relies on three assumptions.

Assumption 1 *A student's placement in 2002 is ex-post stable. That is, even if student i knew the equilibrium cutoff scores in all programs, he would still prefer his program of placement:*

$$p(C_{i2}, L_i) = c_{it}(C_{i2}),$$

Assumption 2 *A student's hypothetical placement in 2001 is ex-post stable. That is, the student preference list in 2002 would result in optimal placement under the cutoffs from 2001:*

$$p(C_{i1}, L_i) = c_{it}(C_{i1}),$$

Assumption 3 *Programs $p(C_{i1}, L_i)$ and $p(C_{i2}, L_i)$ appear in the applicant's submitted list L_i in the order of true preference:*

$$u_{ij_1t} \geq u_{ij_2t} \text{ if } j_1 = p(C_{i1}, L_i) \text{ is listed before } j_2 = p(C_{i2}, L_i) \text{ and vice versa.}$$

Given the number of programs, as the number of students grows to infinity, the uncertainty in the cutoffs vanishes. Asymptotically, students make fewer and fewer mistakes in terms of the cutoffs. This motivates our first assumption, as in Fack et al. (2019). If, in addition, students used last year's cutoffs as their best guess about the next year's cutoffs, then the second assumption would be true. As the cutoff scores from 2001 were included in the information package that all students received before submitting their lists, it is natural that they assign special importance to the previous year's cutoffs. Finally, as Haeringer and Klijn (2009), Chade and Smith (2006) and Shorrer (2019) show, a rank order list that does not respect the true preference order is weakly dominated. This is what motivates the third assumption. Note that we do not assume that everything on the list is truthfully ranked. Our assumption only applies to the programs of placement under 2001 and 2002 cutoffs.

We use the above three conditions to implement a maximum likelihood estimator for the key preference parameters: $\beta, \gamma_t, \sigma_t, t = 1, \ldots, T$. The likelihood function is derived in Appendix B. We estimate the model independently for male and female applicants in three major high school tracks (Science, Turkish-Math and Social Science). To avoid selection issues caused by exam retaking, we only include those applicants who never took the college entrance exam in the past, that is, first-time takers. We exclude applicants who take the optional language part of the exam as they tend to target a very distinct set of programs. The full details of the implementation of the maximum likelihood method are given in Appendix C.

Intuitively our identification strategy hinges on the idea that by switching from the 2002 cutoffs to the 2001 ones, we perform an experiment that slightly manipulates each student's choice set[38] and elicits a new placement response. We estimate our model to reproduce these responses and then use it to predict placements under counterfactual policies in which the changes in the choice sets are not necessarily small anymore. For instance, suppose that the program of placement $j = p(C_{i2}, L_i)$ is not feasible under the 2001 cutoffs for student i because i's score was slightly above the cutoff in j in 2002, but not in 2001. Suppose also that all the other programs remain feasible: $C_{i2} = C_{i1} \cup \{j\}$. When i compiles his preference list L_i, according to Assumption 2 he would make sure that L_i selects his second preferred program within C_{i1}, $c_{it}(C_{i1})$. Effectively, we have the opportunity to "survey" the student on what his second choice would be if his first choice j were not available, similar to what Berry et al. (2004) do for hypothetical second choices in their data on demand for cars.

[38] In most cases, the choice set under the 2002 cutoffs (C_{i2}) significantly overlaps with C_{i1} as the cutoffs changed very little in 2001-2003.

4.2 Demand Estimates

Table 4 presents the estimates of important common parameters, β, by high school track and gender. The first variable is a dummy for the program being a distance program. These programs tend to be not very competitive; moreover, many of them do not even have binding cutoffs. The next variable in Table 4 is an indicator for the program being an evening program. Evening programs seem to be less disliked than distance ones. These are the same programs offered in the day, but as they typically have lower cutoffs, they may be preferred by working students.[39] The next two variables capture the role of geography: distance between the district of the program's campus and that of the high school attended by the student and an indicator of these districts being in the same province. Programs geographically remote from the applicant's high school tend to be less valued. Applicants also prefer to stay in the same province, even after controlling for distance.

The next set of variables, namely an interaction of the program's tuition and the student's income group dummy, capture the role of tuition and income. Applicants have a strong distaste for high tuition. In line with common wisdom, applicants from more well-off families tend to be less sensitive to tuition.

In all three high school tracks, females have a stronger preference for geographic proximity than males. For instance, a male applicant from a low-income family who graduates from the Science track would be willing to pay 1,363 TL to reduce the distance to a program by 1,000 kilometers.[40] A female applicant with the same background would pay 2,271 TL.[41] One explanation for this result is that female students tend to have a hard time getting permission to move away from their home city (Alat and Alat (2011)). This asymmetry may have important implications for gender gaps in placements: if programs in highly valued majors are concentrated in a few geographic locations, they may be relatively less accessible to female applicants from remote parts of Turkey than for male students from the same

[39]Typically, students do not work while attending college in Turkey.

[40]Different programs have different tuitions. Private college programs have higher tuition than public ones. In private colleges, the same program can be offered with a high tuition option and a low tuition one, with the two having different placement score cutoffs. Such variation lets us interpret estimates in money terms.

[41]The above numbers are roughly similar to 950 US dollars for males and 1,500 US dollars for females in 2002.

Table 4: Estimated Demand Parameters, Common Coefficients β

Track	Science		Turkish-Math		Social Science	
Gender VARIABLES	Female	Male	Female	Male	Female	Male
Distance program	-6.82***	-5.40***	-3.73***	-1.35***	-2.06	-5.62*
	(0.90)	(0.50)	(0.57)	(0.43)	(2.28)	(3.04)
Evening program	0.12	0.29***	-0.16**	-0.03	0.12	0.28
	(0.08)	(0.07)	(0.07)	(0.08)	(0.31)	(0.26)
Distance	-2.93***	-1.91***	-2.59***	-1.76***	-2.34***	-1.82***
	(0.12)	(0.09)	(0.13)	(0.12)	(0.40)	(0.34)
Same province	1.22***	1.05***	1.43***	1.39***	1.62***	1.40***
	(0.08)	(0.07)	(0.09)	(0.10)	(0.36)	(0.30)
Tuition×Income=1	-12.90***	-14.01***	-10.28***	-10.30***	-9.56	-11.76
	(1.35)	(1.20)	(1.05)	(1.58)	(5.60)	(4.64)
Tuition×Income=2	-11.09***	-10.73***	-9.22***	-8.46***	-10.21***	-9.26**
	(0.79)	(0.62)	(0.77)	(0.53)	(1.76)	(2.10)
Tuition×Income=3	-7.21***	-6.83***	-4.97***	-4.90***	-6.47***	-6.78***
	(0.59)	(0.54)	(0.35)	(0.33)	(0.95)	(1.32)

Notes: Standard errors (in the parentheses) are obtained using 1,000 bootstrap samples. Significance levels ($*$ — 10%, $**$ — 5%, $***$ — 1%) are obtained using Hall's bootstrap formula.
Variables: Same province — equals one if the applicant's high school and the program are in the same province. Household income categories: 1 — 0-250 TL/month ("new Lira" in 2002), 2 — 250-500 TL/month, 3 — above 500 TL/month. Unreported common controls: full set of university dummies, program's rank and rank squared in 2001 among programs accepting the same score type.
Units: Tuition — 10,000 TL, distance — 1,000 km.

areas.

We also include a rich set of controls to capture the perceived quality of each program. First, we include a dummy for every university in Turkey.[42] This captures the overall preference for being in a particular university. Second, we include every program's ranking in terms of its cutoff score in 2001. Since each program uses a different type of score (SAY, EA, or SÖZ) the ranking differs according to the type of score a program is using. For example, a program using the Y-ÖSS-EA score that has the highest cutoff in 2001 would have a rank of 1 in the EA category, while the other two rank variables (SAY and SÖZ) would be zero for this program. To allow for a flexible mapping from program quality to its cutoff, we also include the square of the ranking.

[42] The estimates for these dummies are available upon request.

Identifying the effect of tuition is aided by the fact that some private university programs are offered at different tuition levels (full, partial, and no tuition). Even though programs with different tuition levels are treated as separate programs and have different cutoffs, we define the ranking for all of them using the cutoff scores of the lowest tuition one as tuition levels enter separately.[43] This way, for instance, a tuition-free Physics program in Bilkent University would enter the demand equation with the same proxies for quality (ranking in 2001 and the Bilkent University dummy) as the full-tuition version of the same program, but the tuition would vary a lot between these two programs. This variation drives our estimates of the coefficients on tuition in Table 4 and prevents tuition from being confounded with program's quality level.

The estimates for γ_t are reported in the Appendix in Tables 6 to 11. Consider Table 6 which reports the estimates for females in the Science track.[44] The estimated coefficients for each latent type t are reported in each of the 8 columns.[45] The probability that an agent is of a particular type is reported in the last row. Some types are more likely than others: types 3 and 8 are more likely than type 1 or 2. The coefficient on the SAY and EA score dummies capture student preferences for programs relying on the respective score. The study major dummies capture preferences for major. The omitted field is education, so a positive coefficient on a field dummy for a given type means that for this type, such programs are better than education. Type 1 students have a positive coefficient on medicine and health service, which means that students of type 1 tend to like medical or nursing schools. The coefficient on Non-placement$\times$Predicted Score reflects the value of the outside option for the student as a function of the student's predicted score based on his demographics and high school GPA. For most students, the outside option is retaking. Controlling for predicted scores captures the fact that retaking is more valuable for students who are likely to score higher in the subsequent attempt.

[43]This makes sense as students admitted to the same program with different tuition levels attend the same classes.

[44]We assume that there are eight unobserved types for this subgroup. Increasing the number of types is computationally costly, we limit the number of types to the number of majors the students place into or eight, whichever is lower.

[45]Note that types are not comparable across tracks or genders. For example, preferences of type $t = 1$ of female Science track students are unrelated to preferences of type $t = 1$ of males from the Turkish-Math track.

To be of any use, our model should approximate substitution patterns well. If it fails to correctly predict how female applicants react to, for instance, adding more engineering programs to their choice sets, it will be useless in policy experiments aimed at reducing the gender gap in engineering. To evaluate the merits of our identification strategy, we compare it to three alternative approaches laid out in Table 5. Column 1 has our preferred specification. In Column 2, we set up and estimate a similar latent class logit model allowing for unobserved heterogeneity in taste (γ_t coefficients), but using ex-post stability of observed placement as the only identifying restriction (Assumption 1, but not Assumptions 2 or 3). Fack et al. (2019) advocate this approach for settings with large numbers of participants. In Column 3, we maintain the identifying Assumptions 1 – 3, but switch to a simple multinomial logit model effectively removing unobserved heterogeneity in γ_t. Finally, in Column 4, we use the multinomial logit and use Assumption 1 only. In each case for the models in Columns 1-4, we estimate the model and then simulate placements based on the estimates. In Column 5, we assume preferences are as given by the student's list and simulate placements based on this using the admission cutoffs from 2001, that is, the previous year.

The last row in Table 5 gives the percentage of students who switched majors from their allocated ones in 2002 using the placement-generating procedure in each column. Thus, the last row in Column 5 says that if we used the list provided as the preferences of the student but used the 2001 cutoffs, 8.6% of the students would switch their major. If Assumption 2 does hold, one can predict placements directly from the reported preference lists treating them as fixed. This provides a model-free benchmark in Column 5. Thus, a model that captures substitution patterns well should predict that roughly 8.6% of students switch majors if the cutoffs change from those in 2002 to 2001. The last row of Table 5 shows that compared to the main specification in Column 1, the alternative ones in Columns 2, 3, and 4 predict higher rates of major switching in response to the change in cutoffs. Compared to the benchmark in Column 5, our preferred approach fares quite well, while the alternatives tend to predict substantially higher rates of major switching. Not surprisingly, the plain logit specifications in Columns 3 and 4 do not perform well. Since they are not designed to capture unobserved heterogeneity in preferences for specific majors, they tend to predict excessive major switching. Houde (2012) and Petrin (2002) explore the mechanics of this phenomenon

Table 5: Alternative Models and Identifying Assumptions

Specification	(1) Main	(2) Alt. 1	(3) Alt. 2	(4) Alt. 3	(5) Fixed list
Unobserved heterogeneity in γ_t	yes	yes	no	no	
Identifying assumptions:					
Ex-post stability in 2002	yes	yes	yes	yes	
Ex-post stability under 2001 cutoffs	yes	no	yes	no	
Truthful ranking	yes	no	yes	no	
Counterfactual experiment: Cutoffs change from the 2002 to 2001 levels:					
Students switching the major of placement	8.6%	13.0%	11.5%	12.0%	8.6%

Notes: Fixed list specification predicts placements treating preference lists in the data as fixed. The outside option (being placed in the omitted exotic programs or not being placed at all) is treated as a distinct major.

and show that the failure to account for relevant factors can inflate the estimated variance of the idiosyncratic shock in the utility function. Using extra data on choices under the 2001 cutoff does not improve the fit. Allowing for unobserved heterogeneity in tastes for majors improves predictions a lot. However, if one does not augment the ex-post stability assumption with Assumptions 2 and 3, the estimator has hard time picking up the correct substitution patterns from the data. Intuitively, the strength of the tastes for majors is identified by how persistently the a person is sticking to the same major in his preference list. Using Assumption 1 alone amounts to dropping the whole preference list except the program of placement. This discards too much information on how strong the individual preferences for majors are.

To look behind these aggregate numbers for switching majors, we also look at where switches occur when we use our preferred model or stated preferences. As discussed in detail in Appendix D, our model performs very well in matching the substitution patterns coming from the benchmark model. We also show how the alternative models in Table 5, Columns 2-4, fare relative to the benchmark. As expected, they do worse.

5 Policies Targeting the Gender Gap in Placements

5.1 Decomposing the Gender Gap: Preferences vs. Performance

In this section, we look at placements by gender under various counterfactual scenarios for students in the three major high school tracks. In each scenario, we manipulate either placement scores (by giving points) or the preferences of female applicants. We then use these preferences and scores to simulate the student placement mechanism. First-time takers from all the academic tracks[46] are our main focus in this exercise. We include repeat takers and students from other tracks in our analysis by keeping their scores and the reported preference lists as fixed.

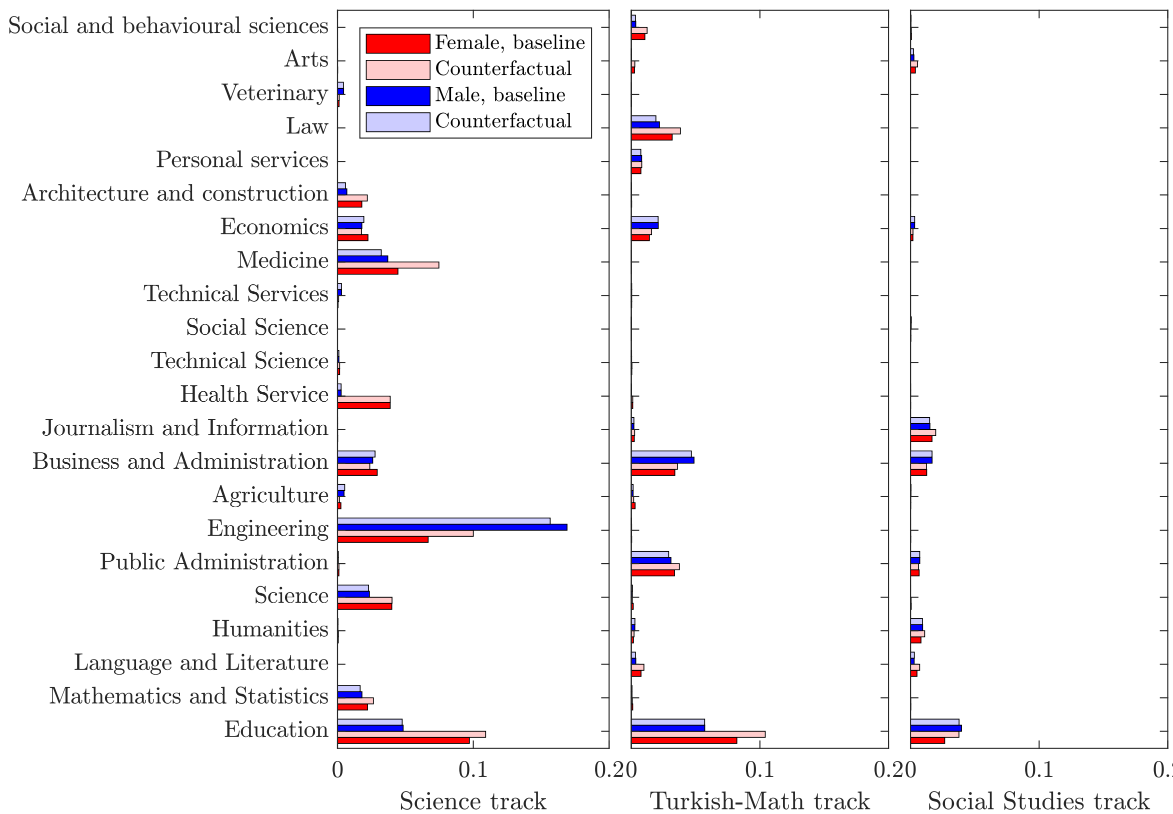
In the first counterfactual experiment, we simulate a policy that eliminates the gender gap in admission scores. We increase the score of every female applicant by the respective estimate in the last column of Table 2. A real-life counterpart of this intervention could be an affirmative action policy granting a score bonus to every female student, or a subsidized preparatory program for females.[47] Figure 3 shows simulated placements by major, high school track, and gender in the counterfactual scenario and the status quo. As we saw earlier, there are large differences by gender in placement. It is also clear that students from the three tracks favor very different subjects by gender. For example, male students from the Science and Turkish-Math tracks are much less likely to be placed in education programs. In contrast, there is no real difference by gender in the counterfactuals for students from the Social Studies track.

Despite giving female applicants a very generous boost in scores, this counterfactual policy fails to close the gender gap in placement to engineering programs. Rather than using their bonus to compete for seats in engineering, most female applicants opt for highly ranked programs in the majors they tend to prefer: medicine, law and education. At the same time, the policy does not lead to a surge in applications in nursing — the least competitive female-dominated major.

[46] The Language track is excluded.

[47] It is worth noting that this bonus would be quite sizable, roughly between one-third and one-half of the standard deviation of the exam score.

Figure 3: Eliminating the Gender Gap in Scores.



Share of placements among first-time takers, by gender, track, and major.

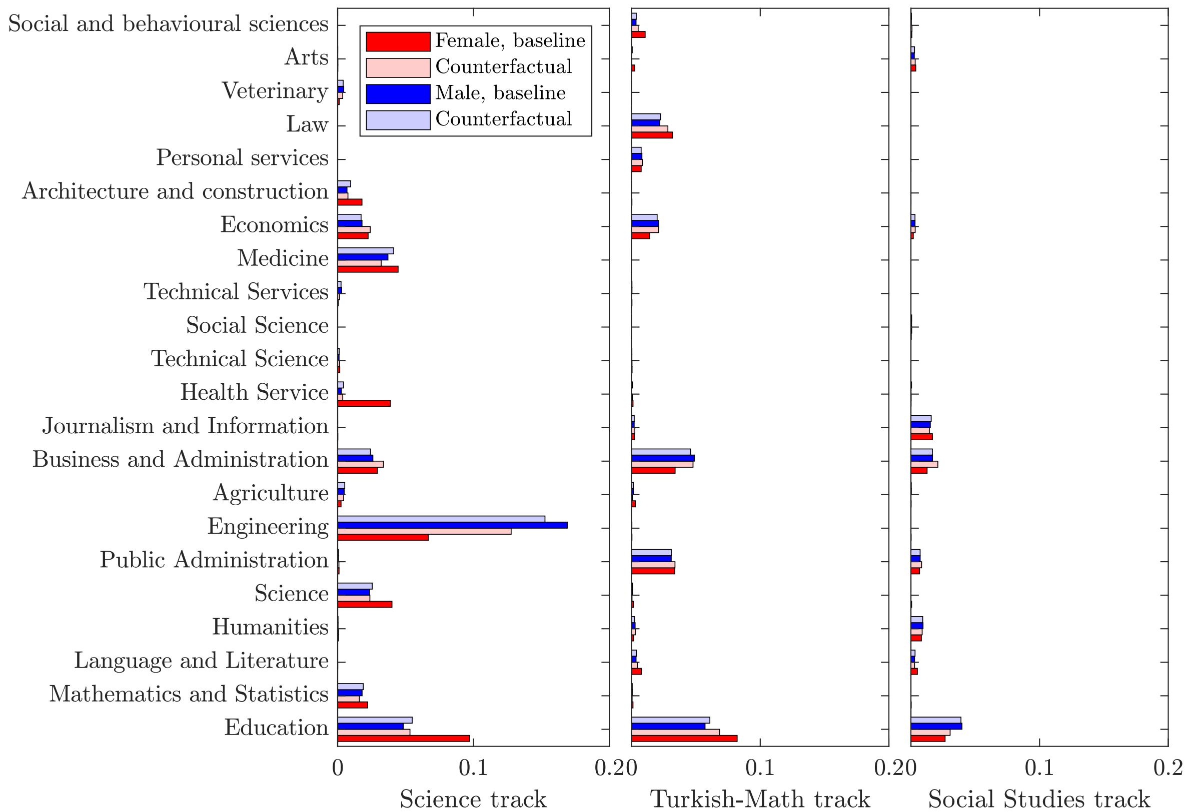
Notes: Baseline — placements predicted by the estimated model, counterfactual — counterfactual policy removing the gender gap in exam score.

In our second counterfactual experiment, we shift focus onto the preference channel. In this scenario, we keep every student's exam score unchanged, but we replace the preference parameters for females with those of males in the same high school track.[48] Figure 4 depicts the predicted placements side by side with those in the status quo scenario. In this scenario, both genders have very similar placement outcomes.[49] Compared to male students, females have a slightly lower chance of getting into competitive majors such as medicine and engineering. This should not be surprising: in the second counterfactual scenario, the existing

[48]For example, for females from the Science track, we use the parameter values in the second column of Table 4 instead of those in the first column.

[49]There is an important caveat: although in this counterfactual experiment males and females in the same high school track have similar placement outcomes, the gender ratio varies a lot between the tracks. Thus, without conditioning on a track, an average female would still differ from an average male in terms of her placement major.

Figure 4: Eliminating the Gender Gap in Preferences.



Share of placements among first-time takers, by gender, track, and major.

Notes: Baseline — placements predicted by the estimated model, counterfactual — counterfactual policy replacing preference parameters for females with those of males.

gender gap in scores gives male students an upper hand. The above two computational experiments suggest that the preference channel shapes most of the observed gender gap in placements. A policy that merely closes the performance gap is unlikely to achieve gender balance in most majors, and in some cases could tip the scales towards even greater gender segregation.

5.2 Policies Targeting Gender Ratios in Engineering

Granting a uniform bonus to all female applicants is a blunt policy tool. In this subsection, we explore more nuanced bonus policies to achieve gender parity in admissions to engineering programs, as these programs are highly popular and very unequal in terms of gender parity. One such policy grants extra score points to females whenever they are considered for admission to engineering programs, but does not raise their scores otherwise. Such a bonus creates an incentive for females to apply for engineering as it does not improve their standing in the admission rankings in other majors. Another possible policy grants extra stipends to all females enrolled in engineering programs, but does not alter their scores. Finally, the third type of policy is a combination of the first two: it uses score and stipend bonuses in conjunction to achieve a given level of gender parity. Under this set of policies, every engineering program offers a bundle of an extra stipend and an exam score bonus to all female first-time takers from the major academic high school tracks. We run this simulation exercise for a range of stipends and score bonus combinations and calculate the female-to-male ratio of placement odds in engineering as the main outcome of interest:[50]

$$\frac{\Pr\{i \text{ placed in engineering} | i \text{ is a female Science track student}\}}{\Pr\{i \text{ placed in engineering} | i \text{ is a male Science track student}\}} \tag{2}$$

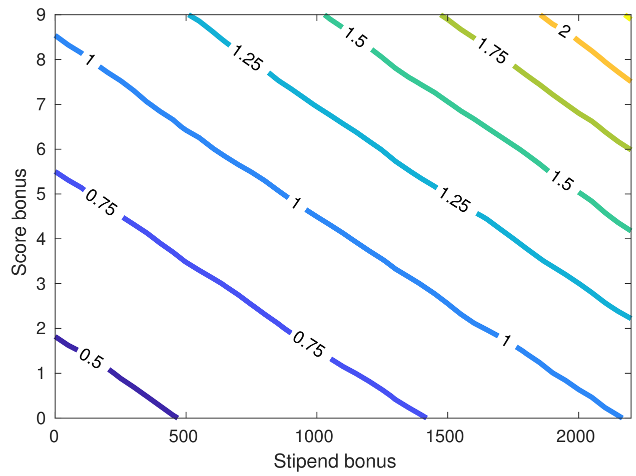
A ratio of unity indicates that females and males coming from the Science track have equal chances of being placed in an engineering program.

We represent the results of all three of these policies succinctly in Figure 5 which shows how the policy parameters affect the odds ratio (2) in equilibrium. The labeled lines correspond to policies that result in the odds ratio reaching 0.5, 0.75, 1, and so on. For example, offering a stipend of around 2150 TL per year (roughly 1400 US dollars in 2002, which is roughly 20% of the full tuition rate charged at the prestigious Bilkent University at the time) would attract enough female applicants to engineering programs to eliminate the gender gap in placements in this major. The policy of giving a bonus of 8.5 extra points to females when they are considered for engineering programs would lead to a similar outcome. These are sizable numbers. The shapes of the policy isolines also suggest that stipends and score bonuses are almost perfect substitutes as the isolines in the figure are nearly straight. This implies that combining the score bonus with the stipend policy would not reduce the magnitude of the required intervention, as there seem to be minimal diminishing returns to using

[50]The numerator and the denominator in this ratio correspond to the red and the blue bar in the first panel of Figure 3 (the line labeled "Engineering").

a given policy.

Figure 5: Female-to-Male Odds Ratio of Being Admitted to an Engineering Program Among Science Track Students



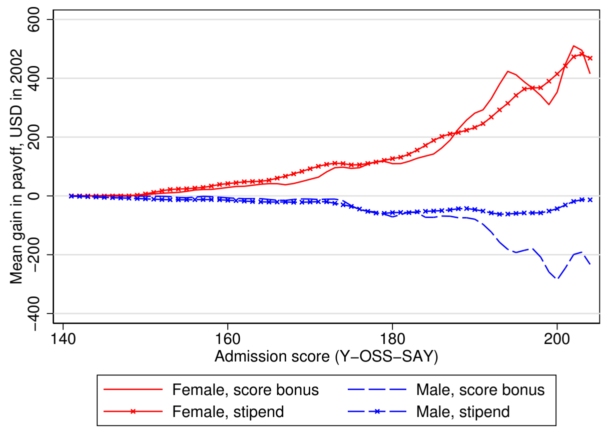
Notes: Each point (x, y) represents a counterfactual policy in which every engineering program offers a stipend bonus of x TL per year and adds y extra points to the entrance exam score for every eligible female applicant. The graph plots odds ratio isolines, i.e., all stipend-score bonus combinations that achieve a given odds ratio of being admitted to an engineering program for females and males from the science high school track. The respective odd ratio is shown by each line's label.

Stipends and score bonuses affect different parts of the student population. To better understand these trade-offs, we simulate and compare the two polar policies described above, one granting the score bonus of 8.5 points and one granting the annual stipend of 1400 US dollars.[51]

Figure 6 shows the expected welfare change over the status quo for first-time takers from the Science track under both policies as a function of admission score and gender. The score bonus policy improves the welfare of females and reduces that of males by roughly the same amount at each level of the admission score. The gains and the losses are especially high at the upper end of the score distribution, as higher-scoring students are more likely to apply to engineering programs. Low-scoring students are unlikely to be affected by the policy, as engineering programs are typically beyond their reach, even with the bonus applied. The

[51]To ensure that both policies are budget-neutral, we assume in the body of the paper, that the stipends are financed by a lump sum tax on all agents in the economy so that these taxes are insignificant for each student. In the Appendix, in Figure 16, we change this assumption so that the stipends are financed by levying a tax on all first-time applicants in the Science track. The tax is only levied on applicants so that the levels of welfare shift by the amount of the tax.

Figure 6: Bonus Policies for Females in Engineering: Expected Changes in Payoffs by Score



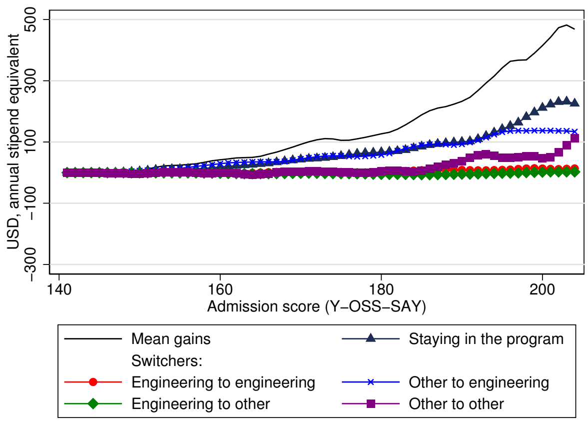
Notes: This graph compares two bonus policies: (a) a stipend for female engineering students, (b) a score bonus granting an admission priority for female applicants in engineering programs. Both policies are calibrated to achieve gender equality in admission probabilities to the engineering major and restricted to first-time takers from the academic track of high school. The payoffs are expressed as annual stipend equivalents in 2002 US dollars.

stipend bonus policy also has greater effects at the upper end of the score distributions, as high-scoring women are the ones who tend to go into engineering. However, in marked contrast to the score bonus, with the stipend bonus policy, the gains for females are not mirrored by equal losses for males. While males do lose overall, their losses are far smaller on average and almost zero for students at the very top of the score distribution, making this more of a win-win policy.

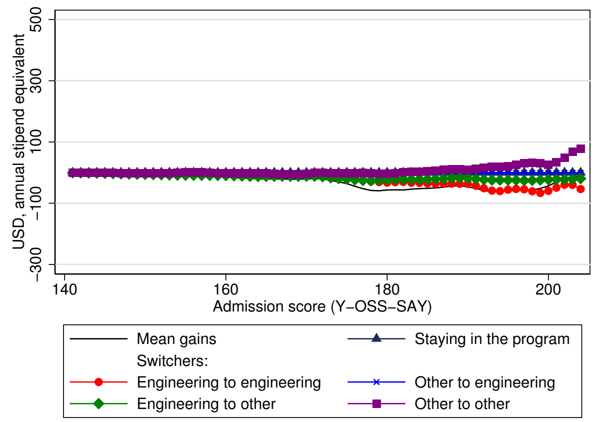
The score and the stipend bonus policies work via very different channels. Figures 7a and 7b present mean gains in student welfare caused by the stipend bonus policy and their decomposition by student groups defined by the choice of study major. Note that females tend to gain more for every one of these groups than males lose. Also, the gains for female students who stay in the same program ("Staying in the program") are largest. They gain just from getting the stipend, and these gains are larger at the upper end of the score distribution as this is where these women are concentrated. The gains for women who move from other programs to engineering are also substantial. The gains for the other groups are much smaller. Women switching from non-engineering to non-engineering majors ("Other to other") gain due to the fall in competition, especially in highly ranked medical

Figure 7: Mean Welfare Gains by Score Under Various Bonus Policies: Main Channels

Stipend Bonus Policy

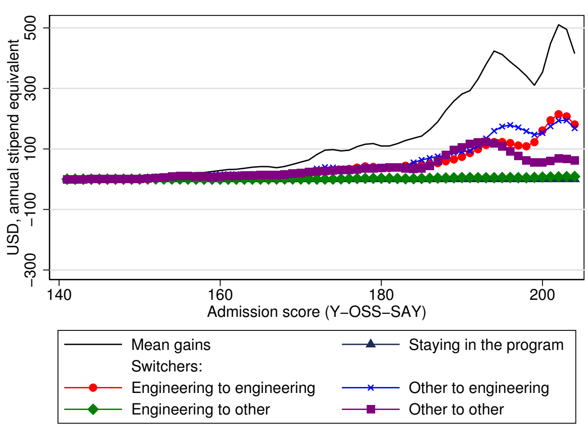


(a) Females

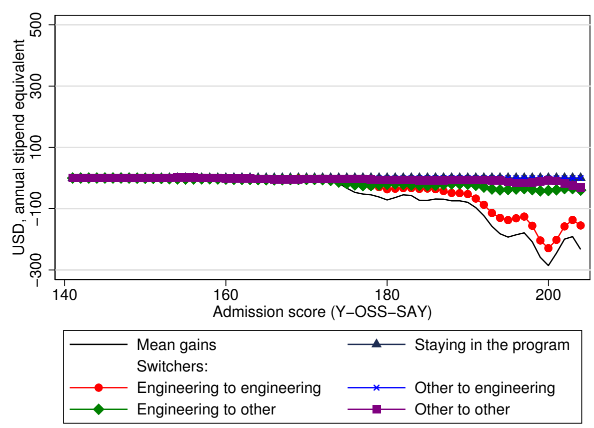


(b) Males

Score Bonus Policy



(c) Females



(d) Males

Notes: The figures depict mean welfare gains after introducing the stipend or the score bonus policy to achieve full gender equality in admissions to the engineering major. The gains are depicted with solid lines. Dotted and dashed lines depict the part of mean gains attributed to students staying in the same program, or switching programs, but staying in the engineering major, or switching major from non-engineering to engineering, and so on. When added up, the dotted lines match the total mean gains.

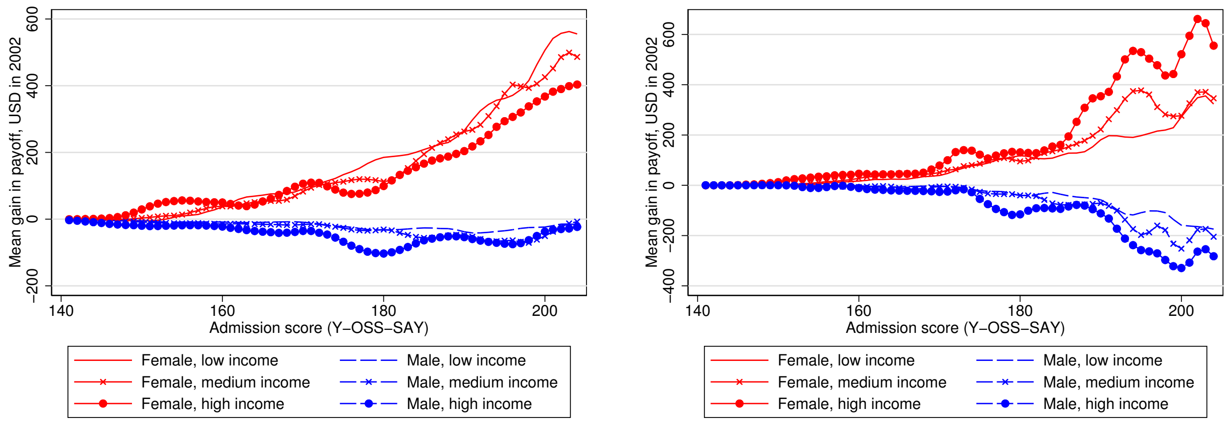
programs. Women who switch from one engineering program to another ("Engineering to engineering") do not gain as this group includes females displaced from their status-quo preferred programs by highly-ranked females switching from other majors. Males at the top of the score distribution can either lose or gain depending on their taste for majors: those who prefer engineering still choose engineering, but are forced to lower-ranked programs and hence lose, while those who prefer medicine upgrade their choices due to the reduction in competition and gain. Other groups are essentially not affected.

Figures 7c and 7d decompose welfare gains under the score bonus policy. Females mostly take advantage of the bonus by upgrading their choices within the engineering field ("Engineering to engineering") or by switching towards engineering from the other fields ("Other to engineering"). As before, some females not interested in engineering ("Other to other") are better off as they have less competition in their chosen fields. The lion's share of payoff changes for male students come from downgrading their choices within the engineering major ("Engineering to engineering").

As the above figures make clear, both policies are wasteful to the extent that they are paid, but do not affect choices. For the stipend policy, the main source of waste is infra-marginal students, that is, female applicants whose program choices are not affected by the policy, but whose stipends have to be paid. Likewise, the score bonus policy displaces many male students to accommodate females who would choose engineering even without the bonus.

Finally, we look at how the two policies affect students as a function of their income. Figures 8a and 8b depict this for the stipend bonus and the score bonus, respectively. As before, the gains for women (and losses for men) are increasing in their placement scores in both (a) and (b). The stipend bonus helps low-income women the most, followed by medium-income ones and high-income ones. This makes sense, as low-income women value the stipend bonus more and are more likely to respond to the policy. The losses on the part of men are smaller than the gains to women and are larger as income rises, suggesting that low-income females are replacing high-income males, especially at the top of the score distribution. The score bonus has the opposite pattern. High-income women gain the most and low-income ones the least, while high-income males lose the most and low-income ones

Figure 8: Mean Welfare Gains by Score and Income Group



(a) Stipend Bonus

(b) Score Bonus

Notes: This graph compares two bonus policies: (a) a stipend for female engineering students, (b) a score bonus granting an admission priority for female applicants in engineering programs. Both policies are calibrated to achieve gender equality in admission probabilities to the engineering major and restricted to first-time takers from the academic track of high school. The payoffs are expressed as annual stipend equivalents in 2002 US dollars.

lose the least. This suggests that out of the two policies, the stipend policy generates higher gains for lower-income females. For males, the stipend policy avoids high concentration of losses in any specific ability or income group.

6 Conclusion

In this paper, we investigate the drivers of the gender gap in college major choice and the effectiveness of policy interventions aimed at reducing this disparity. Using rich administrative data from the Turkish University Entrance Exam and employing a novel approach to estimate preferences, we make several key contributions to the literature on gender differences in college major choice and affirmative action policies.

First, we find that the gender gap in placement is mainly driven by differences in preferences and performance, not due to any reluctance to apply to highly competitive programs on the part of women. Second, while differences in performance can account for a small part of the gender gap in placement, differences in preferences are more important. Third, we evaluate the effects of two targeted policy interventions, stipend bonus and score bonus, that aim to close the gender gap in placements in engineering majors. Our results show that, while both policies reduce the gap, they have very different distributional effects. Score subsidies, which increase women's exam scores, disproportionately benefit high-income women at the cost of similarly scored men. In contrast, stipend bonuses favor low-income women and result in minimal welfare losses for men, making this policy more equitable and pro-poor. These results suggest that the design of gender equity policies should be carefully tailored to consider both their intended beneficiaries and unintended consequences. In this, we contribute to the broader literature on affirmative action by providing a unique comparison between score-based and financial-based affirmative action policies. Our findings indicate that financial-based interventions, such as stipends, can achieve gender parity with fewer trade-offs in terms of overall welfare.

References

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

Akyol, P., Key, J., and Krishna, K. (2022). Hit or miss? Test taking behavior in multiple choice exams. *Annals of Economics and Statistics*, 147:3–50.

Alat, Z. and Alat, K. (2011). A qualitative study of parental resistance to girls' schooling. *Educational Sciences: Theory and Practice*, 11(3):1369–1373.

Arcidiacono, P. (2004). Ability sorting and the returns to college major. *Journal of Econometrics*, 121(1-2):343–375.

Arcidiacono, P. (2005). Affirmative action in higher education: How do admission and financial aid rules affect future earnings? *Econometrica*, 73(5):1477–1524.

Arcidiacono, P., Hotz, V. J., Maurel, A., and Romano, T. (2020). Ex ante returns and occupational choice. *Journal of Political Economy*, 128(12):4475–4522.

Arenas, A. and Calsamiglia, C. (2020). The design of university entrance exams and its implications for gender gaps. *Unpublished Manuscript*.

Artemov, G., Che, Y.-K., and He, Y. (2017). Strategic 'mistakes': Implications for market design research. *Unpublished Manuscript*.

Arum, R., Roksa, J., and Budig, M. J. (2008). The romance of college attendance: Higher education stratification and mate selection. *Research in Social Stratification and Mobility*, 26(2):107–121.

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

Balinski, M. and Sönmez, T. (1999). A tale of two mechanisms: Student placement. *Journal of Economic Theory*, 84(1):73–94.

Berry, S., Levinsohn, J., and Pakes, A. (2004). Differentiated products demand systems from a combination of micro and macro data: The new car market. *Journal of Political Economy*, 112(1):68–105.

Berry, S. T. (1994). Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics*, 25(2):242–262.

Calsamiglia, C., Fu, C., and Güell, M. (2020). Structural estimation of a model of school choices: The Boston mechanism versus its alternatives. *Journal of Political Economy*, 128(2):642–680.

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

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

Cestau, D., Epple, D., and Sieg, H. (2017). Admitting students to selective education programs: Merit, profiling, and affirmative action. *Journal of Political Economy*, 125(3):761–797.

Chade, H. and Smith, L. (2006). Simultaneous search. *Econometrica*, 74(5):1293–1307.

Chari, A. and Goldsmith-Pinkham, P. (2017). Gender representation in economics across topics and time: Evidence from the NBER Summer Institute. Technical report, National Bureau of Economic Research.

Eckel, C. C. and Grossman, P. J. (2008). Men, women, and risk aversion: Experimental evidence. *Handbook of Experimental Economics*, 1(113):1061–1073.

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

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., and Rubin, D. B. (2014). *Bayesian Data Analysis*. Chapman and Hall/CRC, 3rd edition.

Gneezy, U., Niederle, M., and Rustichini, A. (2003). Performance in competitive environments: Gender differences. *The Quarterly Journal of Economics*, 118(3):1049–1074.

Haeringer, G. and Klijn, F. (2009). Constrained school choice. *Journal of Economic Theory*, 144(5):1921–1947.

Hassidim, A., Marciano, D., Romm, A., and Shorrer, R. I. (2017). The mechanism is truthful, why aren't you? *American Economic Review*, 107(5):220–24.

Hassidim, A., Romm, A., and Shorrer, R. I. (2021). The limits of incentives in economic matching procedures. *Management Science*, 67(2):951–963.

He, Y. (2017). Gaming the Boston school choice mechanism in Beijing. Technical report, Toulouse School of Economics (TSE).

Houde, J.-F. (2012). Spatial differentiation and vertical mergers in retail markets for gasoline. *American Economic Review*, 102(5):2147–82.

Hsieh, C.-T., Hurst, E., Jones, C. I., and Klenow, P. J. (2019). The allocation of talent and U.S. economic growth. *Econometrica*, 87(5):1439–1474.

Kahn, S. and Ginther, D. (2018). *The Oxford Handbook of Women and the Economy*, chapter 31, Women and STEM. Oxford Handbooks, Oxford.

Kirkebøen, L., Leuven, E., and Mogstad, M. (2021). College as a marriage market. *NBER Working paper*.

Krishna, K., Lychagin, S., and Frisancho, V. (2018). Retaking in high stakes exams: Is less more? *International Economic Review*, 59(2):449–477.

Larroucau, T. and Rios, I. (2020). Do "short-list" students report truthfully? Strategic behavior in the Chilean college admissions problem. *Unpublished manuscript*.

Lundberg, S. and Stearns, J. (2019). Women in economics: Stalled progress. *Journal of Economic Perspectives*, 33(1):3–22.

Ngo, D. and Dustan, A. (2021). Preferences, access, and the stem gender gap in centralized high school assignment. *American Economic Journal: Applied Economics*, 6(4):257–287.

Niederle, M. and Vesterlund, L. (2007). Do women shy away from competition? do men compete too much? *The Quarterly Journal of Economics*, 122(3):1067–1101.

Niederle, M. and Vesterlund, L. (2011). Gender and competition. *Annual Review of Economics*, 3(1):601–630.

Ors, E., Palomino, F., and Peyrache, E. (2013). Performance gender gap: Does competition matter? *Journal of Labor Economics*, 31(3):443–499.

Petrin, A. (2002). Quantifying the benefits of new products: The case of the minivan. *Journal of Political Economy*, 110(4):705–729.

Rees-Jones, A. (2018). Suboptimal behavior in strategy-proof mechanisms: Evidence from the residency match. *Games and Economic Behavior*, 108:317–330.

Saygin, P. O. (2016). Gender differences in preferences for taking risk in college applications. *Economics of Education Review*, 52:120–133.

Shorrer, R. I. (2019). Simultaneous search: Beyond independent successes. *Unpublished manuscript*.

Shorrer, R. I. and Sóvágó, S. (2023). Dominated choices in a strategically simple college admissions environment. *Journal of Political Economy Microeconomics*, 1(4):781–807.

Stinebrickner, R. and Stinebrickner, T. R. (2014). A major in science? Initial beliefs and final outcomes for college major and dropout. *Review of Economic Studies*, 81(1):426–472.

Taylor, J. (2019). Fairness to gifted girls: Admissions to New York City's elite public high schools. *Journal of Women and Minorities in Science and Engineering*, 25(1).

Turner, S. E. and Bowen, W. G. (1999). Choice of major: The changing (unchanging) gender gap. *Industrial and Labor Relations Review*, 52(2):289–313.

Voyer, D. and Voyer, S. D. (2014). Gender differences in scholastic achievement: A meta-analysis. *Psychological Bulletin*, 140(4):1174–1204.

Wiswall, M. and Zafar, B. (2021). Human capital investments and expectations about career and family. *Journal of Political Economy*, 129(5):1361–1424.

Wu, A. H. (2018). Gendered language on the economics job market rumors forum. In *AEA Papers and Proceedings*, volume 108, pages 175–79.

Wu, A. H. (2020). Gender bias among professionals: An identity-based interpretation. *Review of Economics and Statistics*, 102(5):867–880.

A Additional Tables and Figures

Table 6: Type-Specific Demand Coefficients and Type Shares: Science Track, Female

Placement score used:								
SAY	11.13	-14.36***	-8.28**	-7.51***	-3.41	-6.21***	-2.22	-3.10**
EA	7.70	-3.97	-4.55*	-2.25	-2.64	-6.89**	1.86**	2.53***
Program major:								
Agriculture	4.92***	-0.06	-2.79***	-3.13***	-4.45***	-4.84***	-4.93***	-7.75***
Architecture	7.50***	-2.11**	0.56	-0.00	-3.27***	0.52	-5.12***	-3.83**
Business	1.21	-0.35	-0.17	2.95**	-1.73*	4.07***	-1.00	-6.68***
Economics	1.14	-4.22***	0.46	2.81*	2.50	3.41**	-1.47	-7.55***
Engineering	-2.05***	3.30***	0.28	0.30	-0.75	0.89**	-4.69***	-5.65***
Health Service	6.10***	-0.25	-3.95***	-3.24***	4.50***	-4.73***	-3.07*	-2.32*
Mathematics	3.93***	-1.20	0.22	0.06	-0.92	-0.24	0.08	-6.45***
Medicine	5.32***	3.33*	2.17	-0.74	7.18***	-3.98***	2.50	2.33**
Science	2.79***	-1.46	0.78	-4.60***	0.99	-0.28	-0.08	-5.42***
Other majors	1.19	-1.27	-0.50	0.36	-3.63	0.80*	-7.61***	-10.59***
Non-placement×								
predicted score	5.18	2.63**	2.09*	1.02	3.71***	-0.59	0.66	0.57
Type share	0.02	0.08	0.32***	0.09	0.06	0.08	0.09	0.26***

Notes: Significance levels (* — 10%, ** — 5%, *** — 1%) are obtained using 1,000 bootstrap samples. Placement score dummies indicate programs accepting SAY, EA or SÖZ scores. The interaction of non-placement dummy and predicted score captures the value of the outside option depending on the student's expected score in the main field (SAY for the Science track, EA for the Turkish-Math track and SÖZ for the Social Studies track) predicted using demographic variables and past performance.

Table 7: Type-Specific Demand Coefficients and Type Shares: Science Track, Male

Placement score used:								
SAY	38.53***	-2.61	-5.55	-10.90***	-7.08***	-7.16***	-6.37***	-3.65***
EA	24.35***	-0.64	-3.24	-7.37***	-5.07*	0.30*	-2.06	1.05**
Program major:								
Agriculture	2.03	0.72	0.59	0.08	-1.48	-1.60	-3.19***	-6.86***
Architecture	-1.00	3.57***	-5.25***	0.14	-2.39	-1.54	-3.98***	-4.56***
Business	1.58	3.29***	-0.95	3.89***	2.17	-0.23	3.42***	-4.89***
Economics	2.28**	-2.48***	-0.55	2.88***	2.19	0.66	3.43***	-7.56***
Engineering	-0.55	2.49**	-0.71*	2.38	2.48	4.50***	0.21	-4.35***
Health Service	-3.64***	-3.71***	0.06	-2.03	-3.81***	1.77**	-3.95***	-3.53*
Mathematics	-2.92***	-0.37	-1.05	1.27	-3.34***	6.10***	-0.87	-3.29**
Medicine	17.96***	-1.95***	3.44	2.21	0.46	5.39***	0.05	1.58***
Science	-1.50	-0.32	0.22	0.02	-4.27***	5.39***	-1.26	-5.13***
Technical Science	-2.33*	-2.24	-1.78	2.33*	-3.88***	3.48***	-3.91***	-4.86***
Technical Services	-0.23	-1.51	-0.10	-2.40***	-1.99***	7.28***	-3.13***	-2.74**
Veterinary	6.48***	2.58	1.96	0.58	-0.23	4.17***	-1.89*	-1.12
Other	3.34***	-1.94	-0.20	-5.11***	-3.97***	-2.73**	-1.34	-4.94***
Non-placement×								
predicted score	23.08***	-0.23	1.87	2.33***	0.44	0.89	0.03	-0.48
Type share	0.01	0.05	0.16*	0.29***	0.22***	0.07	0.06	0.13***

Notes: Significance levels (* — 10%, ** — 5%, *** — 1%) are obtained using 1,000 bootstrap samples. Placement score dummies indicate programs accepting SAY, EA or SÖZ scores. The interaction of the non-placement dummy and predicted score captures the value of the outside option depending on the student's expected score in the main field (SAY for the Science track, EA for the Turkish-Math track, and SÖZ for the Social Studies track) predicted using demographic variables and past performance.

Table 8: Type-Specific Demand Coefficients and Type Shares: Turkish-Math Track, Female

Placement score used:								
EA	5.46*	-4.43**	-3.71	-13.38***	-5.73	-3.15	0.22	-0.94**
SOZ	-9.57**	-9.00	-13.51***	-14.11***	-13.28***	-10.52*	-9.60**	-5.91
Program major:								
Arts	5.79***	5.20***	2.92***	-0.02	-0.72	-1.09	-3.80***	-5.44***
Business	2.71***	-4.73***	-3.37	-1.32	1.09	-3.47	-7.96***	-7.49***
Economics	2.03**	-1.57	-6.00***	0.14	1.19	-3.39	-7.32***	-6.97***
Humanities	-3.05*	-1.13	-2.23	4.34***	-2.41	-7.16***	-11.24***	-9.95***
Journalism	-0.78	-2.12	-2.88	6.23***	-4.71*	-8.57***	-12.20***	-10.42***
Language and Literature	-0.00	2.92***	-4.23***	1.04	-1.36	-5.35***	1.64	-3.21*
Law	-0.99**	6.96***	-0.41	11.05***	1.95	3.72**	-7.16***	-6.12***
Personal services	-1.22	-0.62	-3.65**	-0.12	0.36	-6.55***	-7.48***	-7.68***
Public Administration	3.06***	-3.84***	-0.59	10.74***	1.33	-3.27	-9.71***	-5.09
Social sciences	0.71	-4.44***	-0.00	6.59***	-1.05	-6.36***	-7.54***	-8.89***
Other	3.91***	-7.50***	-8.67***	-9.24***	-10.05***	-9.16***	-3.79	-8.85***
Non-placement×								
predicted score	9.33***	-0.29	2.67*	-0.67	0.37	3.09**	0.83	2.44***
Type share	0.03	0.02	0.13*	0.05	0.13	0.29***	0.14	0.22***

Notes: Significance levels (* — 10%, ** — 5%, *** — 1%) are obtained using 1,000 bootstrap samples. Placement score dummies indicate programs accepting SAY, EA or SÖZ scores. The interaction of the non-placement dummy and predicted score captures the value of the outside option depending on the student's expected score in the main field (SAY for the Science track, EA for the Turkish-Math track, and SÖZ for the Social Studies track) predicted using demographic variables and past performance.

Table 9: Type-Specific Demand Coefficients and Type Shares: Turkish-Math Track, Male

Placement score used:								
EA	7.20*	-6.32***	0.70	-7.87**	-9.66***	-4.23**	-2.72**	-1.09*
SOZ	-3.02	-18.11***	0.60	-10.98**	-9.76***	-8.99***	-6.01**	-12.02***
Program major:								
Business	3.80***	2.03*	1.64***	1.32***	-1.70	-3.62***	-4.97***	-8.89***
Economics	2.94**	2.15**	2.85***	1.57	3.84***	-6.82***	-5.90***	-6.84***
Journalism	-2.16***	-5.18***	7.64***	-3.31**	3.05	-6.07***	-7.87***	-6.15***
Language and Literature	7.97***	-0.13	-2.30	-3.93***	-1.34	-5.32***	-1.83	2.63***
Law	5.26***	2.20	-0.02	3.95	6.78***	3.18**	-5.85***	-5.05***
Personal services	-1.95***	0.47	10.22***	-2.91**	-1.47	-3.97	-7.54***	-8.15***
Public Administration	4.11***	1.47	-1.39	-3.20***	7.16***	-2.02	-4.95***	-5.09***
Social sciences	-2.52***	-0.69	8.52***	-4.72***	-2.17	-3.36	-8.77***	-4.33
Other	4.71***	-10.45***	-2.12	-8.86***	-8.73***	-9.82***	-9.92***	-2.60**
Non-placement×								
predicted score	10.34**	0.57	-23.28	4.42**	1.63	3.11***	2.26***	2.62***
Type share	0.02**	0.19***	0.00	0.11*	0.07	0.28***	0.29***	0.04

Notes: Significance levels (* — 10%, ** — 5%, *** — 1%) are obtained using 1,000 bootstrap samples. Placement score dummies indicate programs accepting SAY, EA, or SÖZ scores. The interaction of the non-placement dummy and predicted score captures the value of the outside option depending on the student's expected score in the main field (SAY for the Science track, EA for the Turkish-Math track, and SÖZ for the Social Studies track) predicted using demographic variables and past performance.

Table 10: Type-Specific Demand Coefficients and Type Shares: Social Science Track, Female

Placement score used:								
EA	-42.01***	-7.78	-9.23	-1.65	44.36***	17.22	-14.36**	-12.86***
SOZ	-41.67***	-13.25	-10.19	-8.98*	0.00	14.50***	-11.76*	-5.11
Program major:								
Arts	3.62*	2.07	1.42**	0.05	-0.00	-0.55	-2.13	-5.12***
Business	1.77**	-3.54*	0.66	-7.57***	4.18***	-3.23	-0.71	-2.41
Humanities	-4.83**	-3.24	-6.79***	-5.88***	-0.01	-2.31	-3.34	-2.98
Journalism	1.39	0.01	1.48**	-5.86***	-0.00	0.33	-0.04	-7.77***
Language and Literature	-3.87*	0.67	-5.81**	-4.90***	-0.00	2.66***	2.35***	-4.95**
Public Administration	4.17**	-2.82*	4.30**	-4.89***	4.24***	-5.23***	0.01	-0.05
Other	-4.11**	0.74	-7.27***	-9.11***	2.64***	-2.27	-5.22	-5.47*
Non-placement× predicted score	-62.96***	-3.04	-1.05	0.01	78.83***	72.40***	-1.58	-0.73
Type share	0.02	0.06	0.31***	0.16*	0.03	0.01	0.13	0.28***

Notes: Significance levels (* — 10%, ** — 5%, *** — 1%) are obtained using 1,000 bootstrap samples. Placement score dummies indicate programs accepting SAY, EA or SÖZ scores. The interaction of the non-placement dummy and predicted score captures the value of the outside option depending on the student's expected score in the main field (SAY for the Science track, EA for the Turkish-Math track, and SÖZ for the Social Studies track) predicted using demographic variables and past performance.

Table 11: Type-Specific Demand Coefficients and Type Shares: Social Science Track, Male

Placement score used:							
EA	-6.46	1.47	5.00	-2.10	-9.91**	-0.02	-9.19**
SOZ	-2.27	5.85	0.75	8.93**	0.85	-4.62**	2.41
Program major:							
Arts	3.38**	-0.55**	-0.93	-1.85	-2.26	-5.13***	-5.61**
Business	-0.90	1.34	3.94***	2.62**	-0.43	-7.13***	-0.35
Humanities	-1.27	-1.17	3.23***	-5.88***	0.63	-4.43	-6.48**
Journalism	3.61***	2.78*	-1.29	-9.45***	-2.43	-1.08	-6.36*
Public Administration	6.72***	4.34**	3.20*	3.70***	-0.00	-1.69	-0.01
Other	-5.46***	2.28*	3.23***	-5.11**	-5.68**	-5.08**	-2.98
Non-placement× predicted score	-0.53	14.16	5.18	-7.34	-1.67	-0.96	2.12
Type share	0.16***	0.03	0.04	0.02	0.13	0.22	0.41***

Notes: Significance levels (* — 10%, ** — 5%, *** — 1%) are obtained using 1,000 bootstrap samples. Placement score dummies indicate programs accepting SAY, EA or SÖZ scores. The interaction of the non-placement dummy and predicted score captures the value of the outside option depending on the student's expected score in the main field (SAY for the Science track, EA for the Turkish-Math track, and SÖZ for the Social Studies track) predicted using demographic variables and past performance.

Table 12: Average Monthly Earnings in TL and Employment Probability by Field of Study in 2009

	25-30 years-old				40-50 years-old			
	Earnings		Employment		Earnings		Employment	
Field of Study	Female	Male	Female	Male	Female	Male	Female	Male
Teacher training and education science	1281.24	1405.21	0.74	0.81	1572.70	1686.67	0.73	0.90
Arts	1139.93	1144.00	0.51	0.67	1965.35	1665.00	0.67	0.82
Humanities	1040.10	1350.76	0.65	0.81	1647.96	1619.64	0.77	0.92
Social and behavioral science	1324.96	1575.46	0.56	0.74	1836.75	1823.84	0.62	0.87
Journalism and information	1158.46	1337.50	0.65	1.00	1575.00	2350.00	0.55	1.00
Business and administration	1074.64	1227.87	0.58	0.79	1701.64	1863.27	0.59	0.83
Law	1998.49	2031.44	0.75	0.92	2400.00	2767.08	0.91	0.97
Life science	1046.83	1069.44	0.63	0.66	1461.09	1743.56	0.79	0.88
Physical science	1327.31	1472.16	0.69	0.71	2157.74	2088.06	0.69	0.90
Mathematics and statistics	1042.57	1288.38	0.75	0.82	1583.32	1803.50	0.79	0.97
Computing	1450.17	1239.94	0.59	0.79	2000.00	2045.56	0.25	0.83
Engineering and engineering trades	1419.92	1238.02	0.62	0.83	2052.05	2001.92	0.69	0.92
Manufacturing and processing	1074.75	1287.87	0.55	0.81	1630.00	1741.71	0.53	0.87
Architecture and building	1226.24	1425.72	0.70	0.79	1814.29	2081.39	0.74	0.91
Agriculture, forestry and fishery	980.69	1205.58	0.55	0.75	1747.24	1878.02	0.74	0.93
Veterinary	1561.29	1304.81	0.89	0.79	1798.50	2034.94	0.92	1.00
Health	1592.14	2156.33	0.86	0.88	4031.55	5497.93	0.77	0.95
Personal services	1024.21	1031.26	0.59	0.69	1454.10	1585.42	0.52	0.84
Security services	1895.00	1882.24	0.75	1.00		2166.33		0.75

Note: The Average Dollar-Turkish Lira exchange rate in 2009 is 1.65 TL

Figure 9: Minimum Cutoff Scores

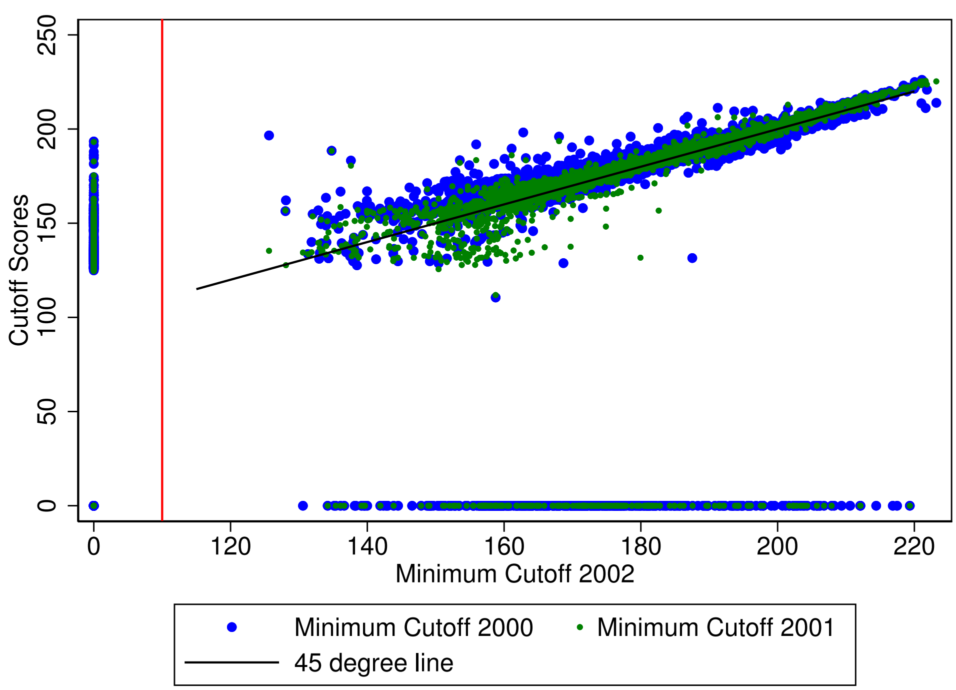


Figure 10: Gender Differences in Major Choice (Science Track)

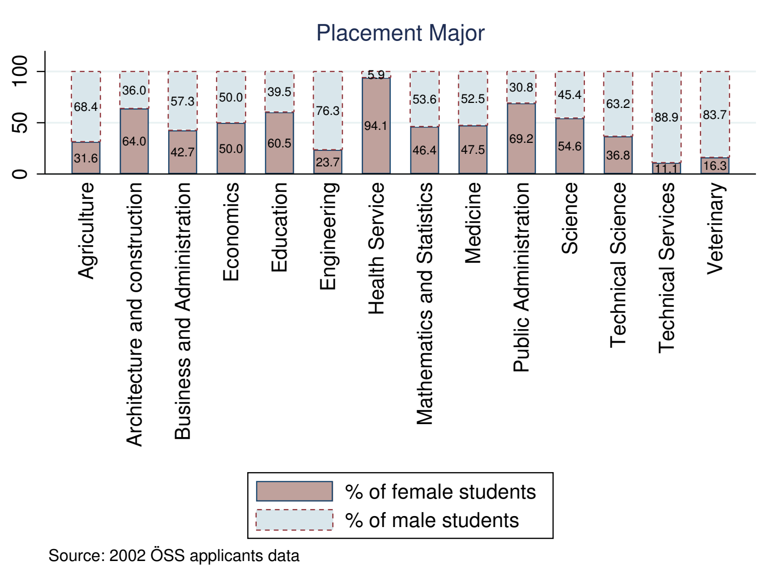


Figure 11: Gender Differences in Major Choice (Turkish-Math Track)

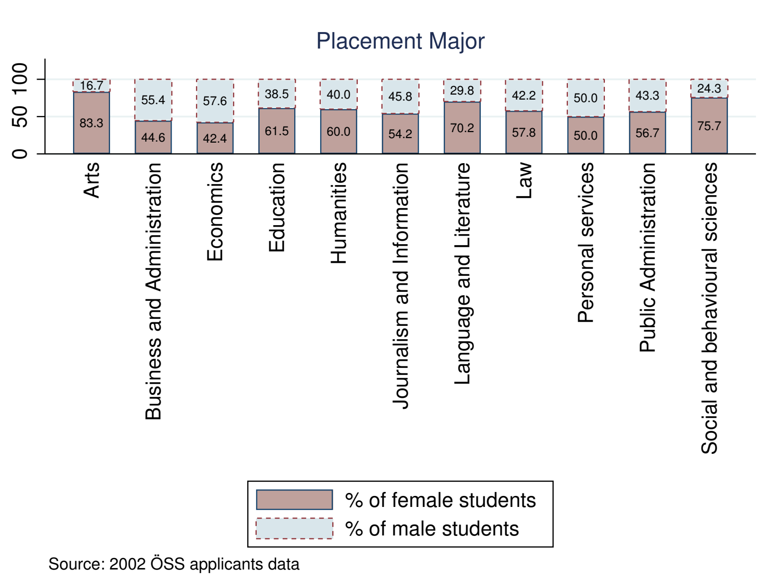


Figure 12: Gender Differences in Major Choice (Social Science Track)

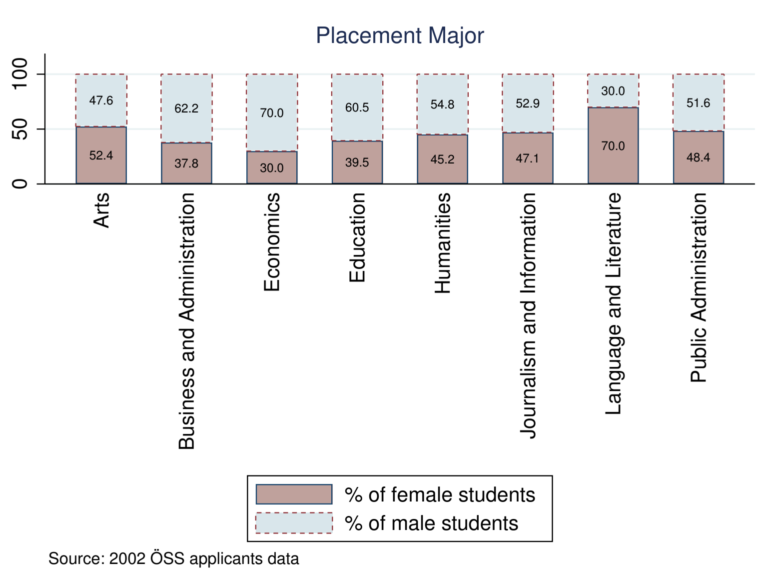


Figure 13: 1st Preference Major (Turkish-Math Track)

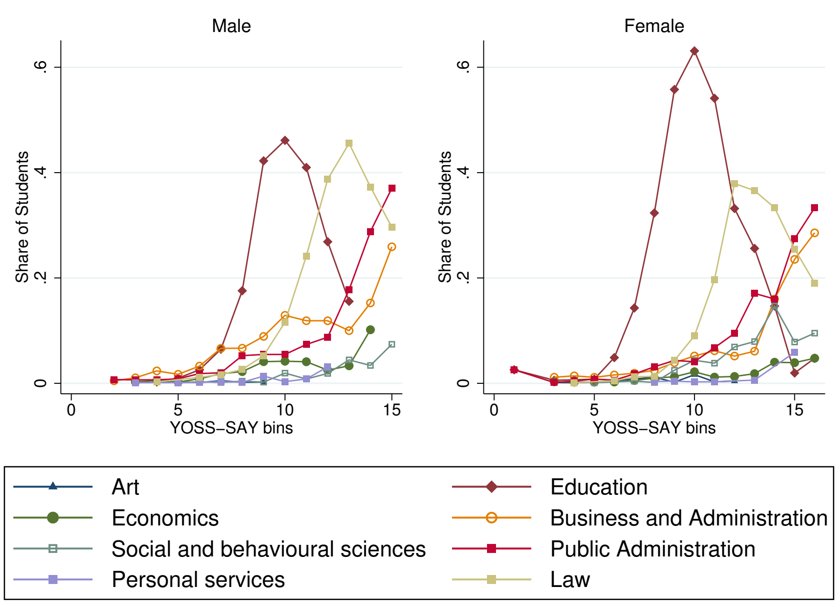


Figure 14: 1st Preference Major (Social Science Track)

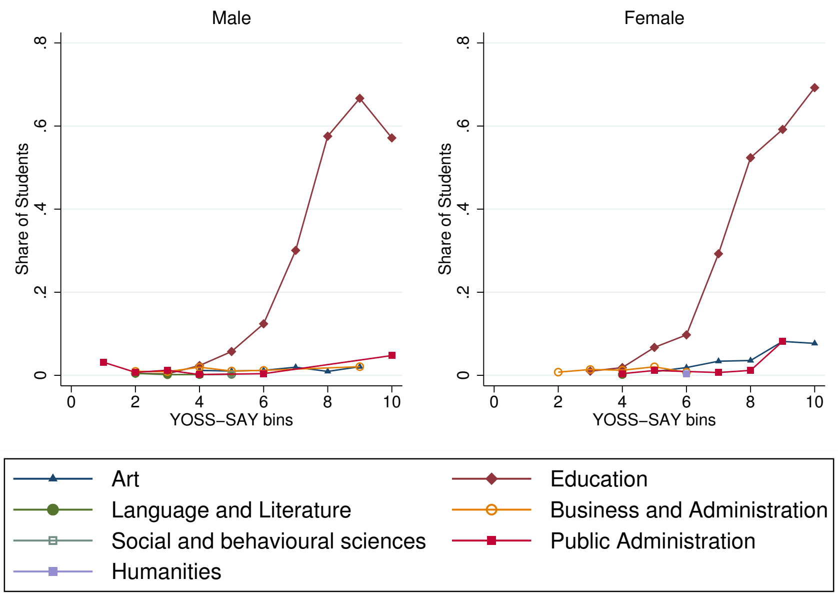


Figure 15: Distribution of Students According to the Share of Dominant Major in Their Preference List

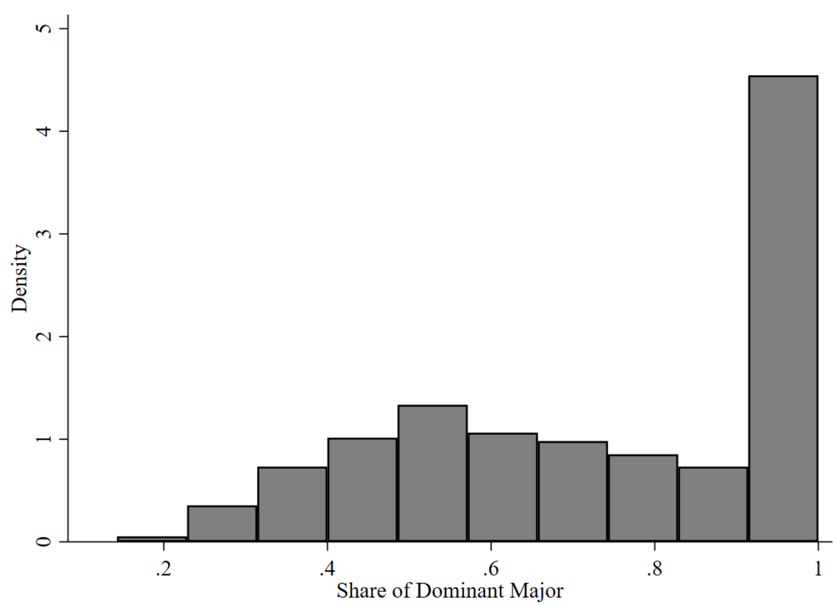
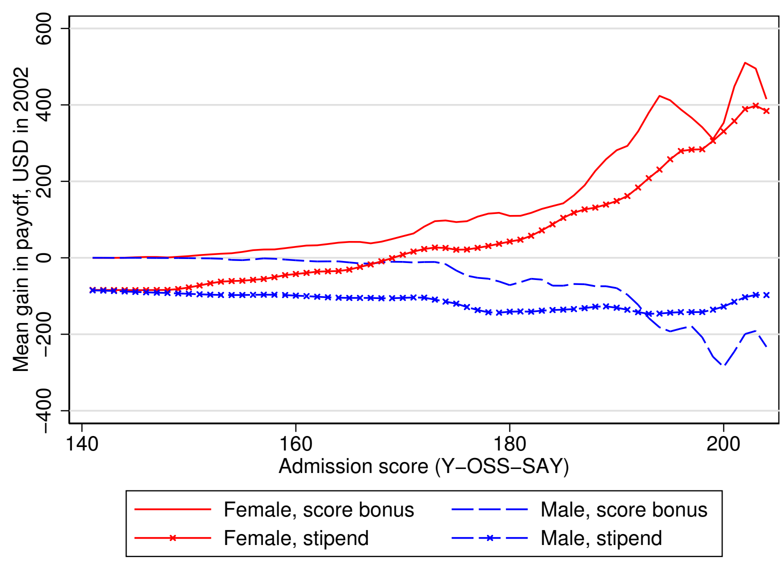


Figure 16: Stipend vs Score Bonus: Expected Changes in Payoffs Net of Tax by Score



Notes: This graph compares two bonus policies: (a) a stipend for female engineering students, (b) a score bonus granting an admission priority for female applicants in engineering programs. Both policies are calibrated to achieve gender equality in admission probabilities to the engineering major and restricted to first-time takers from the academic track of high school. The payoffs are expressed as annual stipend equivalents in the 2002 US dollars net of tax. To finance the stipend policy, a uniform tuition charge is applied to all the admitted students irrespective of gender or major of study.

B Deriving the Likelihood Function

For each student i, we observe the program of placement in 2002, j_{i2}, and the predicted program of placement under the cutoff scores in 2001, j_{i1}, given i's scores and preference list submitted in 2002, s_i and $\mathcal{L}_i$. We also observe whether j_{i1} is ranked above j_{i2} in the student's list L_i.

The likelihood function is defined as the probability of j_{i1} and j_{i2} being ranked in the order given by L_i and being the best choices in the sets of programs ex-post feasible for i in 2001 and 2002, C_{i1} and C_{i2}. Denoting the vector of all parameters as θ, one can express the likelihood function for observation i via a likelihood function conditional on unobserved types:

$$\mathcal{L}_i(\theta; j_{i1}, j_{i2}, L_i, C_{i1}, C_{i2}) = \sum_{t=1}^{T} \sigma_t \mathcal{L}_{it}(\theta; j_{i1}, j_{i2}, L_i, C_{i1}, C_{i2}, t) \tag{3}$$

In what follows, we omit the indices i and t whenever this does not cause confusion. We also use the following notation for the parts of the choice sets: $A_{i1} = C_{i1} \backslash C_{i2}$, $A_{i2} = C_{i2} \backslash C_{i1}$.

Case 1: $j_1 \neq j_2$, $j_1 \succeq j_2$

First, we consider the case in which the choices j_1 and j_2 are different and j_1 is ranked above j_2. This implies that j_1 is the best choice not only in the set C_1, but also in the union of C_1 and C_2. Note that $j_1 \neq j_2$ implies $j_1 \in A_1$ by revealed preference — otherwise, j_1 would be feasible in C_2 and the agent would prefer it to j_2. One can find a closed-form solution for the type- and student-specific likelihood as follows:

$$\begin{aligned}
&\mathcal{L}_t(\theta; j_1, j_2, L, C_1, C_2) = \Pr\{c(C_1 \cup C_2) = j_1, c(C_2) = j_2\} = \\
&\quad = \Pr\{u_{j_1} \geq u_k, u_{j_2} \geq u_l, k \in A_1 \cup j_2 \setminus j_1, l \in C_2\} \\
&\quad = \int \cdots \int I[\varepsilon_k \leq \varepsilon_{j_1} + \delta_{j_1} - \delta_k, k \in A_1 \cup j_2 \setminus j_1] I[\varepsilon_l \leq \varepsilon_{j_2} + \delta_{j_2} - \delta_l, l \in C_2] \prod_j f(\varepsilon_j) d\varepsilon_1 \ldots d\varepsilon_J \\
&\quad = \int \left[\int_{-\infty}^{\varepsilon_{j_1} + \delta_{j_1} - \delta_{j_2}} \prod_{k \in A_1 \backslash j_1} F(\varepsilon_{j_1} + \delta_{j_1} - \delta_k) \prod_{l \in C_2 \backslash j_2} F(\varepsilon_{j_2} + \delta_{j_2} - \delta_l) f(\varepsilon_{j_2}) d\varepsilon_{j_2} \right] f(\varepsilon_{j_1}) d\varepsilon_{j_1}
\end{aligned}$$

$$= \int \left[\int_{-\infty}^{\varepsilon_{j_1}+\delta_{j_1}-\delta_{j_2}} \prod_{l \in C_2 \setminus j_2} \exp(-\exp(-\varepsilon_{j_2} - \delta_{j_2} + \delta_l)) \exp(-\varepsilon_{j_2} - \exp(-\varepsilon_{j_2})) d\varepsilon_{j_2} \right]$$
$$\times \prod_{k \in A_1 \setminus j_1} \exp(-\exp(-\varepsilon_{j_1} - \delta_{j_1} + \delta_k)) \exp(-\varepsilon_{j_1} - \exp(-\varepsilon_{j_1})) d\varepsilon_{j_1}$$
$$= \int \left[\int_{-\infty}^{\varepsilon_{j_1}+\delta_{j_1}-\delta_{j_2}} \exp\left(-e^{-\varepsilon_{j_2}} \sum_{l \in C_2 \setminus j_2} e^{\delta_l - \delta_{j_2}} \right) \exp\left(-e^{-\varepsilon_{j_2}}\right) e^{-\varepsilon_{j_2}} d\varepsilon_{j_2} \right]$$
$$\times \exp\left(-e^{-\varepsilon_{j_1}} \sum_{k \in A_1 \setminus j_1} e^{\delta_k - \delta_{j_1}} \right) \exp\left(-e^{-\varepsilon_{j_1}}\right) e^{-\varepsilon_{j_1}} d\varepsilon_{j_1}$$

One can calculate the inner integral by substituting $z = -e^{-\varepsilon_{j_2}}$:

$$\int_{-\infty}^{\varepsilon_{j_1}+\delta_{j_1}-\delta_{j_2}} \exp\left(-e^{-\varepsilon_{j_2}} \sum_{l \in C_2 \setminus j_2} e^{\delta_l - \delta_{j_2}} \right) \exp\left(-e^{-\varepsilon_{j_2}}\right) e^{-\varepsilon_{j_2}} d\varepsilon_{j_2}$$
$$= \int_{-\infty}^{-\exp(-\varepsilon_{j_1}-\delta_{j_1}+\delta_{j_2})} \exp\left(z \sum_{l \in C_2} e^{\delta_l - \delta_{j_2}} \right) dz$$
$$= \frac{e^{\delta_{j_2}}}{\sum_{l \in C_2} e^{\delta_l}} \exp\left(-\exp(-\varepsilon_{j_1} - \delta_{j_1} + \delta_{j_2}) \sum_{l \in C_2} e^{\delta_l - \delta_{j_2}} \right)$$
$$= \frac{e^{\delta_{j_2}}}{\sum_{l \in C_2} e^{\delta_l}} \exp\left(-e^{-\varepsilon_{j_1}} \sum_{l \in C_2} e^{\delta_l - \delta_{j_1}} \right)$$

Substituting the last line back into the expression for the joint probability yields

$$\mathcal{L}_t(\theta; j_1, j_2, L, C_1, C_2) =$$
$$= \int \left[\int_{-\infty}^{\varepsilon_{j_1}+\delta_{j_1}-\delta_{j_2}} \exp\left(-e^{-\varepsilon_{j_2}} \sum_{l \in C_2 \setminus j_2} e^{\delta_l - \delta_{j_2}} \right) \exp\left(-e^{-\varepsilon_{j_2}}\right) e^{-\varepsilon_{j_2}} d\varepsilon_{j_2} \right]$$
$$\times \exp\left(-e^{-\varepsilon_{j_1}} \sum_{k \in A_1 \setminus j_1} e^{\delta_k - \delta_{j_1}} \right) \exp\left(-e^{-\varepsilon_{j_1}}\right) e^{-\varepsilon_{j_1}} d\varepsilon_{j_1}$$
$$= \frac{e^{\delta_{j_2}}}{\sum_{l \in C_2} e^{\delta_l}} \int \exp\left(-e^{-\varepsilon_{j_1}} \sum_{k \in C_1 \cup C_2 \setminus j_1} e^{\delta_k - \delta_{j_1}} \right) \exp\left(-e^{-\varepsilon_{j_1}}\right) e^{-\varepsilon_{j_1}} d\varepsilon_{j_1}$$

$$= \frac{e^{\delta_{j_2}}}{\sum_{l \in C_2} e^{\delta_l}} \frac{e^{\delta_{j_1}}}{\sum_{k \in C_1 \cup C_2} e^{\delta_k}}$$

The last line is obtained by following the same steps as we used to compute the inner integral.

Case 2: $j_1 \neq j_2$, $j_2 \succeq j_1$

This case is symmetric to the previous one. The conditional likelihood function is obtained from the above formula by changing indices:

$$\mathcal{L}_t(\theta; j_1, j_2, L, C_1, C_2) = \frac{e^{\delta_{j_2}}}{\sum_{l \in C_1 \cup C_2} e^{\delta_l}} \frac{e^{\delta_{j_1}}}{\sum_{k \in C_1} e^{\delta_k}}$$

Case 3: $j_1 = j_2$

In this case, $j_1, j_2 \in C_1 \cup C_2$. Also, j_1 is optimal in C_1 and C_2 if and only if it is optimal in $C_1 \cup C_2$. Thus, the formula boils down to the standard multinomial logit probability:

$$\mathcal{L}_t(\theta; j_1, j_2, L, C_1, C_2) = \Pr\{c(C_1) = c(C_2) = j_1\} = \Pr\{c(C_1 \cup C_2) = j_1\} = \frac{e^{\delta_{j_1}}}{\sum_{k \in C_1 \cup C_2} e^{\delta_k}}$$

C Estimation Details

We estimate the parameters of the model in six sub-populations, defined by gender and three high school tracks: Science, Turkish-Math, and Social Science. Preferences for broad categories of subjects (science vs. humanities) tend to correlate with one's high school track. Preferences may also vary between genders if, for example, certain career paths are incompatible with commonly accepted gender roles.

The set of choice characteristics with common valuation across unobserved types, X_{ij}, includes the following variables:

1. The highway distance between the student's high school and the program's campus.[52] A dummy for the high school and the campus being in the same province.

[52] Obtained from the Directorate of Highways at https://www.kgm.gov.tr/

2. A full set of university dummies and program ranking by the cutoff score in the preceding admission cycle in 2001. These variables control for program quality.

3. Dummies for the type of admission score accepted by the program.

4. Interactions of net tuition with student income dummies capture preference heterogeneity associated with one's income.

The coefficients on the following choice characteristics, Z_{ij}, are allowed to vary across the unobserved student types:

1. A set of dummies for program majors.

2. A dummy variable for the option of not being placed and its interaction with the student's predicted exam score. These terms are meant to serve as a reduced form for the value of retaking the exam in the following year or not attending university at all.

When we implement the maximum likelihood estimator, we are confronted by two practical issues. First, the log-likelihood function in latent class logit models is well-known to have multiple local maxima. Second, latent classes tend to separate in terms of preference for majors. For instance, the estimation algorithm may split the population of students into a latent class that favors medical degrees and never applies for economics and a class that favors economics and never applies to medical programs. This means that the coefficient γ_t on the economics major is nearly minus infinity for the former class, and so is the coefficient on medical majors for the latter one. Moreover, the log-likelihood function is nearly flat for these coefficients, which makes the numerical maximization procedure stop prematurely and produce noisy results. Perfect separation is a well-known issue in estimating latent class discrete choice models.

We tackle the multiple maxima problem in three steps. First, we use the simple multinomial logit instead of the latent class logit to give us the first starting value for the parameter vector β. Second, we set the number of latent classes to the number of majors popular among the students from the sample. The initial values for γ are estimated using simple multinomial logit on the subsample of students who are placed in the respective major; for instance, we run the multinomial logit with no heterogeneity in γ using the subsample of students who are placed in economics, estimate the coefficients on Z_{ij} and use these estimates as a starting value for one of the γ_t's. Third, once we have the starting value for the vector $(\beta, \gamma_1, \ldots, \gamma_T, \sigma_1, \ldots, \sigma_T)$, we generate 100 perturbations of this vector by adding small random shocks to it. We then maximize the log-likelihood function for the fully specified latent class logit model using these 100 random starting values and pick the solution that corresponds to the highest value of log-likelihood. Although we did find that the log-likelihood function has multiple local optima, we could not find visible differences between them in terms of the demand substitution patterns they produce.

To address the preference separation problem, we impose a quadratic penalty on the coefficients β and γ_t:

$$\mathcal{L}_{penalized}(\beta, \gamma) = \mathcal{L}(\beta, \gamma) - \sum_{k} w_{penalty,\beta_k} \beta_k^2 - \sum_{t,l} w_{penalty,\gamma_{tl}} \gamma_{tl}^2$$

The penalty parameters $w_{penalty}$ are set at 0.01 for the coefficients on universities and majors, the main culprits behind the preference separation issue. For all the other coefficients, $w_{penalty} = 0.0001$. One way to view penalized maximum likelihood is that it represents a Bayesian estimator with a vague normal prior. The variance of the prior for a coefficient is inversely related to the penalty placed on this coefficient. This estimator has the usual large-sample asymptotic properties (consistency and normality), and the choice of the weights has vanishing impact on the estimates since the likelihood term $\mathcal{L}(\beta, \gamma)$ becomes dominant on the right hand side as the estimation sample grows in size.[53]

D Predicting Substitution Patterns

We present a "heat map" to illustrate the performance of our model and its competitors relative to the benchmark in Column 5. We first create transition matrices. For each student in a given track of a given gender, we use the associated model to simulate placement. Then, we average the results across all students to produce the transition matrices, which are shown

[53]See Gelman et al. (2014) for more details on large-sample frequentist properties of Bayesian estimators.

in Figures 17 to 19.

In Figure 17, we depict the substitution patterns from the data. The vertical axis depicts the actual major of placement under the 2002 admission cutoffs, while the horizontal axis corresponds to the placements predicted using the preference list of the student but under the cutoff scores in 2001. The programs are ordered in terms of their popularity with the most popular ones at the top. Each colored cell depicts the conditional probability of switching majors, with darker colors representing higher probabilities. The substitution patterns predicted by our preferred model are shown in Figure 18, while the patterns predicted by the models in Columns 2, 3, and 4 of Table 5 are analogously depicted in Figures 19b, 19c and 19d.

Note that our preferred model reproduces the transition matrix for majors quite well. In most cases, students seem to have a strong preference for a specific major, as evidenced by the dark colors on the diagonal: the predicted probability of not switching majors is 91.4% whether we use the fitted model or predict placements using preference lists as given. Programs in education seem to be a backup option for many students, and this is reflected in the fact that whatever major the student was placed in 2002, there is a movement to education with 2001 cutoffs. When our preferred model or its alternatives predict non-negligible switching rates, this usually involves related majors. For instance, economics seems to be a substitute for education, engineering, business, and public administration, subjects that either deal with similar domains or require similar skills.

A feature of the transition matrices that may be puzzling is that they are darker below the diagonal. This comes from the fact that if you are going to switch from one major to another, you are more likely to switch to a popular major than an unpopular one. To draw an analogy to demand for colas, if you were to switch from Coke, you would most likely switch to Pepsi, not RC Cola.

It is hard to see how Figures 17 to 19d differ from one another. To make the differences between the predictions and the data more salient, Figure 20 present a heat map of the differences between Figures 19a to 19d and Figure 17. The solid lines drawn delineate the programs that account for 90% of the placements. The dotted line drawn does the same but for 95% of the placements. Each colored cell represents differences in the transition matrices. White means the differences are close to zero, red shows the difference is positive, and blue shows the difference is negative. Our preferred model (in the upper left panel) overall performs better, as its colors are lighter everywhere than any of the others. More importantly, it does particularly well inside the box delineated by the solid and dashed lines, where most of the action occurs.

E Calculating Placement Score (Y-ÖSS)

The University Entrance Exam placement score (Y-ÖSS) of student i is a function of his ÖSS scores and the weighted normalized high school grade points (AOBP).

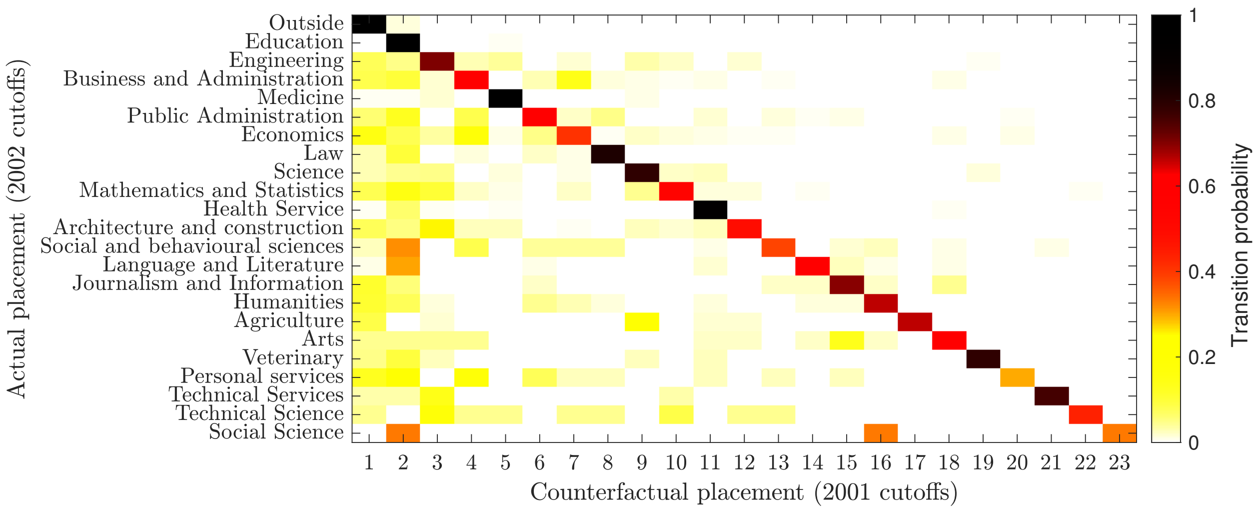
$$Y\_OSS\_X_i = OSS\_X_i + \alpha AOBP\_X_i$$

where $X \in \{SAY, SOZ, EA\}$, and α is a pre-determined constant that changes according to the students' track, preferred department, and whether the student was placed in a regular program in the previous year. The Student Selection and Placement Center (ÖSYM) publishes the lists of departments open to students according to their tracks. When students choose a program from this "open" list, α equals to 0.5. If it is outside the open list, α equals 0.2. If a student graduated from a vocational high school and prefers a department that is compatible with his high school field, α equals 0.65. If a student was placed in a regular university program in the previous year, the student is punished with a 50% penalty to his GPA, so that α equals 0.25, 0.1, and 0.375, respectively in the three above cases.

In turn, the AOBP score of student i from a given track in school j is a function of normalized high school GPA, OBP_j, the minimum and the maximum GPA of the same-school peers, $\min_{i' \in j} OBP_{i'}, \max_{i' \in j} OBP_{i'}$, and the mean ÖSS score in in the respective subject, OSS_X_j, $X \in \{SAY, SOZ, EA\}$, among graduating seniors in that school as in equation (4). Students keep their AOBP over attempts made.

$$AOBP\_X_{ij} = \left[\left(\frac{OSS\_X_j}{80} \times \min_{i' \in j} OBP_{i'} \right) - \left(\frac{OSS\_X_j - 80}{10} \right) \right]$$

Figure 17: Transition Matrix for Majors of Placement, Predicted Using the Preference Data



Notes: Actual major — major of placement in 2002. Counterfactual major — major of placement if the admission cutoffs are the same as in 2001. "Outside" corresponds to not being placed. Counterfactual majors follow the same order as the actual ones (e.g., the label 3 corresponds to engineering). The value in each cell is the mean probability of placement into the counterfactual major conditional on the actual placement. The probabilities are predicted using the preference lists submitted by the students in 2002 and the admission cutoffs from 2001 and 2002.

Figure 18: Transition Matrix for Majors of Placement, Predicted Using the Estimated Model

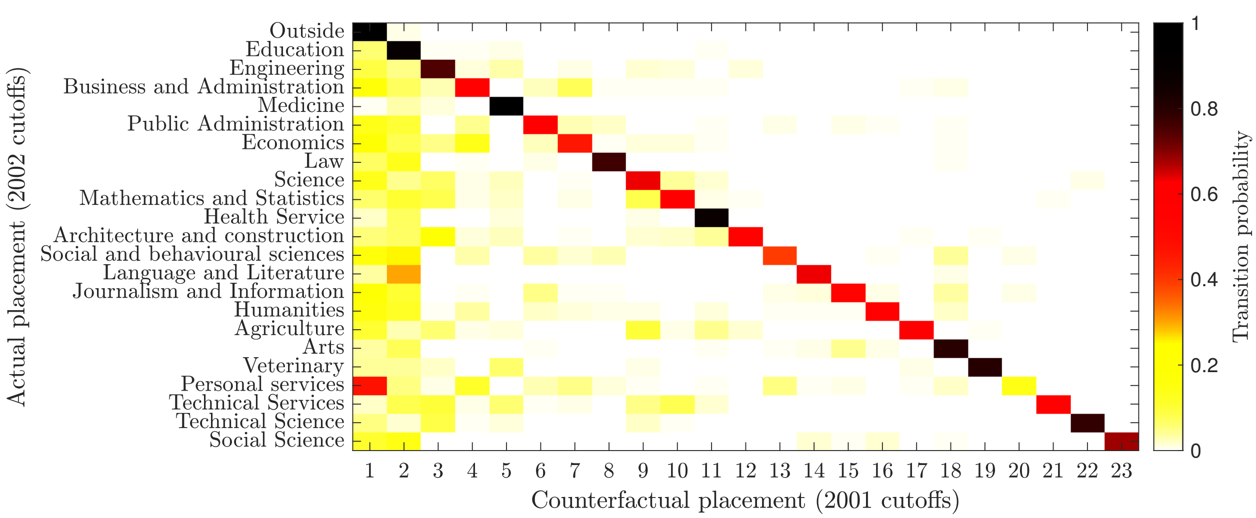
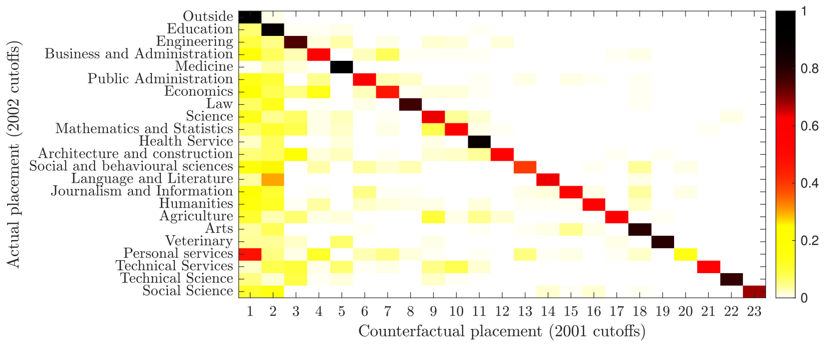
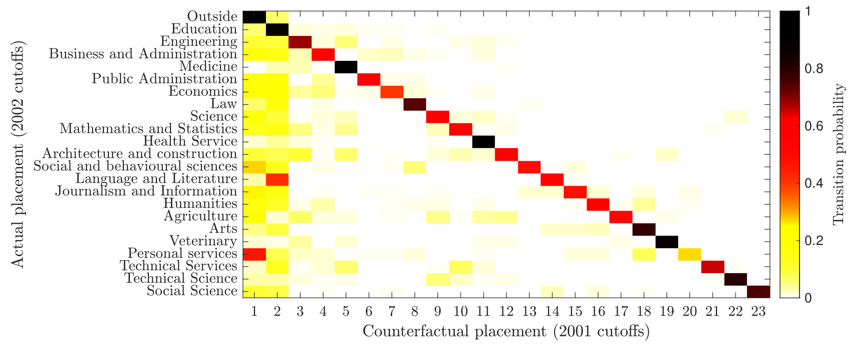


Figure 19: Transition Probabilities for Different Model Specifications

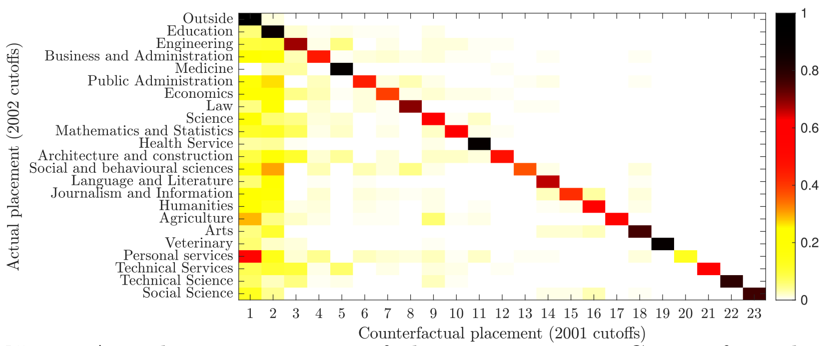
(a) Main Model



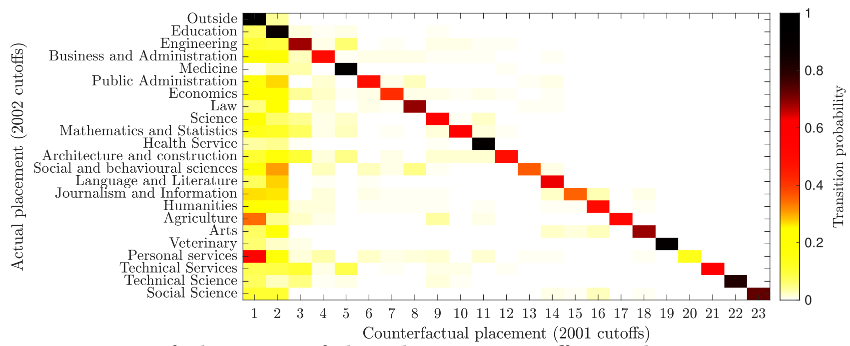
(b) Alternative 1



(c) Alternative 2

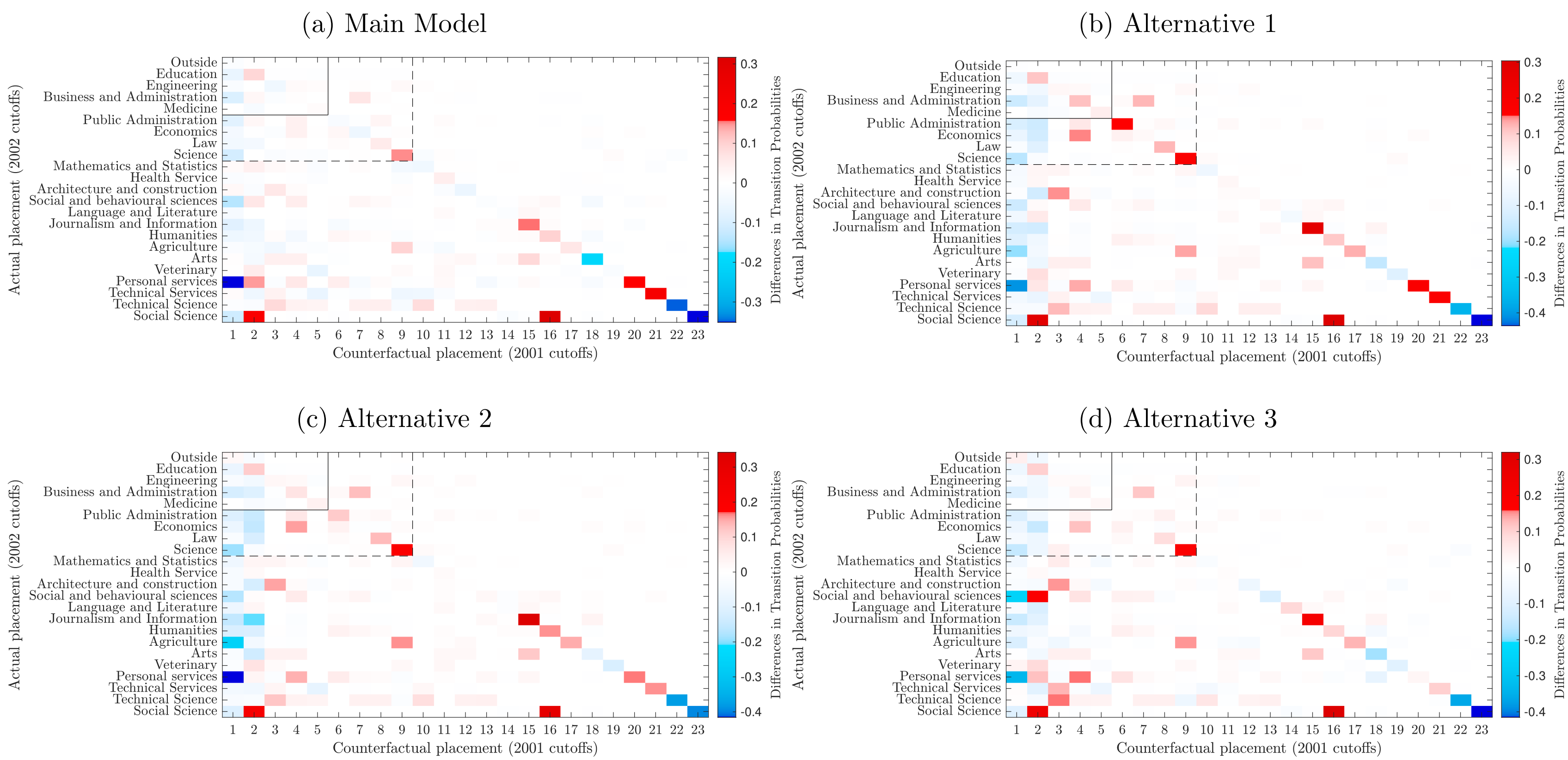


(d) Alternative 3



Notes: Actual major — major of placement in 2002. Counterfactual major — major of placement if the admission cutoffs are the same as in 2001. "Outside" corresponds to not being placed. Counterfactual majors follow the same order as the actual ones (e.g., label 3 corresponds to engineering). The value in each cell is the mean probability of placement into the counterfactual major conditional on the actual placement.

Figure 20: Differences in Transition Probabilities between Model and Data for Different Model Specifications



$$
+ \left[\left(OBP_i \times \frac{OSS\_X_j}{80} \right) - \left(\frac{OSS\_X_j}{80} \times \min_{i' \in j} OBP_{i'} \right) \right]
\times \left[\frac{80 - \left[\left(\frac{OSS\_X_j}{80} \times \min_{i' \in j} OBP_{i'} \right) - \left(\frac{OSS\_X_j - 80}{10} \right) \right]}{\left(\frac{OSS\_X_j}{80} \times \max_{i' \in j} OBP_{i'} \right) - \left(\frac{OSS\_X_j}{80} \times \min_{i' \in j} OBP_{i'} \right)} \right] \tag{4}
$$

We do not observe student AOBP scores in our data set, but we do observe the inputs on the right hand side in (4) other than the minimum and maximum OBP scores in the school.[54] In our sample, we observe the normalized high school GPA (OBP_i) and the raw GPA for all students. ÖSYM calculates OBP as follows:

$$
OBP_i = \max \left\{ 30, \min \left[80, 10 \frac{GPA_i - \mu_{GPA,j}}{\sigma_{GPA,j}} + 50 \right] \right\} \tag{5}
$$

where GPA_i is the students' own GPA, while $\mu_{GPA,j}$ and $\sigma_{GPA,j}$ are the average and the standard deviation of GPA within i's school j. The student's own GPA and OBP are observed in the data. As long as we have at least two students from a given school, we can use equation (5) to solve for $\mu_{GPA,j}$ and $\sigma_{GPA,j}$ for almost almost all the high schools in Turkey.

Since our data set only includes a sample of students, we cannot observe the minimum or the maximum OBP within school for the entire population. To pin down the maximum OBP, we first look at the schools where we have their first-ranking student in our sample (there is a variable that identifies whether the student ranked first or not). In the data set, we observe 445 first-ranked students. This gives us the maximum OBP for 445 schools.

For the remaining schools, we resort to simulations. First, note that the raw GPA is bounded from above by 5. If follows for equation (5) that OBP has an upper bound $\overline{OBP}_j = [80, 10(5 - \mu_{GPA,j})/\sigma_{GPA,j} + 50]$. We assume that OBP scores in each school have a beta distribution with the mean equal to 50, the standard deviation of 10,[55] and the support on $[30, \overline{OBP}_j]$. We find the parameters of this distribution independently for each school. Since the mean and the standard deviation are the same in all schools, parameters

[54] We obtained each school's mean ÖSS scores in each field for the 2002 high school graduates from the OSYM website.

[55] Recall that by definition, OBP is normalized so that its mean within a school is 50 points and the standard deviation is 10 points.

differ in each school only because of the differences in the support of the distribution.

In the second step, we draw from the estimated beta distribution the number of simulated OBP realizations equal to the class size in the school, which is known from the official statistics. We do this S times for each school and then find the average minimum and average maximum OBP over the S draws. We use these averages as our approximation for the minimum and maximum OBP scores:

$$\min_{i \in j} OBP_i = \frac{1}{S} \sum_{k=1}^{S} \min_{i \in j} OBP_i^k$$

$$\max_{i \in j} OBP_i = \frac{1}{S} \sum_{k=1}^{S} \max_{i \in j} OBP_i^k$$

Finally, we match these estimated minimum and maximum OBP scores with our data set. If we observe a lower bound for OBP in our data set than what was simulated, we use it as the minimum OBP for this school. If we observe a higher maximum OBP in the data, we use it as the maximum OBP for this school. Otherwise, we use the simulated minimum and maximum OBP scores.

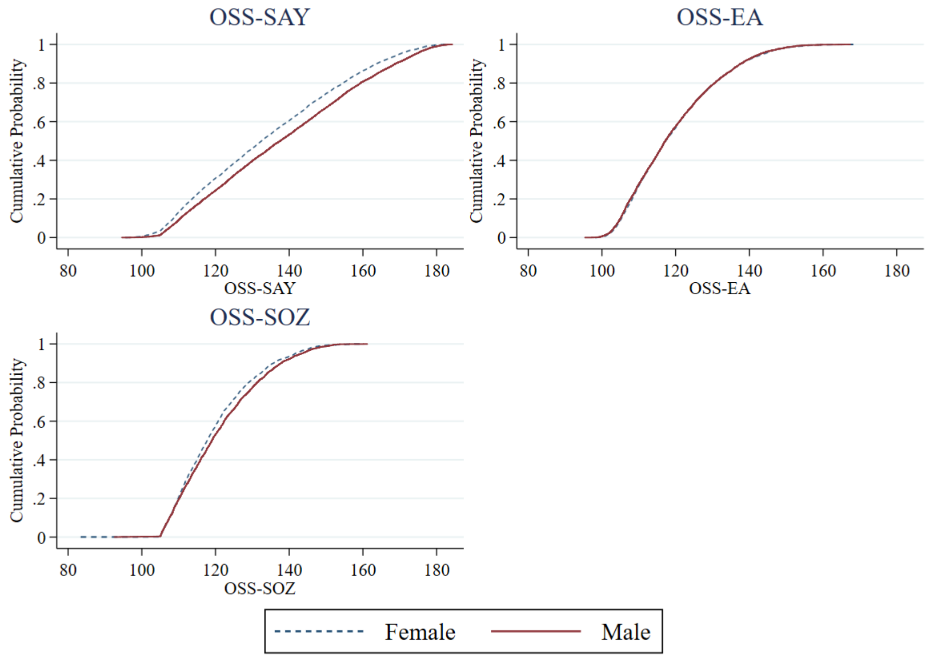
F Score and GPA Distributions

Figure 21 presents the cumulative distribution of exam scores (ÖSS) by gender and high school track. For the score used in the science track programs (ÖSS-SAY) the male students' score distribution (in red) first-order stochastically dominates that of female students. The Kolmogorov-Smirnov test shows this difference is significant. The same pattern holds for ÖSS-SÖZ. On the other hand, for ÖSS-EA, the score usually relevant for students in the Turkish-Math Track, the difference is not as obvious, and the difference in the distributions is not significant (p-value 0.215).

The distributions of high school GPA (AOBP)[56] follow the opposite pattern: women

[56]Since different schools could differ in their grading standards, AOBP score is normalized as explained in Appendix E.

Figure 21: ÖSS Score Distributions by Gender



tend to perform better in school than men do[57] (see Figure 22). Since the placement score (Y-ÖSS) is a mix of the exam score (ÖSS) and the GPA (AOBP), the gap in placement scores is less than that in exam scores.

[57]This pattern, where males do better in high-stakes exams has also been observed in other settings. In a meta-analysis, Voyer and Voyer (2014) show that girls do better than men in high school and have been doing so for quite a while. This pattern is often attributed to women maturing earlier than men.

Figure 22: AOBP Distributions by Gender