Deconstructing the Placement Gender Gap: Performance versus Preferences

Pelin Akyol * Kala Krishna† Sergey Lychagin ‡

October 19, 2023

Abstract

When primary and secondary education is free and of good quality, as well as heavily subsidized at the university level, and admissions are transparent and performance based, one might expect there to be little gender bias in placement at the university level. Yet, the college major choice decisions of students vary considerably by gender. Using Turkish data, we examine what lies behind these differences. Two channels seem to dominate: performance differences and differences in preferences across majors by gender. We estimate a state of the art model of preferences and run counter-factual simulations to evaluate the role of these two channels on the placement gender gap. Finally, we show that policy measures, such as giving women preference in STEM subjects, will not work as well as expected and show that more directed policies are needed.

*Bilkent University, e-mail:pelina@bilkent.edu.tr

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

‡E-mail:lychagin@gmail.com

1 Introduction

In the United States today, over 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 on the grounds that she was female. In Economics among the top 100 US universities, there are more than two men majoring in economics for every woman at the undergraduate level. This fraction 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 for efficiency ones. If intrinsic comparative advantage exists, and there are barriers to entry for women in some fields, large efficiency losses may ensue.[3] Hence, understanding the drivers of gender differences in the choice of college fields is essential for designing effective policies as using different instruments can have dramatically different consequences for the patterns of winners and losers.

The same issues arise when we consider affirmative action by race, caste or ethnicity. It is often argued that such programs end up benefiting the more advantaged rather than the less, which runs counter to the rationale behind the program and creates opposition to them. This is called the "creamy layer problem" in the Indian context where caste based affirmative action is not just endemic but extremely restrictive. More than half the population is now targeted as being scheduled caste (previously called untouchables) or tribes (tribals) or "other backward castes", most of the last being far from disadvantaged. The affirmative action is so restrictive that for top schools score cutoffs for scheduled castes and tribes can be 40 to 50 *percentage points* lower. In this setting, the advantaged are much more likely to qualify creating resentment among upper caste students who see students that

[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 as measured by papers on the program in the NBER summer Institute, also varies substantially. In finance, the share of women is roughly 14.4 percent; in macro & international, it is around 16.4 percent; and in micro, the share is highest, with 25.9 percent of female authors (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.

look much like them unfairly gaining from preferences. This is an issue in the US as well, despite admissions officers basing their decisions on background as well as race. It is argued that affirmative action creates " racial diversity without much economic diversity". An Atlantic article[4], argues that " Seventy-one percent of Black, Latino, and Native American students at Harvard come from college-educated homes with incomes above the national median; such students are in roughly the most advantaged fifth of families of their own race" and that "According to an April 2022 Pew Research Center poll, 74 percent of respondents said race shouldn't be used as even a minor factor in college admissions; majorities of all racial groups opposed such preferences." Thus, the perception that affirmative action polices are "unfair" is a major reason why they are opposed. Despite the recognition that such factors are important, there is, somewhat surprisingly, little quantitative work comparing the affirmative action by point bonuses to the affirmative action created by stipend bonuses.

In this paper we use the Turkish context to shed light on what lies behind gender bias in college placements and what this implies for policy. We choose to look at Turkey because the allocation mechanism is extremely centralized and clear cut. Students list their preferences once they know their scores, and are allocated to their most preferred choice with priority determined by score. Competition is fierce and applicants face considerable stress as a result. There is also significant gender bias in placement. For example, Engineering has over 76% of the students being male. Finally, we have access to administrative data from Turkey at the student level on background, preferences, performance and admission to college programs which allows us to take the model based data driven approach needed to understand the problem.

We start our analysis by examining three potential factors which drive the gender gap in placement: differences in entrance exam scores, differences in preferences, and less aggressive application behavior on the part of women.[5] First, since college seats in Turkey are allocated according to one's placement score, a gender gap in scores could explain the gender gap in

[4]See "The Affirmative Action That Colleges Reall Need" October 26, 2022.

[5]Another channel, documented in Saygin [2016], is retaking: male students tend to be more likely to retake the entrance exam, which potentially improves their exam scores raising the chances of being placed in a good program. While we incorporate differences in retaking by gender into our preference estimation as explained in Section 4, the effects of this channel is not our primary focus here.

placements, especially in the most competitive majors. We do indeed find evidence for such a gender gap. Why might such a score gap occur? It could be that fewer resources are put into their schooling,[6] or that female students under perform in high-stakes environments.[7]

A gender gap in placement could also arise because women's preferences differ. We do find that preferences differ considerably by gender: for example, engineering and technical fields attract few female applicants even after controlling for the entrance exam scores. One reason could be that women tend to avoid highly competitive STEM programs. Women could also have very different preferences for social and cultural reasons or approach the choice of major with the marriage market in mind.[8] Certain fields may be seen as inappropriate for women culturally (veterinary science) or attractive due to being low-pressure and family-friendly (education) even if they are lower paying.[9]

Finally, the gender gaps in placement could arise from women applying more conservatively: that is, they apply to less competitive programs than men, conditional on their exam scores. It has been argued that women are more risk averse and less competitive in a variety of contexts.[10] We show in Section 3.3 that this mechanism is not supported by the data. The difference between the placement score and the cutoff score in the program of placement does seem to be higher for women. However, once we control for major of application, women do not seem to be aiming too low compared to males. This suggests the above gap is explained by differences in preferences rather than competition aversion. For this reason, we do not allow for this channel.

We quantify the importance of preferences versus performance. At the heart of this

[6] Even though we observe gender gaps in performance, we do not find any overt evidence of under-investment by parents or schools into preparing female high school students for competing for college seats.

[7] Taylor [2019] argues that the high stakes testing involved in admission to New York City's Elite Public High Schools disadvantages women. Azmat et al. [2016] use a natural experiment to show that the higher the stakes, the worse the performance of women relative to men. Arenas and Calsamiglia [2020] show that an increase in the stakes of the exam at the end of high school for university enrollment in Spain had a negative effect on female performance, and more so in the exams for which the stakes increase the most.

[8] For instance, see Kirkebøen et al. [2021], and Arum et al. [2008].

[9] Table A.13 in the Appendix shows that, by and large, incomes and the probability of working in Turkey are lower for the kinds of majors women sort into. For example, Teacher training and education pays about 1280 Turkish Lira for women age 25-30 which rises to 1570 at age 40-50. Engineering and engineering trades would given a woman 1420 when young and 2050 when older.

[10] Niederle and Vesterlund [2011], Niederle and Vesterlund [2007a] and Eckel and Grossman [2008] are just a few such examples. Saygin [2016] makes this argument for Turkey.

exercise is a model of college preferences and how these preferences translate into college placement. As explained in detail in Section –, one could use the entire preference list put down by students in estimation. For example, as in Aggarwal (.), one could use the information that the stated preferences are revealed preferred to all other possible lists. This approach rapidly becomes computationally infeasible as the list length and number of alternatives expands. **Larracou (.)** argues that not all alternatives need to be considered, it is sufficient to only replace items one by one. This improves the feasibility of the approach considerably. However, we argue that economically, most alternatives are really not relevant and that it is vital to focus on economically relevant ones, rather than treating all possible alternatives as equally relevant, especially when the list size allowed and the size of the set of alternatives is large[11]. In this we follow Berry et al. [2004] who shows that using data on what consumer's would have bought if their first choice was unavailable improves the ability to match substitution patterns present in the data. Analogously, we make the case that since students are explicitly told about last years cutoffs, their ranking lists should be such that they would obtain their most desired feasible program were cutoffs to be those of the previous year or the realized cutoffs. Our approach thus trades of quality of the data for quantity.

We then design policies aimed at reducing the gender gap in Engineering and simulate the college placement outcomes they induce. We use our estimates to do counterfactual simulations for policy purposes. First we eliminate the gender score gap. We find, contrary to what we expected a priori, that this does little to move them into Engineering, a traditionally male dominated fields. The intuition is simple. Women's preferences differ substantially from those of men. As a result, giving them more points tends to raise the cutoffs in subjects favored by women without reallocating women to Engineering or other STEM majors. Following this, we eliminate the preference gap by giving women the same preferences as men. This does much better, almost doubling the share of women in Engineering. However, it does not fully eliminated the gap. Our findings are in line with those of Arcidiacono [2004] who finds that preferences play a crucial role in student major choices.

[11]The list allowed can include up to 24 programs from the list of feasible alternatives which can be more than 7,000 for students with high scores.

We then compare two policies which reduce the gender bias to a given level in placements in Engineering. The first uses subsidies in terms of stipends, while the second uses subsidies in terms of placement scores for women going into Engineering. Despite a roughly constant trade-off between the two in achieving a given goal in terms of gender bias, we show that the two polices have starkly different outcomes. 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 attained matters for society.

The literature surveyed in Kahn and Ginther [2018] documents gender gaps in performance and the choice of elective courses among the US high school students: males tend to take and complete more math-intensive courses than females do. Boys also seem to have a greater variance in performance which results in a greater fraction of males at the top (and bottom) of the distributions and might help explain the gender gap in STEM. A number of studies attribute gender gaps in placement to student performance in placement tests (Turner and Bowen [1999] is one such example), and to early tracking and the choice of advanced courses in high school (Card and Payne [2021]). In contrast to studies focusing on North America, our research design benefits from the transparent and rigid nature of the Turkish college admission system. Knowing the exact allocation mechanism, the seat quotas in the programs, the priorities of students during the admission, we can credibly simulate placement outcomes as a market equilibrium. This would be much harder to do for the US given the lack of transparency in the college admissions system, especially at the high end.

While our goal here is to show how existing gender gaps in performance and preferences translate into placement outcome in equilibrium, a related line of work focuses on specific mechanisms generating these gaps. A recent literature investigates the role of subjective expectations in preference for major by using direct surveys of US undergraduate students at select institutions. For recent examples of such work, see Wiswall and Zafar [2021], Stinebrickner and Stinebrickner [2014] and Arcidiacono et al. [2020]. Other work looking for specific mechanisms behind the gender gaps in STEM fields include Carrell et al. [2010] who argue that hysteresis may play a role as women are more likely to take STEM courses if their introductory courses in these areas are taught by female professors. A hostile environment for females in the field could be another reason.[12]

The rest if the paper is organized as follows. The next section describes the data, and gives the institutional background on the college entrance system in Turkey. We present reduced form evidence on gender gaps in exam scores, college preferences and competition aversion in Section 3. We then explain how our approach in terms of estimating preferences fits into the literature and show that it does indeed do better than implementable alternatives in terms of matching substitution patterns in the data. In Section 4, we turn to disentangling the impacts of preference and performance gaps on placements. and in Section 5 we use the model to compare counterfactual policies aimed at achieving gender balance. 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.[13] In each track, students study a different curriculum. In their senior year, they take the centralized university entrance exam where their track, GPA, and score in the exam determines their placement score. The university entrance system is highly centralized. Almost every high school senior takes this exam. This exam is conducted by the Student Selection and Placement Center (ÖSYM) once a year. Both high school seniors and past high school graduates can take the exam. Students are free to repeat the exam, but the score obtained in a year can be used only in that year.

This exam includes four tests, Turkish, Social Science, Math, and Science. Students' scores are calculated as a weighted average of their standardized raw scores in each test. For each student, three different scores, Quantitative (OSS-SAY), Turkish-Math (OSS-EA) and Social Science (OSS-SOZ) are calculated. Each score puts more weight on subjects considered

[12]For example, Wu [2018] and Wu [2020] uses textual analysis from the Economics Job Market Rumors website to demonstrate how women are portrayed negatively by men in the profession.

[13]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.

relevant. Each program uses one of these three scores in constructing the placement score. A weighted average of the relevant test score and the high school grade point average (GPA) is used as the placement score (Y-OSS-SAY, Y-OSS-EA, Y-OSS-SOZ) in a program and is the *only* determinant of college admission. Thus, if a student from the science track applies for engineering, their Y-OSS-SAY score would be used, while if they apply for Economics, their Y-OSS-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 OSS 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 own tracks in college.[14]

After the exam, students are informed about their raw scores, weighted scores and placement scores. Students who get at least 120 points in a score type, are eligible to submit preference for all 2-year and 4-year college programs that admit students based on that type of score. Students whose scores are between 105 and 120 are only allowed to submit preference 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. Students are very well informed as they are provided with a booklet with information regarding each program's cutoff admission score in the past year, the rank of the marginal student, the available number of seats, tuition, and the type of the score the program requires.[15] The system was relatively stable in the period 1999–2003 so that it is not unreasonable to think of students having a fairly good idea what their feasible set is.[16]

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] show 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. Students are placed in the order of their scores following the multi-

[14]See Krishna et al. [2018] for details of this process.

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

[16]These admission cutoffs for programs are depicted in Figure A.10. 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.

category serial dictatorship mechanism.[17] 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 which reduce or remove fees. These scholarships are program specific,[18] and are merit, not need, based in contrast to the norm in the US.

2.1 Data

The data used in this study comes from multiple sources. Our main source of data is administrative data on a random sample of 2002 university entrance exam takers and the 2002 University entrance exam candidate survey, which was filled by all students when they are making their application for the exam. This data set includes students' raw test scores in each test, weighted ÖSS test scores, high school ID, track, high school GPA, gender, family background, their ranked preference list, and the college they are assigned, if any. The normalized high school GPA (AOBP) is not directly available, but we are able to recover it by reverse engineering. Details of this process are explained in Appendix F. We have a random sample of around about 40,000 students from each track (Social Science, Turkish-Math, Science).

The second source of data is the booklet published by ÖSYM that includes the minimum cutoff scores and available number of seats in each college programs for the years 2000, 2001, and 2002. This data also includes tuition cost of each department, amount of the scholarship, if provided.[19] In addition, we collected the distance between each of approximately one thousand districts from the General Directorate of Highways.

Summary statistics on first-time takers are presented in Table 1 for each track by gender. Columns 1 and 2 present the means and standard deviations of each variable by gender. Column 3 presents the difference between females and males. Note that ***, **, and * denote that a significant difference at the 1%, 5%, and 10% levels respectively. The same

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

[18] 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.

[19] Tuition cost in public universities does not vary across universities, but it varies according to the major.

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 OSS score relevant for the track (first row) is more prevalent among science track students. The OSS-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 Y-OSS (allocation 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).[20] 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 what is behind the worse performance of women in the university entrance exam. The next row gives the proportion by gender that obtain a scholarship for prep school.[21] 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.[22]

The third group of variables deal 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 for college as expected. The fourth group of variables deal 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 deal with the type of school the students go to. Note that women are not less likely to go to science high schools[23], conditional on finishing high school, but are

[20] Turkey had a hyper inflation 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 Turkish Lira would have been about 1500 US dollars.

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

[22] This is probably because students from non science tracks are very unlikely to get scholarships to begin with.

[23] All students in science high schools are from the Science track, which is why the entries are blank in other tracks.

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 come from the east of Turkey which is seen as being poorer and more conservative than the western part. As expected, the fraction female from the east is significantly less than the fraction 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	(1)-(2)	Mean	Mean	(1)-(2)
VARIABLES	(SD)	(SD)	Diff.	(SD)	(SD)	Diff.	(SD)	(SD)	Diff.
OSS-SAY	134.379	138.586	-4.206***	111.792	112.842	-1.050***	102.331	102.879	-0.548***
	(20.493)	(21.216)		(12.139)	(11.951)		(4.980)	(5.003)	
OSS-EA	127.296	126.987	0.309	119.701	119.479	0.222	110.508	110.403	0.105
	(16.903)	(18.656)		(12.536	(12.538)		(7.472)	(7.628)	
OSS-SOZ	118.607	117.353	1.254***	126.303	125.681	0.623*	119.845	121.050	-1.205***
	(18.404)	(21.535)		(13.573)	(14.077)		(10.934)	(11.538)	
Normalized High School GPA(NHGPA)	55.211	52.005	3.206***	52.758	48.518	4.239***	50.809	48.497	2.312***
	(9.329)	(10.115)		(8.784)	(9.113)		(8.059)	(8.003)	
AOBP_SAY	66.699	63.274	3.425***	62.200	57.933	4.268***	58.223	55.973	2.249***
	(9.076)	(9.535)		(8.695)	(8.508)		8.070	7.818	
AOBP_EA	66.583	63.074	3.509***	62.276	57.948	4.328***	58.486	56.193	2.293***
	(9.084)	(9.572)		(8.623)	(8.477)		7.966	7.734	
AOBP_SOZ	66.49	62.948	3.542***	62.324	57.978	4.346***	58.745	56.437	2.309***
	(9.129)	(9.632)		(8.592)	(8.474)		7.881	7.670	
Prep School Expenditure:									
Missing	0.068	0.078	-0.010*	0.142	0.144	-0.001	0.286	0.276	0.010
	(0.251)	(0.268)		(0.350)	(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	-0.033*
	(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)		(0.307)	(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)	

(continued on next page)

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) (1)-(2) Diff.	Social Science Track (7) Female Mean (SD)	(8) Male Mean (SD)	(9) (1)-(2) Diff.
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	0.438	0.445	-0.007
	(0.425)	(0.437)		(0.466)	(0.470)		(0.496)	(0.497)	
Middle/High School	0.333	0.315	0.018*	0.367	0.342	0.025**	0.345	0.313	0.031*
	(0.471)	(0.465)		(0.482)	(0.475)		(0.475)	(0.464)	
College/Master/PhD	0.324	0.315	0.009	0.204	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	0.168	0.161	0.007
	(0.466)	(0.460)		(0.430)	(0.428)		(0.374)	(0.367)	
Type of the High School:									
Science High Sch.	0.024	0.026	-0.002						
	(0.155)	(0.160)							
Anatolian High Sch.	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 High Sch.	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)	
Type of the Region:									
East 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	13477	6681	5983	12664	2196	2569	4765

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

3 Direct Evidence on Gender Gaps

Our focus is on the gender gap in performance and preferences. We present some direct evidence on this below. It has also been suggested that women tend to do worse in placements because they are less aggressive in applying (see Saygin [2016]). We find no such evidence once one controls for the broad field of application.

3.1 Do Women do Worse in the Entrance Exam?

As is evident from the summary statistics, women do worse in the exam in the science subjects than men. This is more so in the science track. We also present the full distributions of scores for men and women in Appendix G. 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. The GPA distributions by gender and track are reported in Figure G.2.

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. **?**, using the same data we use, show using a novel structural approach that women do seem to be more risk averse than men. Ors et al. [2013] show 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 [2007b] in addition show that in experiments, men choose a tournament compensation system in experiments 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 that "women shy away from competition while men embrace it". Ors et al. [2013] also show that women seem to do better than men in France in undergraduate exams, but worse when it comes to MBA admission exams suggesting that women do worse in competitive exams. Another possible explanation is that girls may be asked to do more chores than boys, and so have less time to spend on their studies.

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 time 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 (OSS-SAY for the Science track, OSS-EA for the Turkish Math track and OSS-SOZ for those in the Social Science track).

The individual level controls, represented by X_{ij}, include factors such as the family background (the parents income group, the level of parents' education) the level of preparedness for the exam (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. Doing so, we are able to account for potential explanations related to parental investment, and high school choice. We also control for preparation for the entrance exam by accounting for prep course expenses, as well as for learning while in high school by including high school GPA in our analysis[24]. Finally, we addressed selection on parental background by controlling for parents' education level. Overall, these controls helped to ensure that our analysis accurately captured 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 underinvestment 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]

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

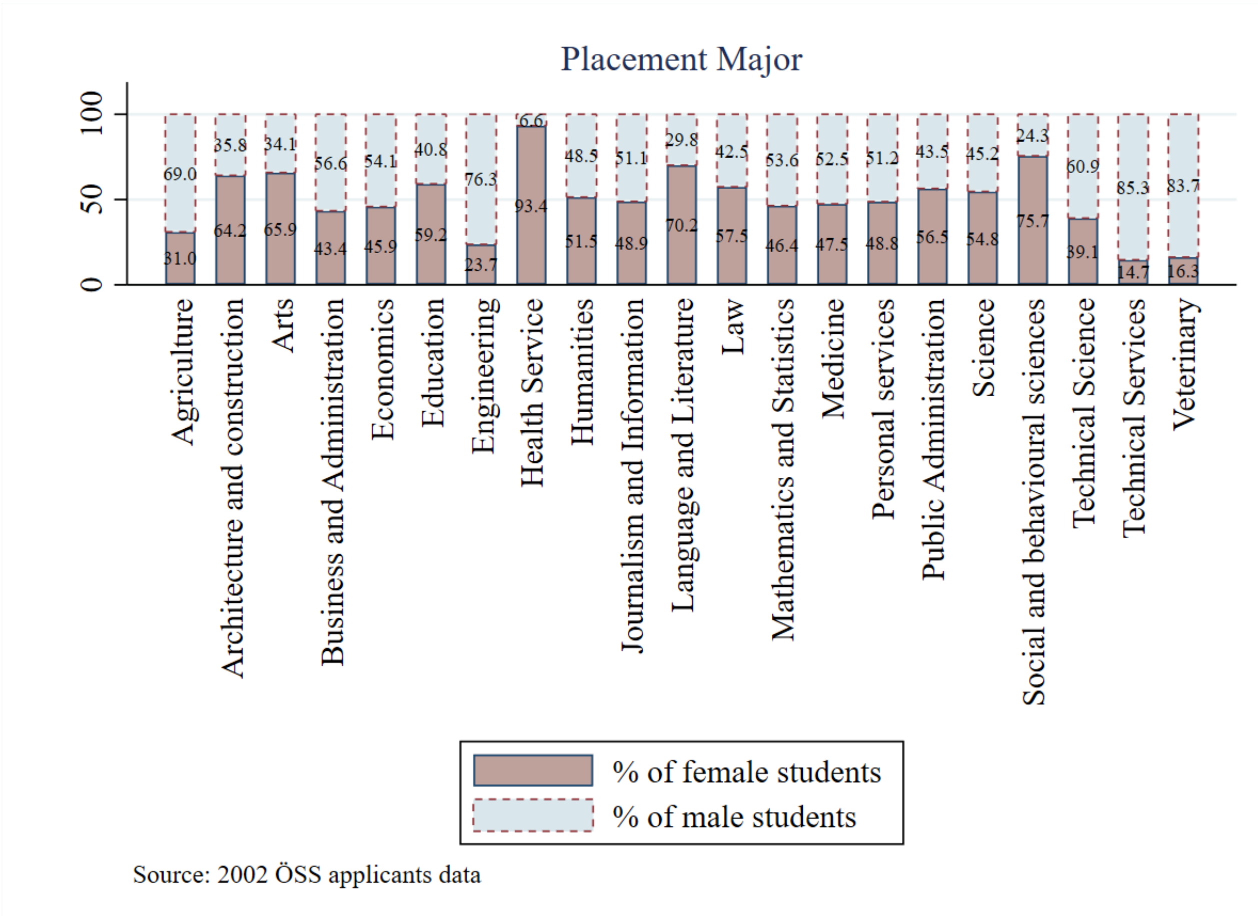
[25]The standard deviations are around 20 points in OSS-SAY for the Science track, 12 points in OSS-EA for the Turkish-Math track and 11 points in OSS-SOZ for the Social Studies track. See Table 1 for more details.

Table 2: Gender Gap in OSS Score

	(1)	(2)	(3)	(4)
	Science Track			
VARIABLES	OSS-SAY Score			
Male	8.909***	9.077***	9.326***	9.188***
	(0.328)	(0.317)	(0.259)	(0.253)
Observations	13,505	13,505	13,505	13,505
	Turkish Math Track			
VARIABLES	OSS-EA Score			
Male	2.990***	3.160***	3.747***	3.571***
	(0.228)	(0.218)	(0.177)	(0.170)
Observations	12,664	12,664	12,664	12,664
	Social Science Track			
VARIABLES	OSS-SOZ Score			
Male	2.610***	2.745***	4.060***	3.795***
	(0.343)	(0.334)	(0.375)	(0.364)
Observations	4,765	4,765	4,765	4,765
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)



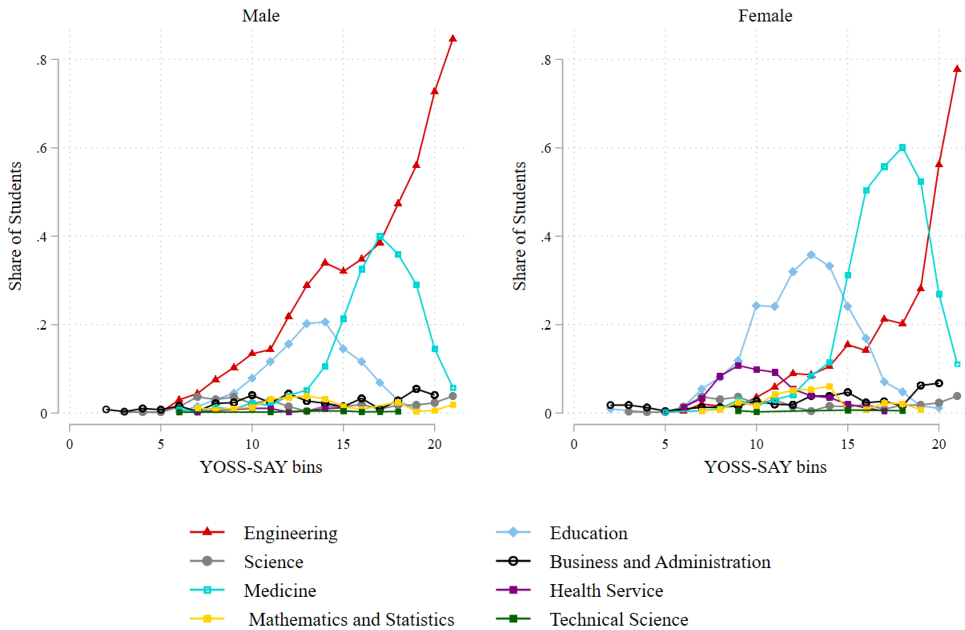
3.2 Are Women's Preferences Different?

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[26] according to placement.[27]. As is evident, there are large differences in 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 visitor) 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.

[26]The subjects that make up these majors are listed in Section E in the Appendix

[27]Figures A.11, A.12, A.13 in the Appendix present the percentage of male and female students in each college major for each of the three tracks separately.

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



These patterns in placements could arise from the difference in scores. For example, women may be under represented 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 by gender as a function of the relevant placement score[28] [29], while Figure A.14 does the same, but according to placement rather than preference. In almost all score bins, male students are more likely than female ones to be placed in and to put engineering programs first on their list. Moreover, the preferences (and placement) of female students varies 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 programs, while those with even lower scores seem to prefer Health Service programs. In addition, women are more likely than men to apply for Medicine at all scores. In contrast, the preference for Engineering programs falls

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

[29]The same graphs for the Social Studies and Turkish-Math track students are presented in Figures A.17 and A.18 in the Appendix.

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?

Do women apply to less selective colleges, perhaps because they are more risk averse or dislike being turned down? Below, we look at the difference in the student's placement score and the cutoff for the school the student was placed in. If women aim lower than men, then this gap should be larger for women than men on average. We show that while this difference is negative and significant on average, 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.[30]

In our analysis, we run the difference in the placement score and the cutoff score in 2001 for the program in which the student was placed on the male dummy and the background controls. 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 effect 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 placed. The effect of this is startling. First, the male dummy 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 dummy is positive and significant for Health Service, Technical Science, Science and Vet Science. This says that all students applying to these

[30]Saygin [2016], using data on Turkey for 2008, runs a regression of the cutoff score for the department placed first on the preference list on gender, characteristics and major fixed effects. She does not control for the student's score. She finds that the cutoff score for the school placed first is not significantly different by gender. However, she finds that the same regression when run on the cutoff score for the school placed last on the list has the dummy for women being negative, suggesting women go further down their list. She connects this to women being more risk averse.

majors tend to be less aggressive. In other words, these majors have a longer right tail in terms of the placement score 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 look 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.[31]

Table 3: Factors affecting difference between Y-OSS Score and Minimum Cutoff (Science Track)

	(1)	(2)	(3)	(4)
VARIABLES	YOSSSAY-Min	YOSSSAY-Min	YOSSSAY-Min	YOSSSAY-Min
Male	-0.580***	-0.684***	-0.676***	0.017
	(0.167)	(0.172)	(0.168)	(0.220)
Income:				
250-500 TL			-0.194	-0.054
			(0.206)	(0.227)
More than 500 TL			0.135	0.476**
			(0.255)	(0.235)
Prep School Expenditure:				
No prep school			1.027	0.763
			(1.362)	(1.382)
Low			-1.622***	-1.496***
			(0.519)	(0.520)
Medium			-1.415***	-1.300**
			(0.523)	(0.507)
High			-0.265	-0.161
			(0.581)	(0.572)
Scholarship			-2.570***	-2.398***
			(0.522)	(0.542)
Parental Education				
Literate			-0.468	-0.992*
			(0.551)	(0.572)
Primary School			-0.099	-0.501
			(0.501)	(0.512)
Middle or High School			-0.238	-0.541
			(0.510)	(0.540)
College/Master/PhD			-0.443	-0.603
			(0.534)	(0.549)
Subject of Major				
Architecture and construction				-0.606
				(1.093)
Education				0.360
				(0.943)
Engineering				-0.822
				(0.926)

(continued on next page)

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

	(1)	(2)	(3)	(4)
VARIABLES	YOSSSAY-Min	YOSSSAY-Min	YOSSSAY-Min	YOSSSAY-Min
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

Standard errors are clustered at the high school city level. *** p¡0.01, ** p¡0.05, * p¡0.1

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

4 Modeling of College Preferences

In order to deconstruct the gap in placements into what comes from performance and preferences, we set up a model of demand for college seats and estimate it. 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. He [2017] estimates household preferences without specifying the distribution of household sophistication types after making certain assumptions. Hwang [2015] set-identifies preferences, assuming certain simple rules on behavior. Agarwal and Somaini [2018] assume that all agents are strategic. Using simple revealed preference arguments, they first pin down the set of preferences that could have generated the observed preference list. They show that a class of mechanisms can be consistently estimated and establish conditions under which preferences are nonparametrically identified. As in the current paper, they exploit the observed assignment outcomes and estimate household preferences without having to solve for the equilibrium. In the application, they estimate a parametric model using data from Cambridge.

However, all of the above work is limited in the size of the choice set that can be dealt with.

Calsamiglia et al. [2015] develops a method applicable to a wide range of choice mechanisms and which allows for a much richer choice set for agents. Their approach also allows them to incorporate different types of agents (eg. strategic and non strategic) and estimates both preferences and the distribution of strategic types in a parametric model. Larroucau and Rios [2018] extends Agarwal and Somaini [2018] by showing that it is sufficient to take a "first order approach", that is, to compare the stated list with lists where only one item is replaced by alternatives ("one-shot swaps") to ensure its optimality. This helps mitigate the "curse of dimensionality", though with over 7,000 options to choose from in our setting, even this approach is problematic to implement.

We propose a novel way of estimating preferences. We extend the ex-post stability approach of Fack et al. [2019] who assume that the system is in steady state and all cutoffs are known (there are so many students that cutoffs do not vary) when preferences are stated.[32] As a result, students put their most preferred feasible (given the current cutoffs) option at the top of their list and only the top listed feasible choice is relevant. This is essentially what Fack et al. [2019] do and they show that their estimates perform well in fitting the data and in generating cutoffs close to the actual ones using data on schools in Paris. However, it is well-documented in the empirical IO literature that identifying complex substitution patterns is extremely hard if one only uses realized choices.[33] Building on this insight we use variations in placement cutoffs over the years which makes other programs on the students preference list relevant. In this way, we match substitution patterns implied by the student rank order lists and small deviations of the admission cutoff scores. We rely on administrative data from the 2002 admissions cycle including student scores and submitted preference lists. We also use realized cutoff scores for each program in 2002, as well as the cutoffs in 2001.

Note that our approach does not involve considering all other alternatives: only those that might be relevant given the likely variation in cutoffs and so is computationally much simpler[34]. We try to strike a balance between quantity and quality of data. On the one

[32]This assumption makes sense in Turkey as there are very many 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.

[33]For this reason, Berry et al. [2004] use data on what consumer's would have bought if their first choice was unavailable.

[34]Our approach also lets us incorporate unobserved heterogeneity in a standard manner.

hand, one could use the entire preference lists. However, there is a literature that suggests that using the entire list might be problematic. For example, students often make obvious mistakes ranking programs.[35] In addition, this approach is computationally infeasible when the number of choices is in the thousands as in our case.[36] On the other hand, a more conservative approach only uses ex-post stability on the premise that doing so relies on data that is much less prone to error. Our contribution on the methodological front is thus a novel 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 innovation in estimation yields considerable rewards (at low cost) as our model performs far better than its competitors in terms of matching the substitution patterns in the data. This suggests that the additional information on preferences we use is instrumental for capturing the substitution patterns correctly.

Our method excludes preference list items that have extremely low probabilities of being reached. Fack et al. [2019] show that such items are unreliable as a source of information on true preferences. At the same time, we do use information on substitution patterns from the preference lists instead of relying just on placement outcomes. It is well documented in the empirical IO literature that identifying complex substitution patterns is extremely hard if one only uses realized choices. For this reason, Berry et al. [2004] use data on what consumer's would have bought if their first choice was unavailable. In the same vein, we argue that since students know last years cutoffs, their ranking lists should be such that they would obtain their most desired feasible program were cutoffs to be those of the previous year or the realized cutoffs. Thus, variation in cutoffs can make the actual placement unavailable or make preferred programs feasible. This gives us variation like that used by Berry et al. [2004] which we use in our estimation.

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

[35]Hassidim et al. [2017], Hassidim et al. [2021], Rees-Jones [2018] and Artemov et al. [2017] provide examples for a variety of settings.

[36]The size of the choice set we consider is roughly 7000 compared to 13 in Agarwal and Somaini [2018] and 1400 in Larroucau and Rios [2018] .

except for a vanishingly small share. Second, we assume that the placement generated by 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 have past minimum admission scores in each program in front of them. 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. This assumption is quite innocuous as a rank order list that does not respect their true order is weakly dominated, see Haeringer and Klijn [2009]. We show that our approach does much better at reproducing the substitution patterns found in the data than alternative ones. It is vital to show that substitution patterns are captured in by the model since if they are not, our counterfactual exercises will be completely 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 fit the data.[37]

4.1 Notation and Identifying Assumptions

Applicants $i = 1, .., I$ choose between programs $j = 1, .., J$. Each applicant has a set of exam scores, s_i, which determines i's priority in the allocation mechanism. Programs are characterized by the major of study, the level of tuition, the distance to applicant's high school and other observable characteristics.

Each student may belong to one of T unobservable types: $t = 1, .., T$. 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.[38] The shares of

[37] For example, see Houde [2012].

[38] Figure A.19 shows density of students according to the share of dominant major in their ranked preference list. It is clear from the figure that students fill their preference list with certain type of majors. This means that there are different types of students: some who say like medicine or engineering and only put such programs on their list, and other types who are more willing to substitute between possibly different subsets of programs. This motivates our allowing for unobserved heterogeneity in our estimation.

types in the population are denoted as σ_t. We use X_{ij} to denote choice characteristics that have the same valuation across the student types.

By choosing program j, the student obtains utility

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

The term ε_{ijt} captures idiosyncratic preferences and is drawn from the standard type-2 Gumbel 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}.

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, student's 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:

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

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 years cutoffs as their best guess about the next years 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 apply 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 C. 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. 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 implementing the maximum likelihood are given in Appendix D.

4.2 Demand Estimates

Table 5 presents the estimates of the 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 not be very competitive; moreover, many of them do not even have binding cutoffs. The next variable in Table 5 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 for these districts being in the same province. Programs geographically remote from the applicant's high school tend to be valued less. 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 1,438 Turkish liras for reducing distance to a program by 1,000 kilometers.[40] A female applicant with the same background would pay 2,291 Turkish liras.[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 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 SOZ) the ranking will differ according to the type of score a program is using. For example,

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

[40]Different programs have different tuitions. Private schools have higher tuition than public ones. In private schools, 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 958 US dollars for males and 1,527 US dollars for females in 2002.

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

a program using the Y-OSS-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 SOZ) would be zero for this program. In order to allow for a flexible mapping from program quality to its cutoff, we also include the square of the ranking. Finally, we also include a variable no rank, which indicates that the program was not ranked in 2001, most likely because it was new. Last but not least, some private university programs are offered at different tuition levels (full, part, and no tuition). Even though these 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 as explained below. This makes sense as students admitted in the same program with different tuition levels attend the same classes. Note that the fact that the same program has different cutoffs depending on the scholarship offered gives us variation used in estimating demand elasticity with respect to tuition.

In order 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. These are laid out in Table 4. 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 Column 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 students list and simulate placements based on this using 2001 cutoffs, not the 2002 cutoffs.

The last row in Table 4 gives the percentage of students who switch majors from their actually 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 major if the cutoffs change from those in 2002 to 2001. The last row of Table 4 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 **(due to increasing vaiance of shocks?)**. Using extra data on choices under the 2001 cutoffs forces 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 is having hard time picking up the correct substitution patterns from the data. Intuitively, the degree of how strong the tastes for majors are is identified by how persistently the 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 are occurring when we use our preferred model or stated preferences. As discussed in detail in Appendix B, our model performs extremely well in matching the substitution patterns coming from the benchmark model. We also show how the alternative models in Table 4, Columns 2-4, fare relative to the benchmark. As expected, they do worse.

Table 4: 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 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.

Table 5: Estimated demand parameters, some common coefficients β

Track	Science		Turkish-Math		Social Science	
Gender	Female	Male	Female	Male	Female	Male
Variable						
Dist program	-6.82	-5.40	-3.73	-1.35	-2.06	-5.62
Distance	-2.93	-1.91	-2.59	-1.76	-2.34	-1.82
Evening program	0.12	0.29	-0.16	-0.03	0.12	0.28
Placement score: EA * norank	-3.28	-3.96	1.12	1.78	1.32	-8.88
Placement score: EA * rank	-19.98	-23.36	0.34	0.99	-3.26	-8.85
Placement score: EA * rank^2	27.83	31.68	6.69	7.53	10.57	12.34
Placement score: SAY * norank	2.92	1.48				
Placement score: SAY * rank	0.14	-6.18				
Placement score: SAY * rank^2	14.81	20.83				
Placement score: SOZ * norank			8.03	5.43	6.62	-0.31
Placement score: SOZ * rank			4.54	-0.61	10.07	-14.57
Placement score: SOZ * rank^2			10.60	14.30	2.34	23.68
Same province	1.22	1.05	1.43	1.39	1.62	1.40
Tuition *Income=1	-12.90	-14.01	-10.28	-10.30	-9.56	-11.76
Tuition *Income=2	-11.09	-10.73	-9.22	-8.46	-10.21	-9.26
Tuition *Income=3	-7.21	-6.83	-4.97	-4.90	-6.47	-6.78

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 Turkish lira/month ("new lira" in 2002), 2 — 250-500 TL/month, 3 — above 500 TL/month.

Units: Tuition — 10,000 TL, distance — 1,000 km, rank — varies from 0 (lowest cutoff among programs accepting the same type of score) and 1 (highest cutoff). No rank – an indicator variable for programs not included in the 2001 ranking.

5 Policies Targeting the Gender Gap in Placements

5.1 Decomposing the Gender Gap: Preferences vs Performance

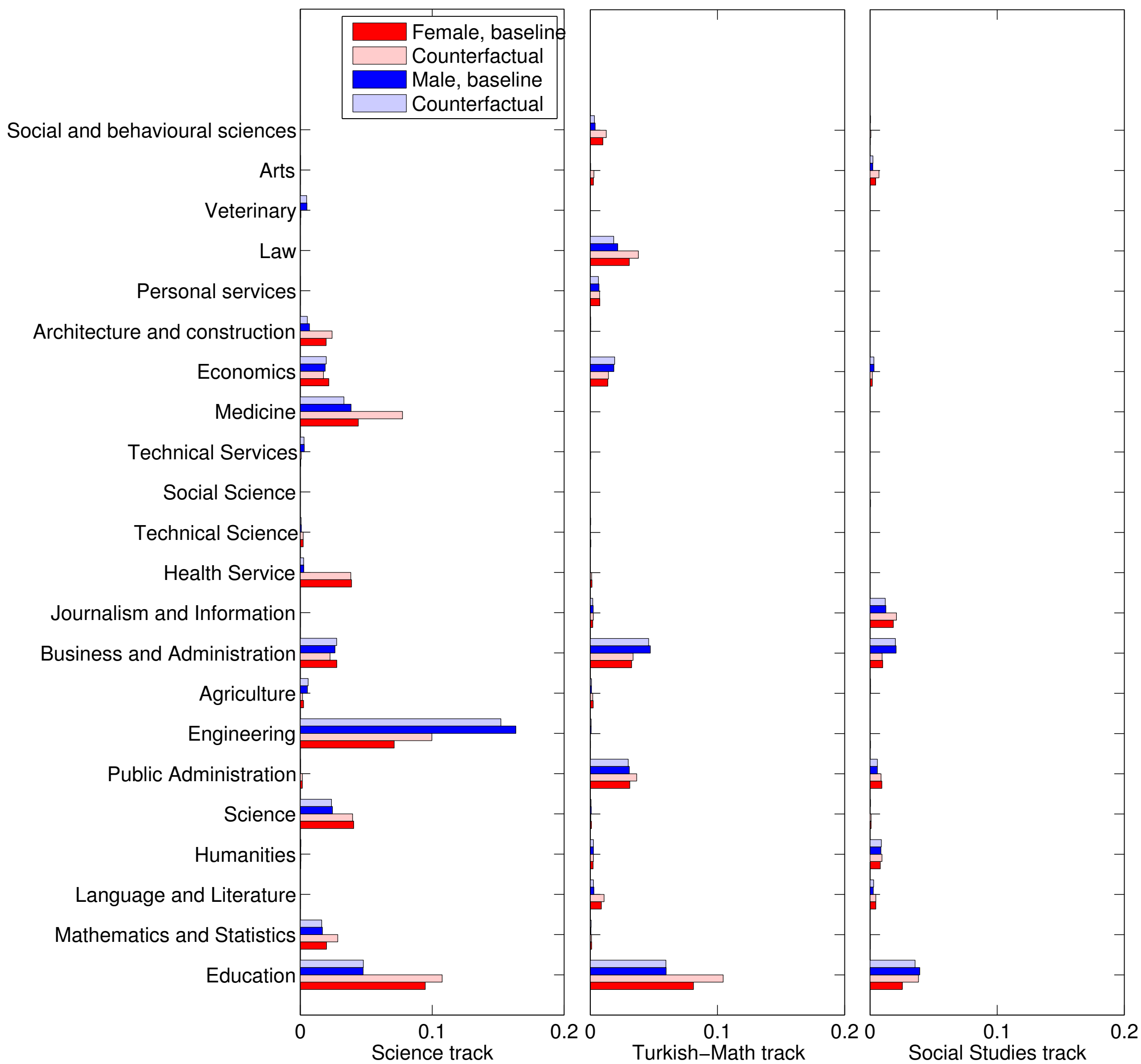
In this section we look at placements by gender under various counterfactual scenarios for students in the three major high school tracks. In each scenario, we manipulate either placement scores or preferences of female applicants. We then use these preferences and scores to simulate the student placement mechanism. First time takers from the academic tracks is our main focus in this exercise. We deal with repeat takers and students from other tracks as a fringe in these simulations. In all exercises, we keep their scores and reported preference lists fixed.[43]

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.[44] 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, for students from the Science and Turkish-Math track, men are much less likely to be placed in education programs as the brightest blue and red bars in each table give the shares under the status quo for males and females respectively. In contrast, for students from the Social Studies track, there is no real difference.

Despite giving female applicants a very generous boost to 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-

[43]Appendix **??** covers the simulations in more detail.

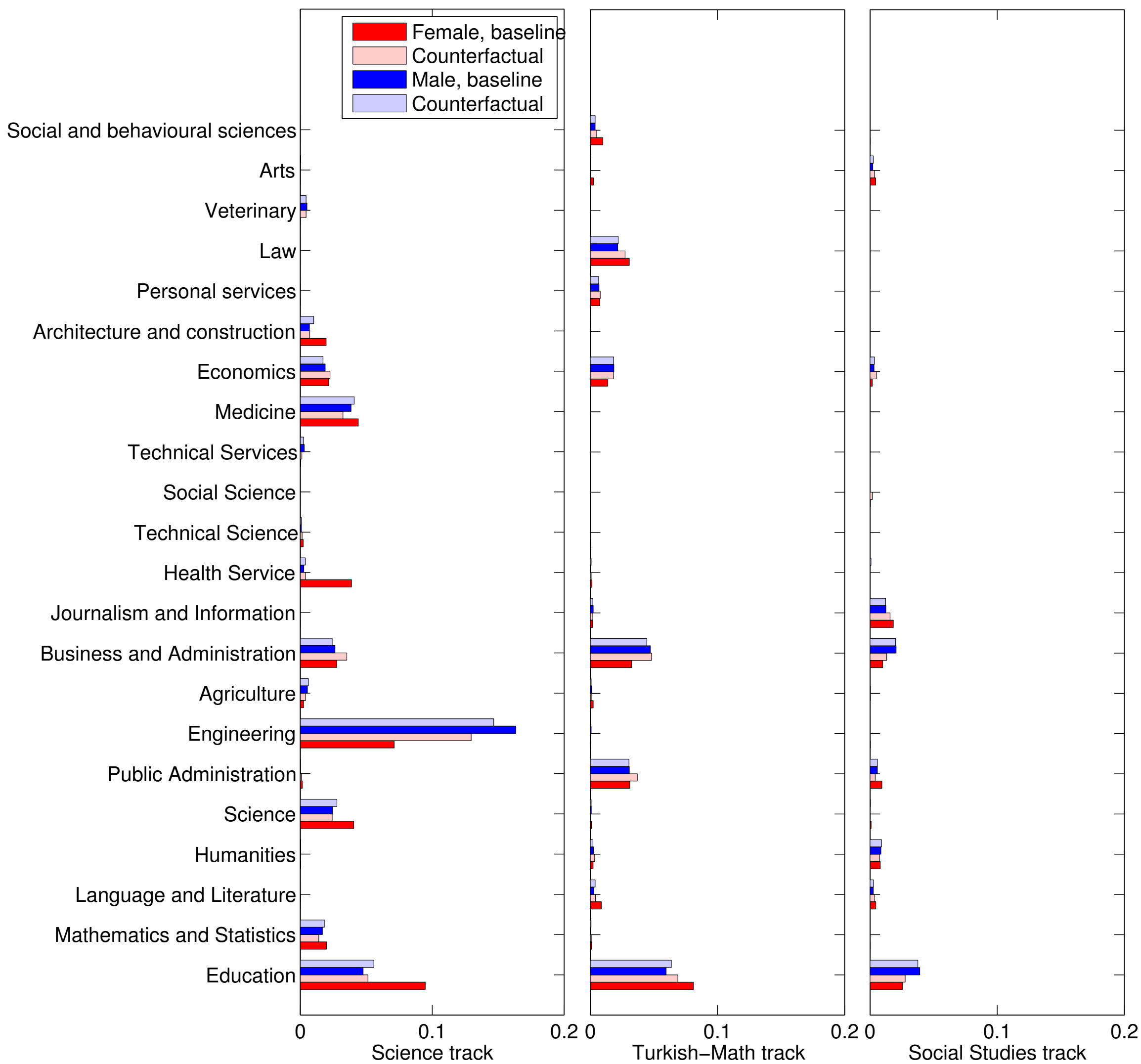
[44]It is worth noting that this bonus would be quite sizable, roughly between one third and one half of the standard deviation of the exam score.



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.

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 replacing preference parameters for females with those of males.

Figure 4: Eliminating the gender gap in preferences.

dominated major.

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.[45] 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.[46] Compared to male students, females have slightly lower chances of getting into competitive majors such as medicine and engineering. This should not be surprising: in the second counterfactual scenario, the existing gender gap in scores gives male students an upper hand.

5.2 Policies Targeting Gender Ratios in Engineering

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

Granting 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. One such policy grants extra score points to females whenever they are considered for admission in 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 stipend 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.

Under this set of policies, every engineering program offers a bundle of an extra stipend

[45] For example, for females from the science track, we use the parameter values in the second column of Table 5 instead of those in the first column.

[46] 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.

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:[47]

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

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

Figure 5 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. Offering the stipend of 2150 Turkish liras per year (roughly 1400 U.S. 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 similar outcomes. These are sizable numbers. The shapes of the policy isolines suggest that stipends and score bonuses are nearly 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

Stipends and score bonuses affect different parts of the student population. To better understand the tradeoffs, 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. To ensure that both policies are budget-neutral, we assume that the stipends are financed by levying tax on all first-time applicants in the science track.[48]

Figure G.3 shows the expected welfare change over the status quo for both policies as a function of admission score and gender. The score bonus policy is improving welfare for females and reducing that for males. The gains and the losses are especially high at the

[47] 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").

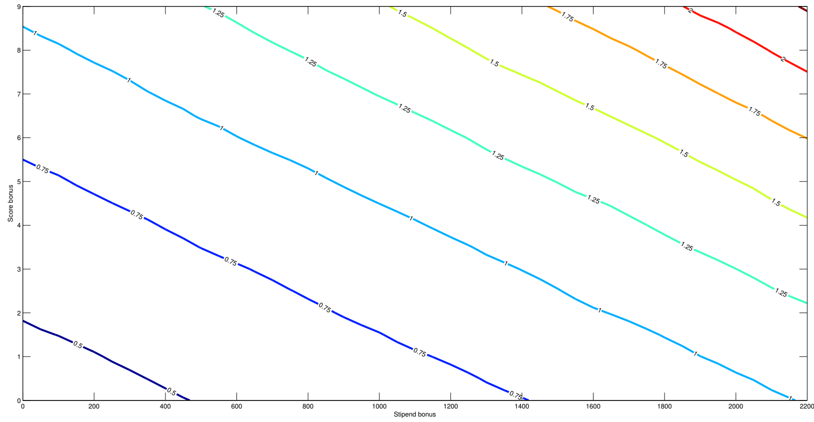
[48] A tax slightly below 100 US dollars in yearly stipend equivalent (i.e., 100 USD paid annually for 4 years) would be sufficient to finance the extra stipends introduced by the policy.

upper end of the score distribution as higher scoring students are more likely to apply to engineering programs. In contrast, low scoring students are unlikely to benefit or lose from the policy as engineering programs are typically beyond their reach even with the bonus applied. In contrast to the score bonus, the stipend bonus policy has a strong regressive effect: only higher-scoring females are in position to win the stipend, but all students have to pay the stipend tax.

While the score and the stipend bonus policies look very similar in terms of aggregate gains and losses, they work via different channels. Figure 7 presents mean gains in student welfare caused by the stipend bonus policy and their decomposition by student groups defined by major choices. Under this policy, female student gain mostly from getting the stipend without changing the program choice, switching from non-engineering to engineering and, at the very top, from upgrading within non-engineering majors due to the fall in competition in high-ranked medical programs. Males lose uniformly from paying the "stipend tax", while those at the top 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, while those who prefer medicine upgrade their choices due to the reduction in competition, similar to females with strong taste toward non-engineering majors.

Figure 8 decomposes welfare gains under the score bonus policy. Females mostly take advantage of the bonus by upgrading their choices within the engineering field or by switching towards engineering from the other fields. As before, some females not interested in engineering are better off from having less competition in their fields. The lion's share of payoff changes for male students comes from downgrading their engineering choices.

As the above figures make clear, both policies have features leading to waste. For the stipend policy, the main source of waste are inframarginal students — that is, female applicants whose program choices are not affected by the policy, but whose stipends have to be paid according to the rules. The tax used to finance these stipends falls disproportionately on males and lower-scoring females for whom engineering is out of reach. Likewise, the score bonus policy displaces many male students to accommodate females who would choose engineering even without a bonus.

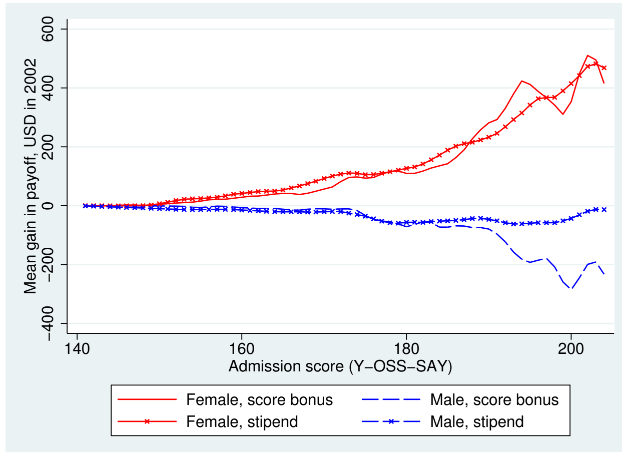


Notes: Each point (x, y) represents a counterfactual policy in which every engineering program offers a stipend bonus of x Turkish liras per year and adds y extra points to the entrance exam score for every eligible female applicant. The graph plots odds ratio isolines, i.e., all stipend-score bonus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

Figure 5: Female-to-male odds ratio of being admitted to an engineering program among science track students

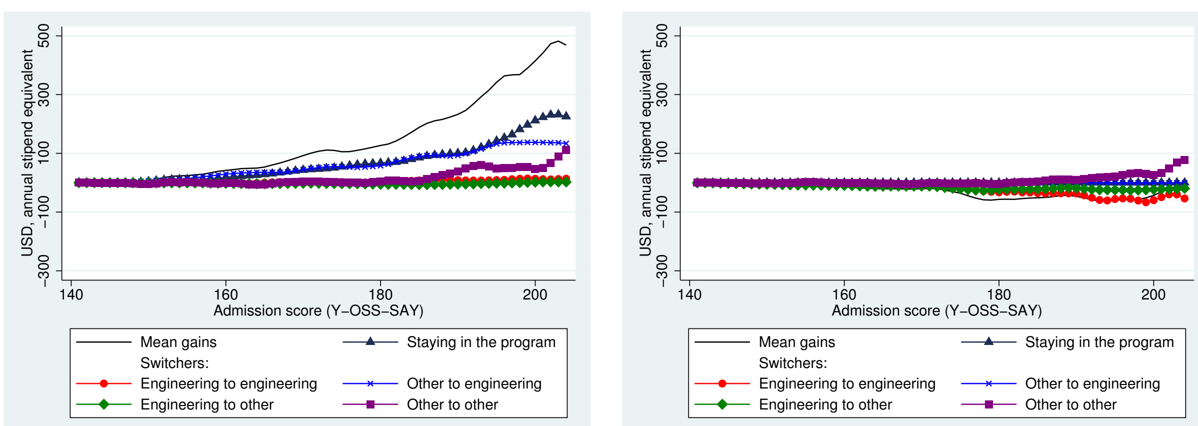
6 Conclusion

This paper has four main conclusions. First, that the main drivers of the placement gender gap are performance and preferences, not any lack of aggression in applications on the part of women. Second, that while differences in performance can account for a small part of the placement gender gap, differences in preferences are more important. Third, giving bonus points to all women is less effective in targeting the placement gender gap than giving directed preferences. In future work we hope to explore in more detail what drives these differences in both performance in the university entrance exam and in preferences by gender.



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.

Figure 6: Expected changes in payoffs by score

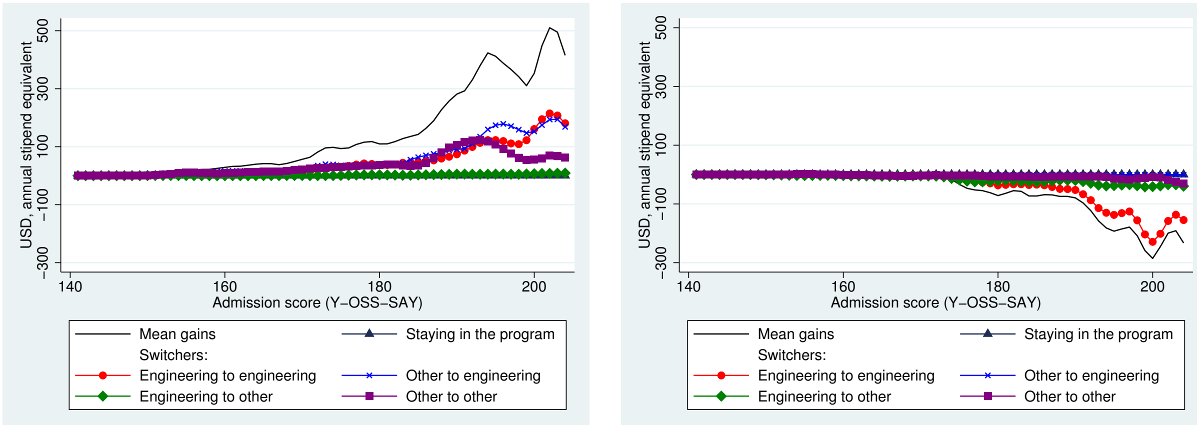


(a) Females

(b) Males

Notes: Mean welfare gains after introducing the stipend bonus policy to achieve full gender equality in admissions to the engineering major. The gains are depicted with the solid lines. Dotted lines depict the part of mean gains attributed to student 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 mean gains.

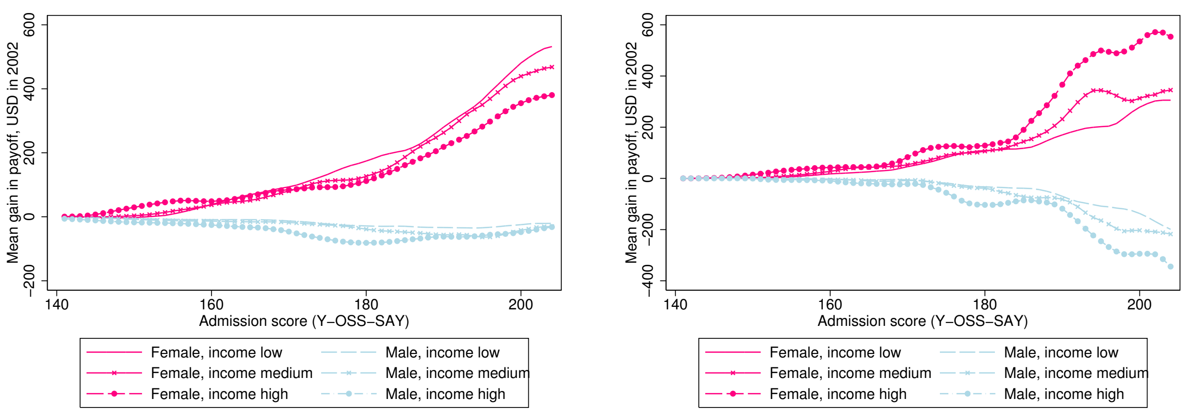
Figure 7: Mean welfare gains by score under the stipend bonus policy: main channels



(a) Females (b) Males

Notes: Mean welfare gains after introducing the stipend bonus policy to achieve full gender equality in admissions to the engineering major. The gains are depicted with the solid lines. Dotted lines depict the part of mean gains attributed to student 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 mean gains.

Figure 8: Mean welfare gains by score under the score bonus policy: main channels



(a) Stipend Bonus (b) Score Bonus

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.

Figure 9: Mean welfare gains by score and income group

References

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

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

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

Peter Arcidiacono, V. Joseph Hotz, Arnaud Maurel, and Teresa Romano. Ex ante returns and occupational choice. *Journal of Political Economy*, 128(12), 2020.

Andreu Arenas and Caterina Calsamiglia. The design of university entrance exams and its implications for gender gaps. *Unpublished Manuscript*, 2020.

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

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

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

Michel Balinski and Tayfun Sönmez. A tale of two mechanisms: student placement. *Journal of Economic theory*, 84(1):73–94, 1999.

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

C. Calsamiglia, C. Fu, and M. Güell. Structural estimation of a model of school choices: The boston mechanism versus its alternatives. *Journal of Political Economy*, 128(2), 2015.

David Card and A. Abigail Payne. High school choices and the gender gap in stem, 2021.

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

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

Anusha Chari and Paul Goldsmith-Pinkham. Gender representation in economics across topics and time: Evidence from the nber summer institute. Technical report, National Bureau of Economic Research, 2017.

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

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

Uri Gneezy, Muriel Niederle, and Aldo Rustichini. Performance in competitive environments: Gender differences. *The quarterly journal of economics*, 118(3):1049–1074, 2003.

Guillaume Haeringer and Flip Klijn. Constrained school choice. *Journal of Economic theory*, 144(5):1921–1947, 2009.

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

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

Y. He. Gaming the boston school choice mechanism in beijing. 2017.

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

Chang-Tai Hsieh, Erik Hurst, Charles I Jones, and Peter J Klenow. The allocation of talent and us economic growth. *Econometrica*, 87(5):1439–1474, 2019.

S. Hwang. A robust redesign of high school match. 2015.

Shulamit Kahn and Donna Ginther. *The Oxford Handbook of Women and the Economy*, chapter Women and STEM. Oxford Handbooks, Oxford, 2018.

Lars Kirkebøen, Edwin Leuven, and Magne Mogstad. College as a marriage market. *NBER Working paper*, 2021.

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

Tomás Larroucau and Ignacio Rios. Do “short-list” students report truthfully? strategic behavior in the chilean college admissions problem. Technical report, Technical report, Working paper, 2018.

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

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

Muriel Niederle and Lise Vesterlund. Do women shy away from competition? do men compete too much? *The quarterly journal of economics*, 122(3):1067–1101, 2007b.

Muriel Niederle and Lise Vesterlund. Gender and competition. *Annu. Rev. Econ.*, 3(1): 601–630, 2011.

Evren Ors, Frédéric Palomino, and Eloic Peyrache. Performance gender gap: does competition matter? *Journal of Labor Economics*, 31(3):443–499, 2013.

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

Perihan Ozge Saygin. Gender differences in preferences for taking risk in college applications. *Economics of Education Review*, 52:120–133, 2016.

Ran I Shorrer. Simultaneous search: Beyond independent successes. Technical report, Pennsylvania State University, 2019.

Ralph Stinebrickner and Todd R. Stinebrickner. A major in science? initial beliefs and final outcomes for college major and dropout. *Review of Economic Studies*, page 426–472, 2014.

Jonathan Taylor. 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), 2019.

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

Daniel Voyer and Susan D Voyer. Gender differences in scholastic achievement: a meta-analysis. *Psychological bulletin*, 140(4):1174, 2014.

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

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

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

A Appendix

A.1 Additional Tables and Figures

Table A.6: Factors affecting difference between Y-OSS Score and Minimum Cutoff (Turkish Math Track)

VARIABLES	YOSSEA-Min	YOSSEA-Min	YOSSEA-Min	YOSSEA-Min
Male	-0.924***	-1.126***	-1.240***	-0.223
	(0.263)	(0.294)	(0.340)	(0.242)
Income:				
250-500 TL			-0.290	-0.002
			(0.291)	(0.236)
More than 500 TL			0.763	1.313**
			(0.588)	(0.526)
Prep School Expenditure:				
No prep school			-0.741	-0.005
			(1.113)	(0.822)
Low			-1.205	-0.482
			(0.808)	(0.538)
Medium			-2.517***	-1.316**
			(0.851)	(0.574)
High			-1.788*	-0.552
			(0.955)	(0.639)
Scholarship			-0.643	0.245
			(1.116)	(0.792)
Parental Education				
Literate			0.724	0.068
			(1.005)	(0.921)
Primary School			0.017	-0.393
			(0.699)	(0.616)
Middle or High School			-0.010	-0.260
			(0.451)	(0.540)
College/Master/PhD			-0.142	-0.115
			(0.512)	(0.712)
Subject of Major				
Economics				-1.301***
				(0.352)
Education				6.072***
				(0.419)
Humanities				0.278
				(1.044)
Journalism and Information				5.672*
				(3.322)
Law				3.493***
				(0.450)
Personal services				-0.483
				(0.580)
Public Administration				0.481
				(0.349)
Social and behavioural sciences				-0.301
				(0.609)
Other				13.116***
				(4.148)
Observations	2,004	2,004	2,004	2,004
High School City FE	NO	YES	YES	YES

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

(continued on next page)

	(1)	(2)	(3)	(4)
VARIABLES	YOSSEA-Min	YOSSEA-Min	YOSSEA-Min	YOSSEA-Min

Standard errors are clustered at the high school city level

Table A.7: Estimated demand parameters, type-specific coefficients and type shares, science track, female

Placement score: SAY	11.13	-14.36	-8.28	-7.51	-3.41	-6.21	-2.22	-3.10
Placement score: EA	7.70	-3.97	-4.55	-2.25	-2.64	-6.89	1.86	2.53
Major: Agriculture	4.92	-0.06	-2.79	-3.13	-4.45	-4.84	-4.93	-7.75
Major: Architecture and construction	7.50	-2.11	0.56	-0.00	-3.27	0.52	-5.12	-3.83
Major: Business and Administration	1.21	-0.35	-0.17	2.95	-1.73	4.07	-1.00	-6.68
Major: Economics	1.14	-4.22	0.46	2.81	2.50	3.41	-1.47	-7.55
Major: Engineering	-2.05	3.30	0.28	0.30	-0.75	0.89	-4.69	-5.65
Major: Health Service	6.10	-0.25	-3.95	-3.24	4.50	-4.73	-3.07	-2.32
Major: Mathematics and Statistics	3.93	-1.20	0.22	0.06	-0.92	-0.24	0.08	-6.45
Major: Medicine	5.32	3.33	2.17	-0.74	7.18	-3.98	2.50	2.33
Major: Science	2.79	-1.46	0.78	-4.60	0.99	-0.28	-0.08	-5.42
Major: Other	1.19	-1.27	-0.50	0.36	-3.63	0.80	-7.61	-10.59
Outside*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

Table A.8: Estimated demand parameters, type-specific coefficients and type shares, science track, male

Placement score: SAY	38.53	-2.61	-5.55	-10.90	-7.08	-7.16	-6.37	-3.65
Placement score: EA	24.35	-0.64	-3.24	-7.37	-5.07	0.30	-2.06	1.05
Major: Agriculture	2.03	0.72	0.59	0.07	-1.48	-1.60	-3.19	-6.86
Major: Architecture and construction	-0.99	3.57	-5.25	0.14	-2.39	-1.54	-3.98	-4.56
Major: Business and Administration	1.58	3.29	-0.95	3.89	2.17	-0.23	3.42	-4.89
Major: Economics	2.28	-2.48	-0.55	2.88	2.19	0.66	3.43	-7.56
Major: Engineering	-0.55	2.49	-0.71	2.38	2.48	4.50	0.21	-4.35
Major: Health Service	-3.64	-3.71	0.06	-2.03	-3.81	1.77	-3.95	-3.53
Major: Mathematics and Statistics	-2.92	-0.37	-1.05	1.27	-3.34	6.10	-0.87	-3.29
Major: Medicine	17.96	-1.95	3.44	2.21	0.46	5.39	0.05	1.58
Major: Science	-1.50	-0.32	0.22	0.02	-4.27	5.39	-1.26	-5.13
Major: Technical Science	-2.33	-2.24	-1.78	2.33	-3.88	3.48	-3.91	-4.86
Major: Technical Services	-0.23	-1.51	-0.10	-2.40	-1.99	7.28	-3.13	-2.74
Major: Veterinary	6.48	2.58	1.96	0.58	-0.23	4.17	-1.89	-1.12
Major: Other	3.34	-1.94	-0.20	-5.11	-3.97	-2.73	-1.34	-4.94
Outside*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

Table A.9: Estimated demand parameters, type-specific coefficients and type shares, Turkish-Math track, female

Placement score: EA	5.46	-4.43	-3.71	-13.38	-5.73	-3.15	0.22	-0.94
Placement score: SOZ	-9.57	-9.00	-13.51	-14.11	-13.28	-10.52	-9.60	-5.91
Major: Arts	5.79	5.20	2.92	-0.01	-0.72	-1.09	-3.80	-5.44
Major: Business and Administration	2.71	-4.73	-3.37	-1.32	1.09	-3.47	-7.96	-7.49
Major: Economics	2.03	-1.57	-6.00	0.14	1.19	-3.39	-7.32	-6.97
Major: Humanities	-3.05	-1.13	-2.23	4.34	-2.41	-7.16	-11.24	-9.95
Major: Journalism and Information	-0.78	-2.12	-2.88	6.23	-4.71	-8.57	-12.20	-10.42
Major: Language and Literature	-0.00	2.92	-4.23	1.04	-1.36	-5.35	1.64	-3.21
Major: Law	-0.99	6.96	-0.41	11.05	1.95	3.72	-7.16	-6.12
Major: Personal services	-1.22	-0.62	-3.65	-0.12	0.36	-6.55	-7.48	-7.68
Major: Public Administration	3.06	-3.84	-0.59	10.74	1.33	-3.27	-9.71	-5.09
Major: Social and behavioural sciences	0.71	-4.44	-0.00	6.59	-1.05	-6.36	-7.54	-8.89
Major: Other	3.91	-7.50	-8.67	-9.24	-10.05	-9.16	-3.79	-8.85
Outside*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

Table A.10: Estimated demand parameters, type-specific coefficients and type shares, Turkish-Math track, male

Placement score: EA	7.20	-6.32	0.70	-7.87	-9.66	-4.23	-2.72	-1.09
Placement score: SOZ	-3.02	-18.11	0.60	-10.98	-9.76	-8.99	-6.01	-12.02
Major: Business and Administration	3.80	2.03	1.64	1.32	-1.70	-3.62	-4.97	-8.89
Major: Economics	2.94	2.15	2.85	1.57	3.84	-6.82	-5.90	-6.84
Major: Journalism and Information	-2.16	-5.18	7.64	-3.31	3.05	-6.07	-7.87	-6.15
Major: Language and Literature	7.97	-0.13	-2.30	-3.93	-1.34	-5.32	-1.83	2.63
Major: Law	5.26	2.20	-0.02	3.95	6.78	3.18	-5.85	-5.05
Major: Personal services	-1.95	0.47	10.22	-2.91	-1.47	-3.97	-7.54	-8.15
Major: Public Administration	4.11	1.47	-1.39	-3.20	7.16	-2.02	-4.95	-5.09
Major: Social and behavioural sciences	-2.52	-0.69	8.52	-4.72	-2.17	-3.36	-8.77	-4.33
Major: Other	4.71	-10.45	-2.12	-8.86	-8.73	-9.82	-9.92	-2.60
Outside*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

Table A.11: Estimated demand parameters, type-specific coefficients and type shares, Social Science track, female

Placement score: EA	-42.01	-7.78	-9.23	-1.65	44.36	17.22	-14.36	-12.86
Placement score: SOZ	-41.67	-13.25	-10.19	-8.98	0.00	14.50	-11.76	-5.11
Major: Arts	3.62	2.07	1.42	0.05	-0.00	-0.55	-2.13	-5.12
Major: Business and Administration	1.77	-3.54	0.66	-7.57	4.18	-3.23	-0.71	-2.41
Major: Humanities	-4.83	-3.24	-6.79	-5.88	-0.01	-2.31	-3.34	-2.98
Major: Journalism and Information	1.39	0.01	1.48	-5.86	-0.00	0.33	-0.04	-7.77
Major: Language and Literature	-3.87	0.67	-5.81	-4.90	-0.00	2.66	2.35	-4.95
Major: Public Administration	4.17	-2.82	4.30	-4.89	4.24	-5.23	0.01	-0.05
Major: Other	-4.11	0.74	-7.27	-9.11	2.64	-2.27	-5.22	-5.47
Outside*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

Table A.12: Estimated demand parameters, type-specific coefficients and type shares, Social Science track, male

Placement score: EA	-6.46	1.47	5.00	-2.10	-9.90	-0.02	-9.19
Placement score: SOZ	-2.27	5.85	0.75	8.93	0.85	-4.62	2.41
Major: Arts	3.38	-0.55	-0.93	-1.85	-2.26	-5.13	-5.61
Major: Business and Administration	-0.90	1.34	3.94	2.62	-0.43	-7.13	-0.36
Major: Humanities	-1.27	-1.17	3.23	-5.88	0.63	-4.43	-6.48
Major: Journalism and Information	3.61	2.78	-1.29	-9.45	-2.43	-1.08	-6.36
Major: Public Administration	6.72	4.34	3.20	3.70	-0.00	-1.69	-0.01
Major: Other	-5.46	2.28	3.23	-5.11	-5.68	-5.08	-2.98
Outside*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

Table A.13: 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 A.10: Minimum Cutoff Scores

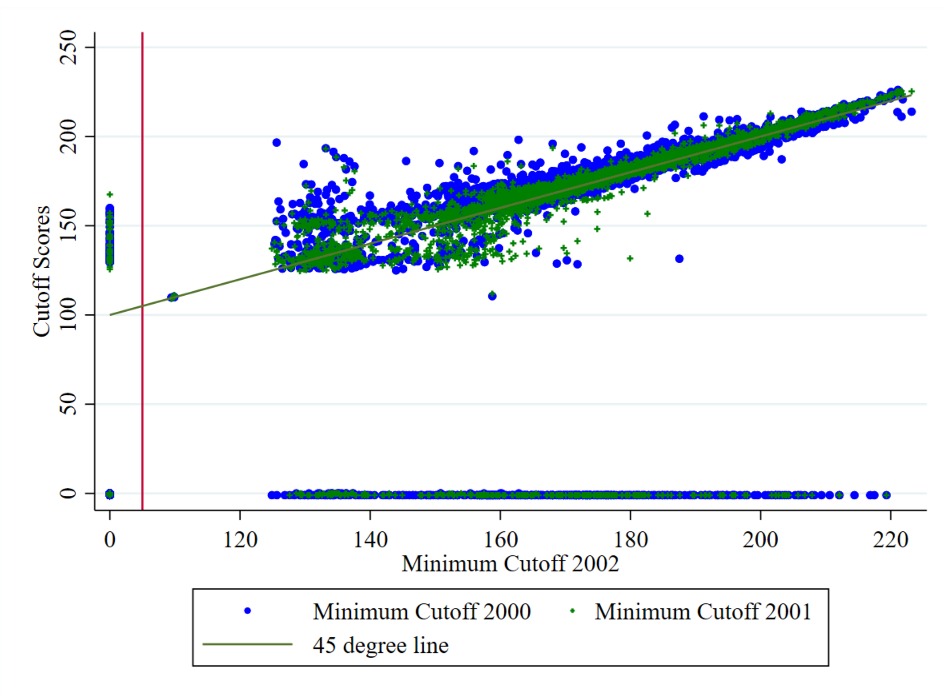


Figure A.11: Gender Differences in Major Choice (Science Track)

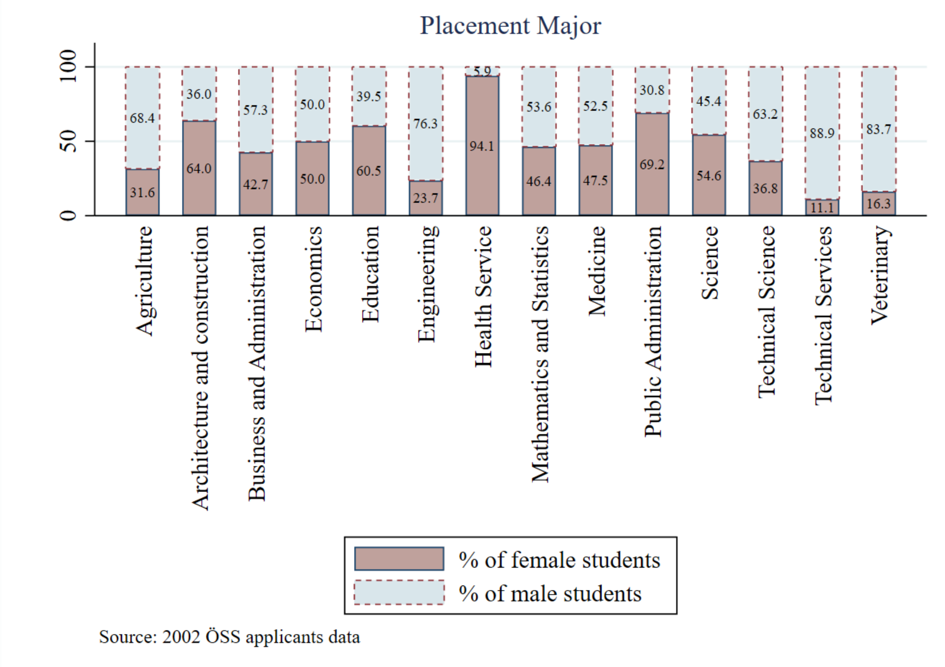


Figure A.12: Gender Differences in Major Choice (Turkish-Math Track)

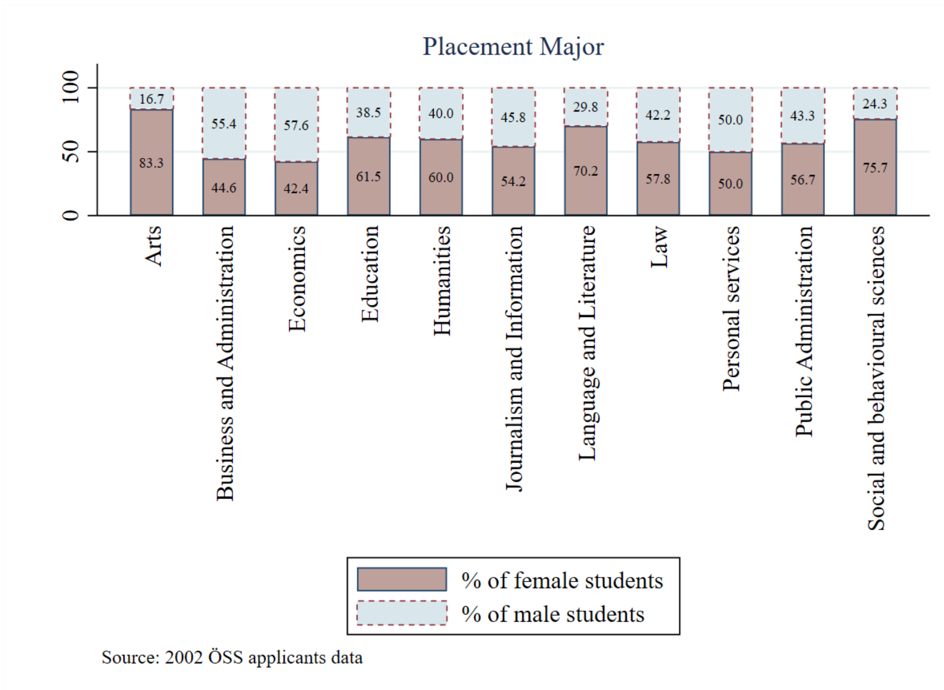


Figure A.13: Gender Differences in Major Choice (Social Science Track)

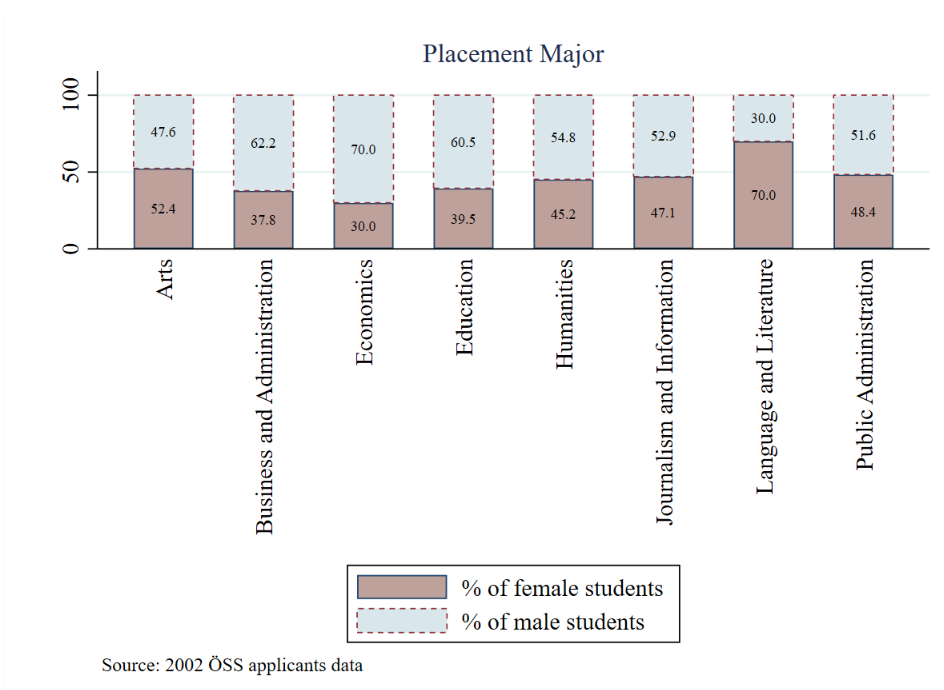


Figure A.14: Placement Majors (Science Track)

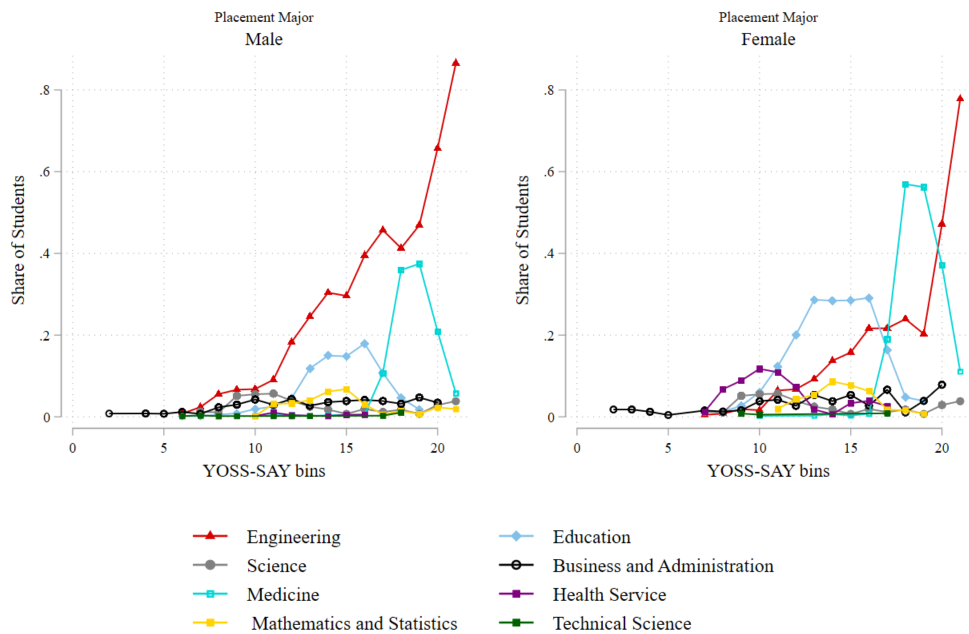


Figure A.15: Placement Majors (Turkish-Math Track)

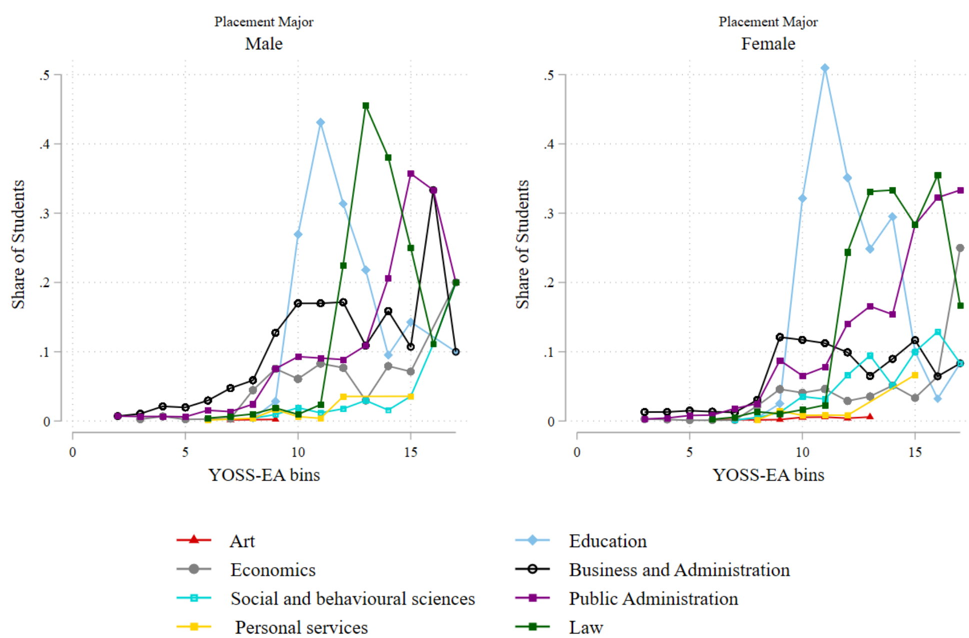


Figure A.16: Placement Majors (Social Science Track)

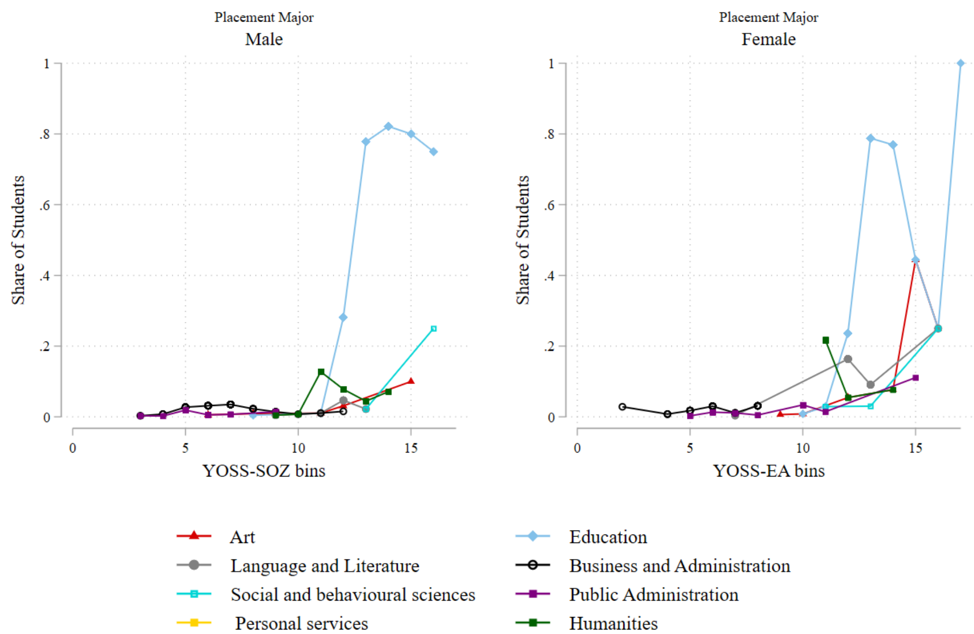


Figure A.17: 1st Preference Major(Turkish-Math Track)

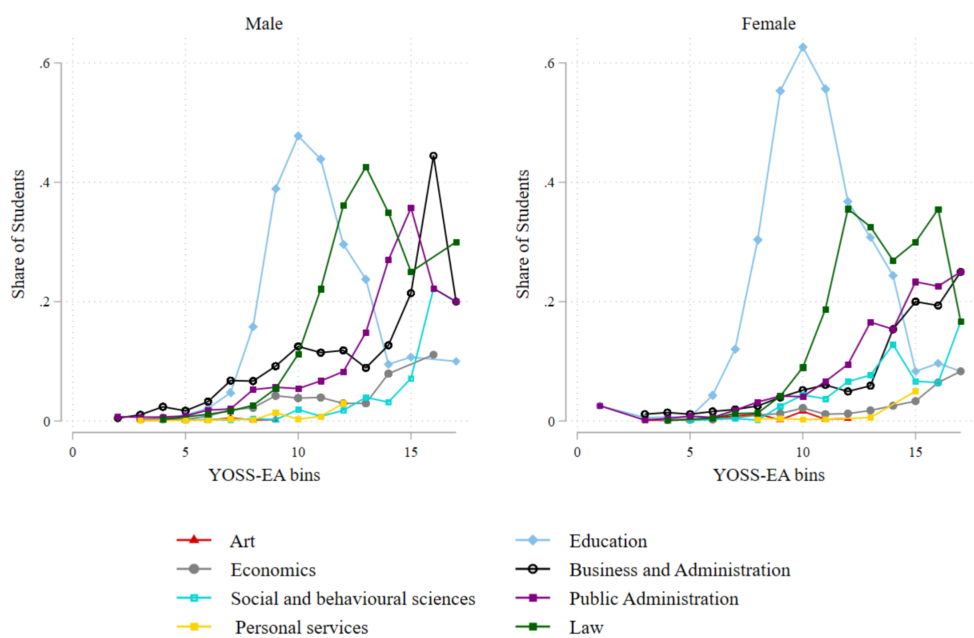


Figure A.18: 1st Preference Major (Social Science Track)

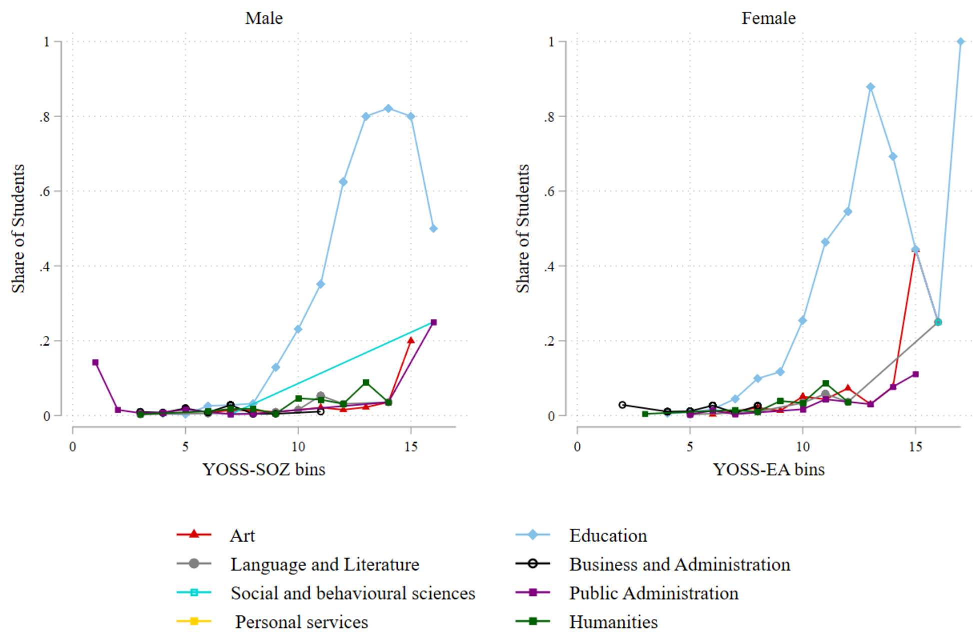


Figure A.19: Distribution of Students according to Dominant Share of Major in Their Preference List

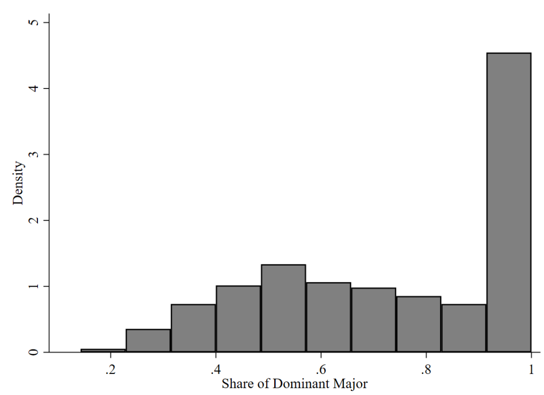
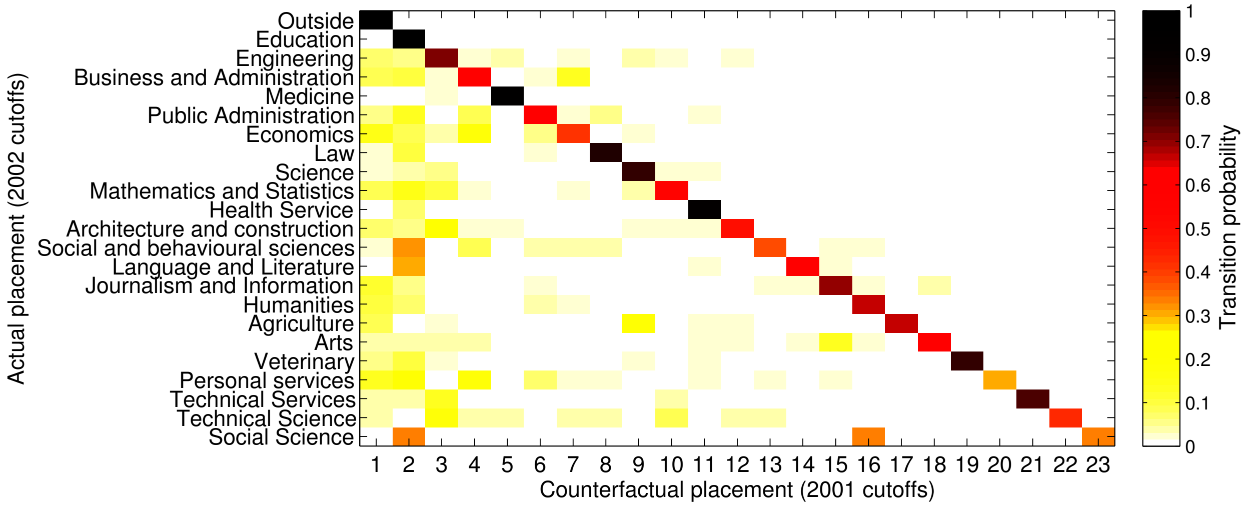
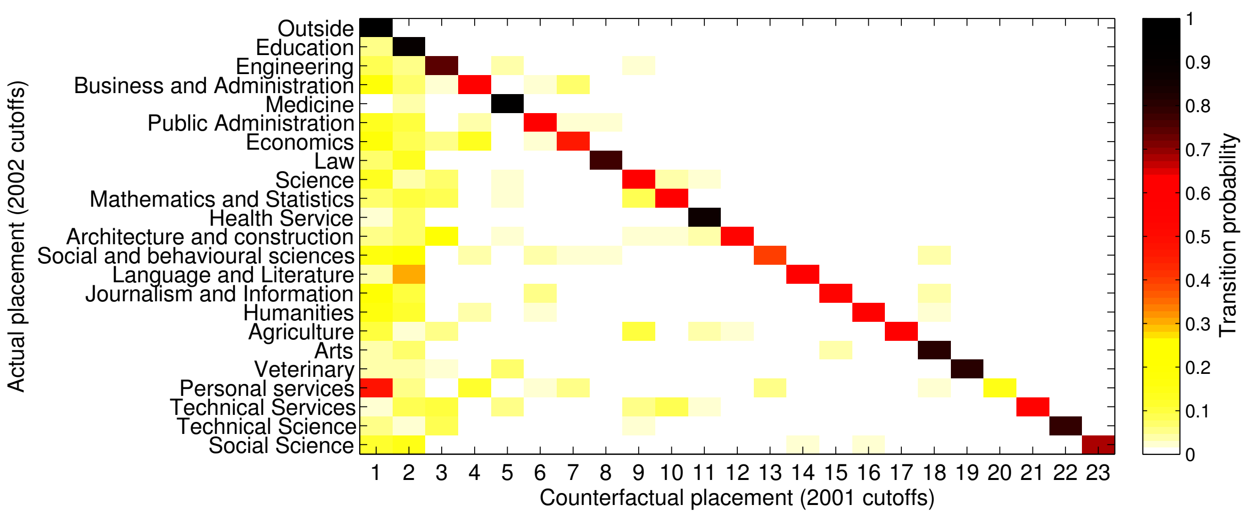


Figure A.20: Transition matrix for majors of placement, predicted using the preference data



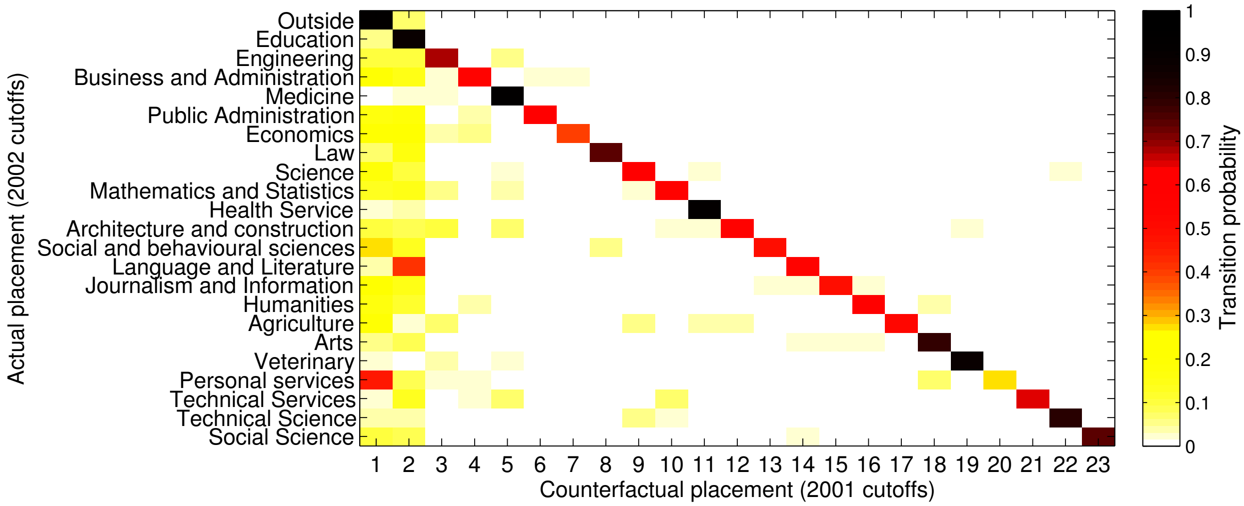
Notes: Actual major — major of placement in 2002. Counterfactual major — major of placement if the admission cutoffs are the same as in 2001. "Outside" corresponds to not being placed. Counterfactual majors follow the same order as the actual ones (e.g., the label 3 corresponds to Engineering). The value 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 A.21: Transition matrix for majors of placement, predicted using the estimated model



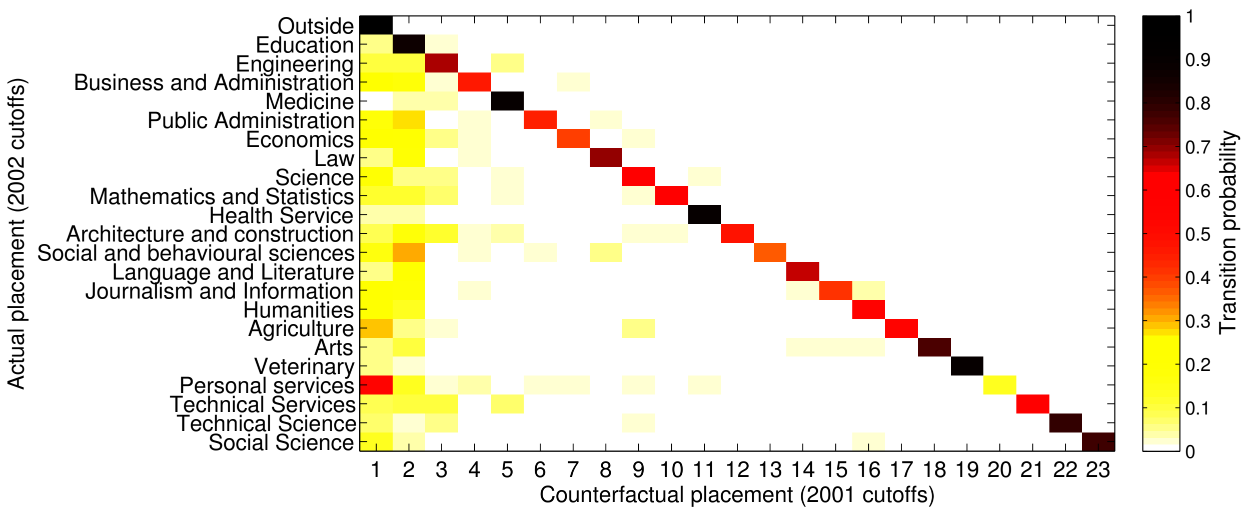
Notes: Actual major — major of placement in 2002. Counterfactual major — major of placement if the admission cutoffs are the same as in 2001. "Outside" corresponds to not being placed. Counterfactual majors follow the same order as the actual ones (e.g., the label 3 corresponds to Engineering). The value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

Figure A.22: Transition matrix for majors of placement, using latent class logit and ex-post stability in 2002 (alternative specification 1)



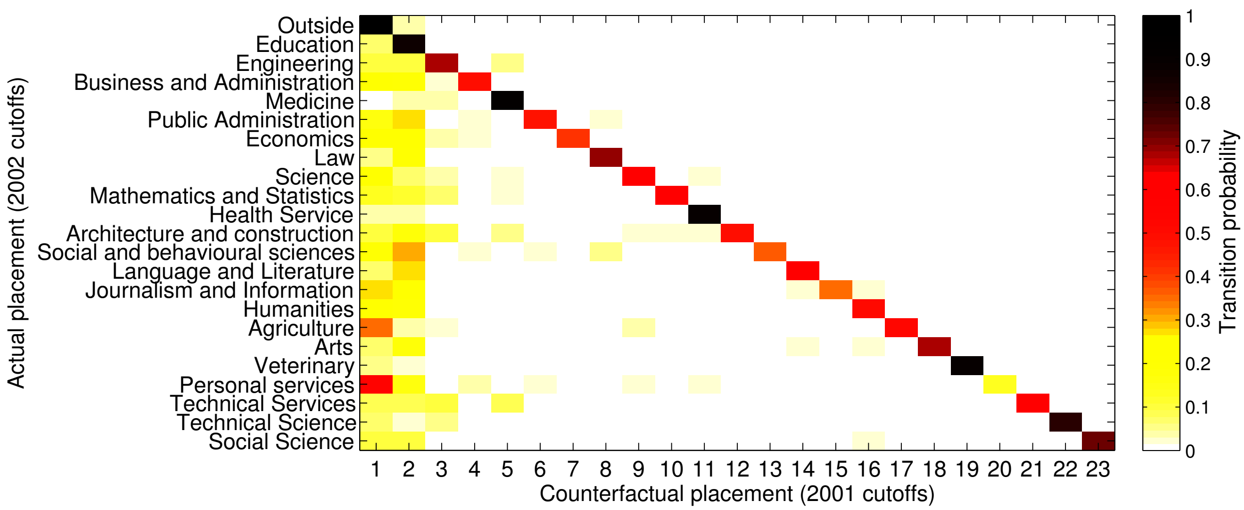
Notes: Actual major — major of placement in 2002. Counterfactual major — major of placement if the admission cutoffs are the same as in 2001. "Outside" corresponds to not being placed. Counterfactual majors follow the same order as the actual ones (e.g., the label 3 corresponds to Engineering). The value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

Figure A.23: Transition matrix for majors of placement, using standard multinomial logit and assumptions 1–3 (alternative specification 2)



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.

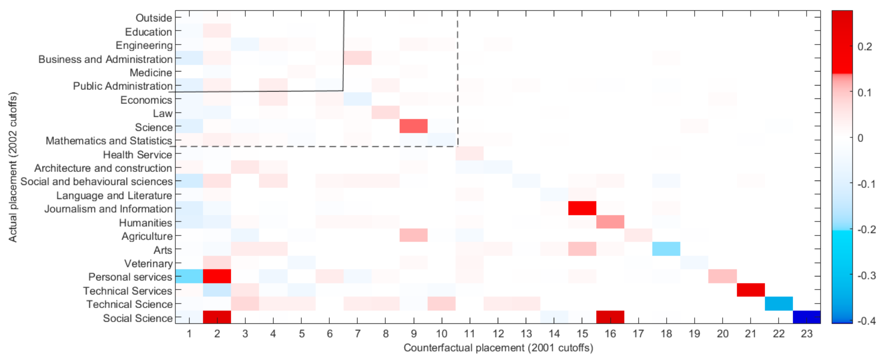
Figure A.24: Transition matrix for majors of placement, using standard multinomial logit and ex-post stability in 2002 (alternative specification 3)



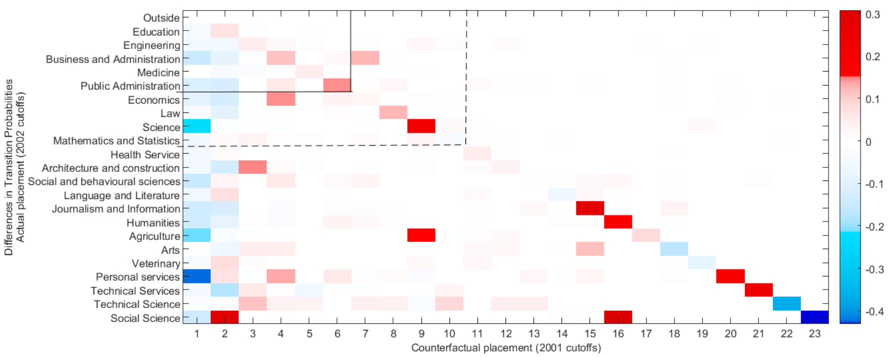
Notes: Actual major — major of placement in 2002. Counterfactual major — major of placement if the admission cutoffs are the same as in 2001. "Outside" corresponds to not being placed. Counterfactual majors follow the same order as the actual ones (e.g., the label 3 corresponds to Engineering). The value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

Figure A.25: Differences in Transition Probabilities between Model and Data for Different Model Specifications

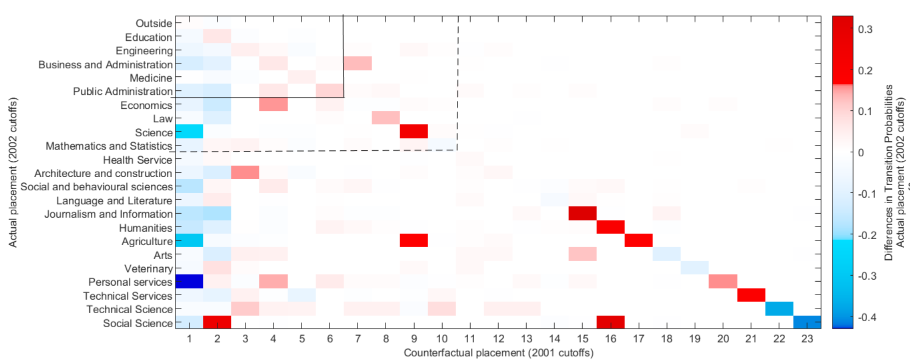
(a) Main Model



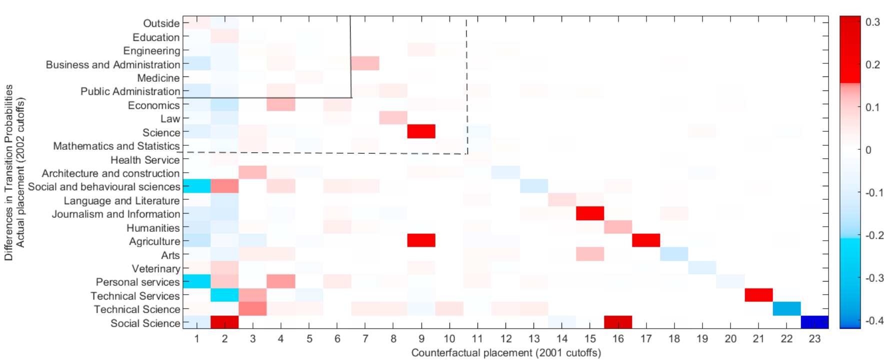
(b) Alternative 1



(c) Alternative 2



(d) Alternative 3



B Predicting Substitution Patterns

We plot a “heat map” representation of where our model (and its competitors) do well and where they do badly 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 over all students to generate the transition matrices. These are to be found in Figures A.20 to A.24.

In Figure A.20 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. Each colored cell depicts conditional probability of switching majors, with darker colors representing higher probabilities.

The substitution patterns predicted by our preferred model are depicted in Figures A.21. The vertical axis depicts major of placement from our preferred model under the 2002 admission cutoffs, while the horizontal axis corresponds to the placements predicted using our preferred model but under the cutoff scores in 2001. Each colored cell depicts conditional probability of switching majors, with darker colors representing higher probabilities. The programs are ordered in terms of their popularity with the most popular ones at the top. The substitution patterns predicted by the models in Columns 2, 3 and 4 of Table 4 are analogously depicted in Figures A.22 to A.24.

Note that our preferred model reproduces the transition matrix for majors quite well. In most cases, students seem to have 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 the major the student was placed in 2002, there is movement to education with 2001 cutoffs.[49] When our preferred model or its alternatives predicts non-negligible switching rates, this usually involves related majors. For instance, economics seems to be a substitute to education, engineering, business, public administration - subjects

[49] This is so no matter which model presented in the columns of Table 4 is used.

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 a 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 A.20 to A.24 differ from one another. To make the difference easier to see we present a heat map of the differences in Figures A.21 to A.24 and Figure A.20. A.25 presents the differences in transition matrices for the data minus those for the model, in question. The vertical axis depicts majors of placement under the 2002 admission cutoffs, while the horizontal axis corresponds to the counter-factual placements predicted under the cutoff scores in 2001. The programs are ordered by popularity with the most popular one being the outside option, followed by education,... The solid lines drawn delineates the programs that account for 90% of the placements. The dotted line drawn does the same but for 95% of the placements. It is easy to see that there are many programs that have a small share of placements.

Each colored cell depicts 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. We present all four comparisons. The difference in transition matrices for the preferred model (column 1) versus the data (column 5) is at the top left. It is very clear that our preferred model does better overall as its colors are lighter everywhere than any of the others. More important, it does particularly well inside the boxed delineated by the solid and dashed lines where most of the action occurs.

C 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:

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

$$\times \exp\left(-e^{-\varepsilon_{j_1}} \sum_{k\in A_1\backslash j_1} e^{\delta_k-\delta_{j_1}}\right) \exp\left(-e^{-\varepsilon_{j_1}}\right) e^{-\varepsilon_{j_1}} d\varepsilon_{j_1}$$

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

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

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

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

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:

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

D 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 student's high school and the program's campus.[50] A dummy for the high school and the campus being in the same province.

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.

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

4. Interactions of net tuition, dummies for evening and distance programs with student income dummies. These controls 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, its interactions with the student's exam scores, the high school GPA and their squares. These terms are meant to serve as a reduced form for the value of retaking the exam in the following year or not attending college 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 population of students may have a latent class that favors medical degrees and never applies for economics and a class that favors economics and never applies for medical degrees. 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 to stop prematurely and produce noisy results.

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 to the respective major; for instance, γ_t for the "economics" latent class t is obtained by running multinomial logit on students who are placed to programs in economics. Third, we create 100 starting values by adding small random shocks to the starting values obtained above. We then maximize the log likelihood function for the fully specified latent class logit model and pick the solution that corresponds to the highest value of log likelihood. Although we did find that the log-likelihood function has multiple solutions, we could not find visible difference between them in terms of the demand substitution patterns they produce.

In order 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 calibrated to be most restrictive for the coefficients on universities and majors, the main culprits behind the preference separation issue. 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. In this sense, the penalties we use roughly correspond to the prior that each university gets at least one applicant from the sample, while each latent class sends at least $\frac{1}{\text{size of the latent class}}$ applicants to each major.

E Department Classification

Language and Literature:

Comparative literature

Eastern Language and Literature

Western Language and Literature

Ancient Language and Literature

Language acquisition

Literature and linguistics

Language and Literature

Interpreting and Translating

Turkish Language and Literature

Sociology and cultural studies

Linguistics

Engineering:

Aircraft Engineering

Biomedical Engineering

Environmental Engineering

Textile Engineering

Electricity Engineering

Electronics Engineering

Industrial Engineering

Physics Engineering

Ships Engineering

Food processing Engineering

Civil Engineering

Chemical Engineering

Mining Engineering

Mechanics Engineering

Material Engineering

Mathematics Engineering

Metallurgical Engineering

Nuclear Energy Engineering

Forestry Engineering

Motor vehicles Engineering

Petrol Engineering

Textile Engineering

Natural Science Engineering

Education:

Vocational Education

Language Education

Pre-School Education

Technical Education

STEM Education

Education science

Social Science Education

Turkish Language Education

Business and Administration:

Finance, banking and insurance

Business, administration and law

Accounting

Marketing and advertising

Management and administration

Wholesale and retail sales

Economics

Mathematics and Statistics:

Mathematics

Statistics

Health Service:

Health Service

Nursing and midwifery

Humanities:

History and archeology

History

Philosophy and ethics

Religion and theology

Sociology and cultural studies

Medicine:

Dental studies

Medicine

Pharmacy

Science:

Earth sciences

Physics

Biochemistry

Biology

Chemistry

Science

Technical Science:

Software and applications development and analysis

Database and network design and administration

Technical Service:

Environmental protection technology

Mining and extraction

Earth sciences

Food processing

Motor vehicles, ships and aircraft

Social and behavioral sciences:

Psychology

Political sciences and civics

Sociology and cultural studies

Public Administration:

International Relations

Political Science

Public Administration

Political sciences and civics

Personal services:

Hotel, restaurants and catering

Transport services

Travel, tourism and leisure

Journalism and Information:

Audio-visual techniques and media production

Journalism and reporting

Library, information and archival studies

Agriculture:

Agriculture

Fisheries

Crop and livestock production

Architecture and construction:

Architecture

Fashion, interior and industrial design

Architecture and town planning

Arts:

Fashion, interior and industrial design
Music and performing arts

Audio-visual techniques and media production

F Allocation Score (Y-ÖSS)

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

$$\text{Y-OSS-X}_i\text{=OSS-X}_i\text{+}\alpha\text{AOBP-X}_i$$

where $X, \in \{SAY, SOZ, EA, DIL\}$, and α is a pre-determined constant which changes according to the students' track, preferred department and whether student placed in a regular program in the previous year or not. OSYM 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 to 0.2. If student graduated from a vocational high school and prefers a department that is compatible to his high school field, α equals to 0.65. If student placed in a regular university program in the previous year, the student is punished and α equals to 0.25, 0.1, and 0.375, respectively, that is, for such students, the α coefficient is equal to half of the regular α.

In turn, the AOBP score (of student i from a given track in school j in programs that require OSS-SAY, OSS-SOZ or OSS-EA) is a function of normalized high school GPA (OBP_j), minimum and maximum normalized high school GPA in the high school the student graduated from ($\min_{i\in j}(OBP_i), \max_{i\in j}(OBP_{ij})$), and the mean OSS score in $k = SAY, SOZ, EA$ (OSS_{jk}) among graduating seniors in that school as in equation (4). Students keep their AOBP over attempts made.

$$\begin{aligned}
&AOBP_{ijk}\\
&= F[OBP_{ij}, \min_{i\in j}(OBP_i), \max_{i\in j}(OBP_i), OSS_{jk}]\\
&= [(\frac{\ddot{O}SS_{jk}}{80} \times \min_{i\in j}(OBP_i)) - (\frac{\ddot{O}SS_{jk} - 80}{10})]
\end{aligned}$$

$$+[(OBP_{ij} \times \frac{\ddot{O}SS_{jk}}{80}) - (\frac{\ddot{O}SS_{jk}}{80} \times \min_{i \in j}(OBP_i))]$$
$$\times \left[\frac{80 - [(\frac{\ddot{O}SS}{80} \times \min_{i \in j}(OBP_{ij})) - (\frac{\ddot{O}SS-80}{10})]}{(\frac{\ddot{O}SS}{80} \times \max_{i \in j}(OBP_{ij})) - (\frac{\ddot{O}SS}{80} \times \min_{i \in j}(OBP_{ij}))} \right] \tag{4}$$

We don't observe students' AOBP score in our data set. However, we know the rule as above, as well as the inputs into the rule[51] other than the minimum and maximum OBP scores in the school. In our sample, we observe the normalized high school GPA (OBP_{ij}) and GPA for all students. OSYM calculates OBP as follows:

$$OBP_{ij} = 10\frac{gpa_i - \mu_{gpa,j}}{\sigma_{gpa,j}} + 50 \tag{5}$$

where gpa_i is the students' own GPA, $\mu_{gpa,j}$ Average GPA in school j and $\sigma_{gpa,j}$ = Standard deviation of GPA within School j. The student's own GPA and OBP are observed in the data. Thus, 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 $\sigma_{gpa,j}$.Thus, for almost all schools, we can obtain $\mu_{gpa,j}$ and $\sigma_{gpa,j}$. The OBP is forced to be between 30 and 80 (This is a rule of OSYM, if the OBP formula gives a number less than 30, it is set to 30 and if it is more than 80, it is set to 80). The OBP formula suggests that average student in each school gets 50 as OBP. Therefore, the maximum OBP cannot be less than 50 and the minimum OBP cannot be more than 50.

The other missing component is the minimum and maximum OBP in each school. To pin down the maximum OBP in a school, 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 GPA for 445 schools. The summary statistics of OBP of these students are as follows:

	# of Obs	Mean	Std	Min	Max
OBP	445	71.03213	5.557276	55.538	80

[51]We obtained each schools' mean OSS scores in each field for the 2002 high school graduates from OSYM website.

This means that on average, the first ranked students are two standard deviations away from mean GPA. Note that GPA is bounded from above by 5. From equation (5) we see that, depending on mean GPA, maximum OBP also bounded from above. Also, if mean GPA in a school is very high, the maximum OBP is smaller. To find the highest possible OBP in a school (where we don't observe first ranking student) we calculate the OBP of the student with GPA 5. Notice that we don't know whether there exists a student with GPA 5, but we know that max OBP cannot be higher than calculated OBP for this hypothetical student. This calculated maximum is denoted by $max_{obp,j}$.

In the next step, we assume that OBP scores in each school has a beta distribution with mean 50, standard deviation 10, and support $[30, max_{obp,j}]$

In the first step, for each school we find the parameters of the distribution for each school, given the mean and standard deviation across schools to be 50 and 10 as forced by OSYM. Since mean and standard deviation across schools are same in all schools, parameters differ in each school only because of the different support of the distribution.

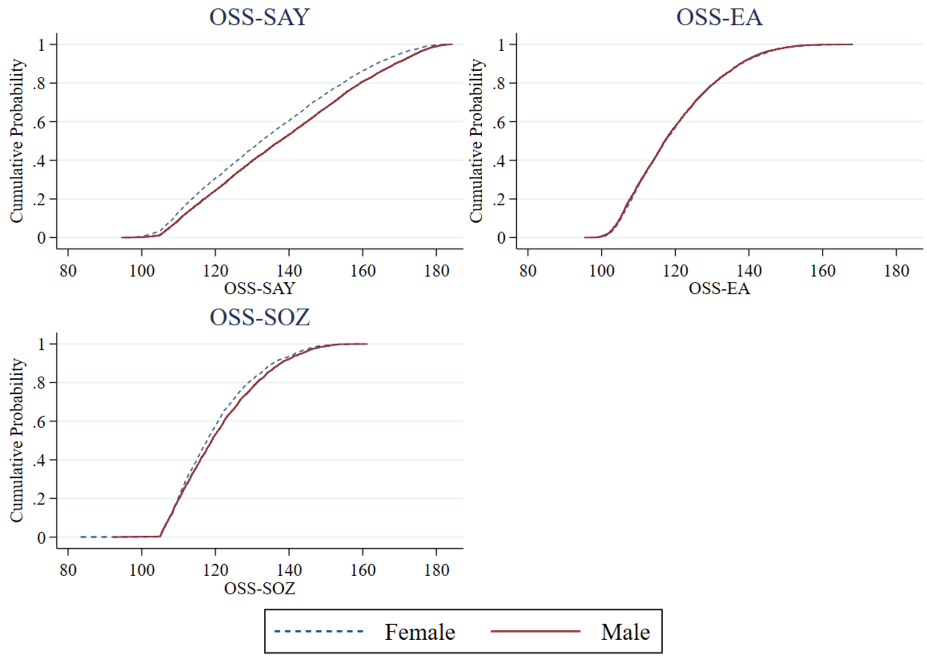
In the second step we draw from the beta distribution the number of draws that correspond to the class size in the school using the parameters estimated in the first step. We do this S times for each school. We then find the average minimum and average maximum OBP over the S draws which we use as our estimate 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 min OBP for this school.If we observe higher maximum OBP in the data, we use it as the max OBP for this school. Otherwise, we use the simulated minimum and maximum OBP scores.

Figure G.1: OSS Score Distributions by Gender



G Score and GPA Distributions

Figure G.1 presents the cumulative distribution of exam scores (OSS) by gender for first time takers and do so separately for each track. For students in each track, the weights used to calculate the placement score are those corresponding to their track.[52] For the score used in the Science track programs (OSS-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 OSS-SOZ. On the other hand, for OSS-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).

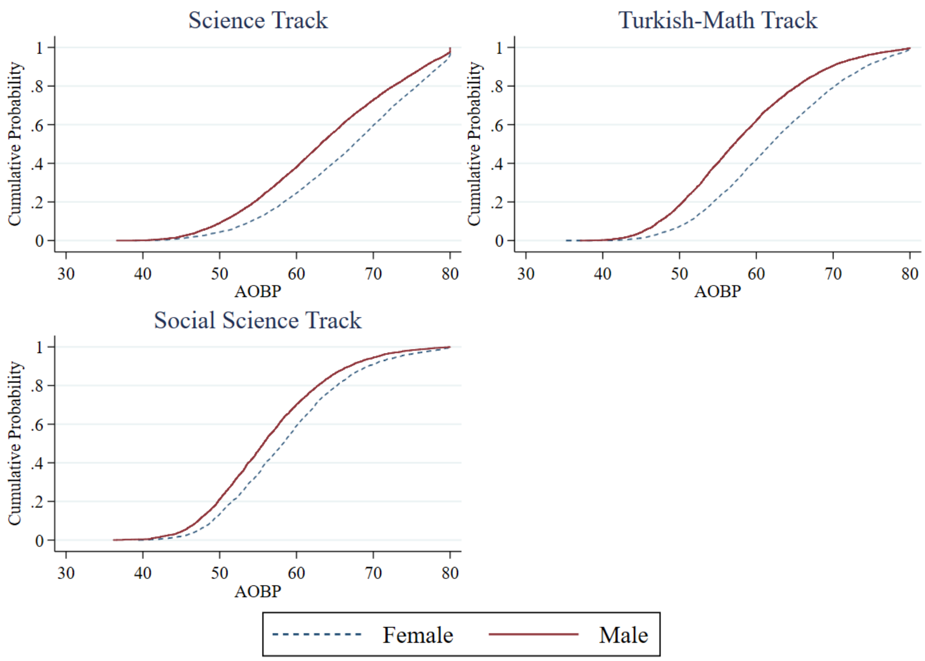
The distributions of high school GPA (AOBP)[53] follow an opposite pattern: women tend to perform better in school than men do[54] (see Figure G.2). Since the placement score (Y-

[52]Recall, each student has three placement scores as different programs have different weights in calculating their placement score.

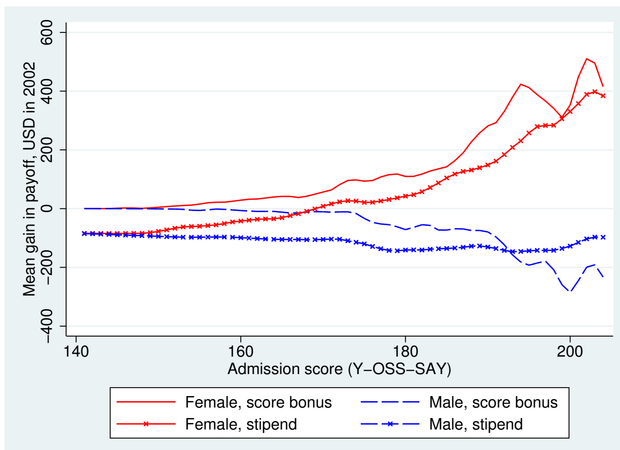
[53]Since different schools could differ in their grading standards, the HSGPAs are normalized by school performance. This is called the AOBP score.

[54]This pattern, where males do better in high stakes exams has also been observed in other settings. Voyer and Voyer [2014], in a meta analysis 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 then men.

Figure G.2: AOBP Distributions by Gender



OSS) is a mix of the exam score (OSS) and the GPA (AOBP), the gap in placement scores is less than that in exam scores.



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 science track of high school. The payoffs are expressed as annual stipend equivalents in the 2002 US dollars. The stipend policy is financed by a lump-sum tax; the expected payoffs are net of the tax.

Figure G.3: Expected changes in payoffs by score, net of taxes used to finance the stipend