Deconstructing the Placement Gender Gap: Performance versus Preferences

Pelin Akyol *   Kala Krishna†   Sergey Lychagin ‡

This is a very preliminary draft. Please do not cite or circulate without authors' permission. Comments are very welcome.

Abstract

When good primary and secondary education is free, and at the university level, is heavily subsidized, 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 by gender and differences in preferences across majors. We then 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
‡Central European University, , e-mail:lychagin@gmail.com

1 Introduction

In the United States today, over 50% of entering law school students are female. In 1958-59 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 by Supreme Court Justice Felix Frankfurter for a clerkship on the grounds that she was female. In Economics among the top 100 US universities, at the undergraduate level[2], there are more than two men majoring in economics for every woman. This fraction is roughly the same at the Ph.D level.[3]

Reducing the gender gap is important, not just for equity reasons, but for efficiency. If, as one might believe, intrinsic comparative advantage exists, there may be large losses from barriers to entry for women in some fields. Hsieh et al. [2019] take a structural approach and use a Roy model together with data on the US to show 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. If gender neutrality/equality is the ultimate goal, we need to understand what drives the difference in the choice of college fields, which is where these gender gaps start, to better understand the appropriate policies to reduce these gaps.

This is the focus of this paper. Our contribution to the literature is two fold. First, on the methodological side, we propose an innovation, similar in spirit to that of Berry et al. [2004] who use not just the product chosen, but the alternative choice had this product not available. We show that doing so allows us to fit substitution patterns in the data very well. This is important because this patterns are notoriously difficulty to match. Second, on the policy side, we show that because of differences in preferences by gender in general equilibrium settings, the obvious policy measures (like affirmative action) are relatively ineffective.

In Turkey, as in most countries, men are over represented in STEM programs. These correspond to engineering and technical science programs. We ask where might this be

[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] Only 25% of assistant professors and only 13% of full professors are women (Lundberg and Stearns [2019]).

[3] 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]).

coming from? The gender gap in placement could come from four sources: exam scores, less aggressive application and retaking behavior, and preferences.

First, it could be that women have lower exam scores. This could be because women have less to gain from a higher score, given their attachment to the labor force tends to be lower, and the value of their outside option, i.e., marriage, higher.[4]. Especially if more competitive fields with higher placement cutoffs tend also to be seen as family unfriendly, women could rationally decide to study less for the university entrance exam and hence do worse there. We focus on the differences in exam score by gender. We focus not just on the *average gender gap* in score, as is usually done, but on the gender gap function: how the gender gap in score varies by AOBP, the normalized high school GPA used in the placement score. We find a hump shaped gender gap as a function of AOBP, with an average gap of about 9 points (the maximum possible score is 220 points) for students in the Science track.

Second, gender gaps in placement could arise from women applying to less competitive programs than men, given their score. There is a large literature in Economics that argues that women are more risk averse and less competitive, see for example (Niederle and Vesterlund [2011], Niederle and Vesterlund [2007a], Eckel and Grossman [2008]). We show in Section 3.2 that this does not seem to be the case in our data: in fact, such an interpretation looks like it is coming from a composition effect where women apply more to majors where there is a large average gap in the placement score and the admission cutoff score. Hence, we do not focus on this aspect.

Third, lower placement scores could arise because women retake less often. Families could be less supportive of women retaking. It could also be that women wish to retake less as they are more risk averse. Retaking is a risky business, since the last score obtained is used. Saygin [2016] suggests that this channel is important. Her work exploits the fact that to retake without penalty one must be unassigned, i.e., not put down any school that is feasible. She shows a marked average preference on the part of women to not be unassigned, which she interprets as a signal they will not retake. Retaking is the main focus of Krishna et al. [2018], who build and estimate a structural dynamic model of retaking using the same

[4]In Turkey, the female labor force participation rate was 30 percent while the male labor force participation was 75 percent in 2002.

data as used in our paper. The retaking channel is not the primary focus of this paper. However, we do incorporate differences in retaking by gender into our preference estimation as explained in Section 5.

Fourth, the gender gap in placement could arise because women seem to prefer very different fields, with the observed gender gaps just a reflection of this. One reason for different preferences, it has been argued, is that STEM fields like Engineering are more competitive and women prefer less competition. Women could also have very different preferences for social/cultural reasons. Women may attend university with the marriage market in mind (see Kirkebøen et al. [2021]). Certain fields may be seen as inappropriate/appropriate for women culturally (Veterinary Science) or low pressure and family friendly (Education). This is a core part of our analysis and we find very large effects coming from this channel[5]. We estimate preferences by building on state of the art methods to extend this frontier. We use our estimates to do counterfactual simulations for policy purposes. We find, contrary to what we expected a priori, that giving additional points to women, does little to move them into male dominated fields or stem subjects. The intuition is simple. Women's preferences differ substantially from that of men. As a result, giving them more points tends to raise the cutoffs in subjects favored by women, without reallocating women to more competitive STEM fields. Giving women points only for STEM subjects does a bit better, but its efficacy is also limited. It is worth pointing out that we do not incorporate the choice of track in High School in our analysis. Policies, especially those that give points to women in STEM subjects only, could also affect this margin. As a result, our counter-factual estimates are likely to be on the conservative side.

We focus on the two channels that seem to be the crucial ones in our setting: performance and preferences. What does the existing literature say in this regard? The literature surveyed in Kahn and Ginther [2018] suggests that there is a gender gap in performance in Math in the US by the end of High School with males being more likely to take and do well in more math intensive subjects like Physics while females are more likely to take Biology. Boys also

[5]We do not focus on less aggressive application and retaking behavior as the former does not seem to exit, while we incorporate the latter, at least partially, into our preference estimation. We focus on the extent to which the gender gap in placement is driven by differences in performance in the university entrance exam which we call the score gender gap, and/or preferences.

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. Turner and Bowen [1999] find that SAT score differences account for 45% of the gender gap in terms of enrollment in math-physical science fields and 32% of the gender gap in engineering under the assumption that the college and major are chosen together. This may be different in Turkey, as the educational institutions as well as the role of women in society (which may impact their preferences) may differ in Turkey and the US. The US system is much more flexible than the Turkish one: students choose their majors after they start college and can switch until quite late in their studies. The time of choosing a major, the number of required and elective courses taken before the major choice decision, vary across settings and may affect program choice decisions, see (Stinebrickner and Stinebrickner [2014]). In addition, norms regarding work and family may differ in the two countries which may well impact behavior. Performance gender gaps in Math and the Physical Sciences are very important in Turkey as scores in these subjects are heavily weighed in determining placement scores in STEm subjects.

Why might a gender gap in preferences across fields and performance exist and why might it matter? The reasons offered for a gender gap in preferences are many. Carrell et al. [2010] 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. In some countries safety might be an issue as women as often subjected to harassment, especially while traveling. A hostile environment for females in the field could be another reason. In Economics for example, Wu [2018] and Wu [2020] showed, using textual analysis from the Economics Job Market Rumors website, how women are portrayed negatively by men. The absence of affordable child care may also drive women into areas where such flexibility is easier. We do not focus on why preferences might differ by gender, and the consequences of such difference. Kahn and Ginther [2018] provide an excellent overview of what seems to drive differences in female representation in STEM subjects focusing on the US. This gender gap in preferences exists: Zafar [2013] uses information from a survey of 161 Northwestern University students in Arts and Sciences to separate the role of preferences, enjoyment of the course work, and the ability to balance work and life. Estimating a structural model he finds that these three account for 60%, 27% and less than 5% of the gender gap in Engineering respectively. It matters at the very least because of the large earning differences in fields chosen by men and women. Arcidiacono [2004] find large earning differences across majors in chosen by males and females in the US even after controlling for selection consistent with differences in preferences driving such choices. Loury [1997]. Loury and Garman [1995] also document large earning differences across majors.[6]

We study this question using data on Turkey. Admission to university in the Turkish system is based only on performance (high school GPA and university entrance exam scores) and stated preferences, it provides a simpler setting in which to study the role of these two factors in driving gender differences across majors.[7] In contrast, admission decisions in the US are more opaque[8] and major choices are made after admission, not before making such an analysis more difficult. We use a representative sample of 2002 University entrance exam takers in Turkey. We only use the sample of first time takers in our data and estimation as retakers select into doing so and we do not want to focus on retaking. We ask, how much of a role do these two factors, performance in the entrance exam and preferences, play in explaining the observed gender gaps? What policies might be effective in closing them? These two questions are at the heart of this paper. In order to answer these questions, we need to have a way to represent preferences that can be taken to the data. It is well understood, see for example Berry et al. [2004], that using information on observed choices alone (i.e. placement in our setting) and a Logit structure gives substitution patterns that are not close to those in the data. We do have the entire preference list of students. This might push one to use exploded Logit. However, this would be the wrong thing to do. From

[6]In Turkey, women tend to major in subjects with lower returns. While males are more likely to major in engineering, females are more likely to major in education and the humanities.Table A.15 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.

[7]After students learn their scores they submit a preference list and are allocated to college programs with higher scores given priority.

[8]Recent events made Harvard's closely guarded admissions procedure public revealing its innate biases. Arcidiacono et al. [forthcoming] looks at how Harvard University makes its admissions choices regarding Legacy and Athletic admissions. Arcidiacono et al. [2020b] look at its recruitment of African American students, while Arcidiacono et al. [2020c] look at whether they discriminate against Asian Americans in admissions.

Chade and Smith [2006], we understand that in the presence of costs of adding to the list and/or limits to the length of the list, while choices on the list will be correctly ranked, there are likely to be "holes" in the list so that exploded logit would be inaccurate in its assumptions.

Our contribution on the methodological front is a novel strategy to identify demand for college seats using information on substitution patterns from student preference lists. 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.

Berry et al. [2004] argue that "models without unobserved heterogeneity (but with observed consumer attributes) do a bad job of reproducing observed substitution patterns." For this reason we incorporate unobserved heterogeneity in our estimation.We allow for retaking possibilities in our estimates of unobserved heterogeneity. If women were less likely to retake, this would show up as their being more likely to be assigned to unobserved types with a low outside option associated with retaking. In this way, we account for such difference in retaking preferences in our preference examination.

Our method is based on three identifying assumptions. First, we follow Fack et al. [2019] and assume that observed placements are asymptotically ex-post stable. This means that the placements observed in 2002 are optimal under the realized admission cutoffs for all students except for a vanishingly small share. Second, we assume that the placement generated by 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; i.e. $j_1 \succ j_2$ if j_1 is listed before j_2 and vice versa. 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 logit estimates do not capture the substitution patterns well, the random shocks are blown up in an attempt to fit the data. For example, Houde [2012] shows that without accounting for commuting patterns, the logit estimates of demand for gasoline does not capture substitution patterns well. In order to account for high sales of gas stations on highways, which are far from where consumers live, the cost of distance is biased downwards, the extent of substitution biased downwards (the variance of shocks is biased upwards).

Agarwal and Somaini [2018] in their insightful paper showed how to use revealed preference insights along with parametric assumptions on the error term to estimate preferences. Their approach is unfortunately not suited to our setting because the number of choices available in our data are orders of magnitude larger than those available in their data. Larroucau and Rios [2018] show that one can estimate preferences restricting attention to single level replacements, i.e., a first order approach. We choose not to do so as there is a small literature that shows that in the school choice area, students put dominated choices above undominated ones, suggesting that using the entire list might be problematic. For example, students often make obvious mistakes like ranking programs with scholarships below those without scholarships, even though the two variants have the same cutoffs. See Hassidim et al. [2017] for a short survey. Laboratory experiments show that a significant fraction of subjects do not report their preferences, see for example, Chen and Sönmez [2006]. Mistakes also occur in real world settings, see Hassidim et al. [2021] for an illustration using graduate programs in psychology in Israel, Rees-Jones [2018] using the medical resident match in the United States, and Artemov et al. [2017] using the Australian university admissions.

Using stability, i.e., the assumption that each student is placed in their most favored program in their feasible set, is a more conservative approach. Fack et al. [2019] argue that stability is a good assumption since when school number and capacities are fixed and the number of students increases proportionally, the fraction of students not matched with their favorite feasible school tends to zero. They show that stability based estimates perform well in fitting the data and in generating cutoffs close to the actual ones using data on schools in Paris. In this paper we assume that each student's placement is stable under the current and previous years cutoffs. This makes sense as students are given the previous years cutoffs as part of the application process.

The rest if the paper is organized as follows. The next section describes the data, and gives the necessary background information regarding the college entrance system in Turkey. We present patterns found in the data in section 3. We do so graphically and by using some simple regressions. Section 4 looks at the factors that lie behind the differences placement score in exam scores. We show that women do not seem to be less aggressive than men in applying to programs, an argument often made to explain the gender gap in placement. We argue that a composition effect is what lies behind this average gap. Once we control for major of placement, this gap is no longer significant.

We then look at the gender gap in placement. To do so we regress the exam score on students characteristics such as background, expenditure on prep school, AOBP (the normalized High School GPA used in the placement score), the school attended, and a dummy for being male. The coefficient on this dummy gives the average gender gap once we have controlled for all observables. The average gender gap is positive in line with the work of Saygin [2016]. We argue that we need to look at the gender gap as a function of the placement score. When we do so, we find that the average gender gap masks masks large differences across students of different abilities as measured by AOBP. The gender gap is significantly negative at low abilities and significantly positive at high abilities. Using the estimates obtained, we can evaluate the importance of the variables included in the regression on the performance gap for the placement gender gap function we depict. We find that in the science track, these variables do not make much of a difference in the gender gap function, while they do in the Turkish math and Social Studies tracks suggesting a difference in students in the three tracks.

Finally, we turn to preferences. In Section 5, we lay out our model for estimating preferences and explain the nature of, and how we deal with, several issues that arise. Following that we present our estimates of preferences and run a horse race between are approach and several alternatives to show that our approach does best in terms of matching the substitution patterns in the data. We than do counter-factuals using our estimated preferences in combination with a general equilibrium model that gives cutoff scores. We show how much the different channels present in the gender score gap would account for in terms of the gender placement gap. We find that eliminating the preference channel alone would do a lot more: it would remove roughly 60%. of the gender gap in placement while removing the gender gap in score would only remove 30% of the gender gap in placement. Section 6 then looks at the policy implications. We consider various policies including affirmative action that is undirected (giving bonus points to women) as well as directed (bonus points for only certain subjects or to students coming from certain regions). Section 7 concludes.

2 Data and Background

In Turkey, a year after students start high school, they choose one of the four - Science, Turkish-Math, Social Studies, or Language - tracks. 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 constructed where each score puts different weights on each of the four tests. 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.[9]

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.[10] The system is relatively stable in this period so that it is not unreasonable to think of students having a fairly good idea what their feasible set is.

Students face fierce competition, especially at the top. A placement score of 211.20 (out of a maximum of 224) would allow an applicant in Engineering to enroll in the 6th ranked school and those with a worse ranking. A score of 200.70 would allow only schools ranked 17th and worse to be feasible. In medicine, placement score of 207.9 would allow an applicant to enroll in the 3rd ranked school and worse. A score of 201.99 would allow only

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

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

schools ranked 8th and worse to be feasible. 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. Student are placed in the highest ranked program available to them once higher ranked students have been placed. This is equivalent to the deferred acceptance algorithm being used. 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,[11] and are merit, not need, based in contrast to the norm in the US. The placement score is again used to rank students.

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 test scores (ÖSS_X), high school ID, track, high school GPA, gender, if they submit a preference list, their ranked preference list, and the college they are assigned, if any. We construct the normalized high school GPA (AOBP) by reverse engineering as explained in the Appendix E. The survey data includes information on students' family background. 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 minimum cutoff scores, maximum admission scores and available number of seats in all college programs for the years 2000, 2001, and 2002. This data also includes tuition cost of each department, amount of the scholarship, if provided. [12] In addition, we collected the distance between all cities from the General Directorate of Highways.

Summary statistics are presented in Table 1 for each track by gender. Columns 1 and 2 present the means and standard deviations of each variable for females and males, respectively. Column 3 presents the difference between females and males. Note that ***, **, and

[11]Admission is to a program in a university, not to the university more broadly.

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

* denote that a significant difference at the 1%, 5%, and 10% levels respectively. The same statistics are presented in columns 4 to 6 for Turkish-Math track students and in columns 7 to 9 for Social Science track students. Note that the gender gap in the 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 relative to 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).[13] 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.[14] 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.[15]

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[16], conditional on finishing high school, but are

[13]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.

[14]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.

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

[16]All students in science high schools are from the Science track, which is why the entries are blank in

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 form 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)	

(continued on next page)

other tracks.

VARIABLES	Science Track (1) Female Mean (SD)	(2) Male Mean (SD)	(3) (1)-(2) Diff.	Turkish-Math Track (4) Female Mean (SD)	(5) Male Mean (SD)	(6) (1)-(2) Diff.	Social Science Track (7) Female Mean (SD)	(8) Male Mean (SD)	(9) (1)-(2) Diff.
Scholarship	0.044	0.057	-0.013***	0.023	0.019	0.005	0.014	0.015	-0.001
	(0.205)	(0.232)		(0.151)	(0.135)		(0.118)	(0.122)	
Highest Parental Education:									
Missing	0.072	0.052	0.020***	0.069	0.050	0.019***	0.065	0.040	0.025***
	(0.259)	(0.222)		(0.254)	(0.218)		(0.246)	(0.195)	
Literate	0.034	0.062	-0.027***	0.042	0.090	-0.048***	0.058	0.112	-0.054***
	(0.182)	(0.241)		(0.200)	(0.286)		(0.234)	(0.316)	
Primary School	0.237	0.256	-0.019*	0.317	0.330	-0.013	0.438	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 Data Patterns

In Turkey, around 1.5 million students take the University Entrance Exam every year, and only one third of these are offered a place a university program. Students face fierce competition. On the other hand, there is no limit on retaking the exam. Students, who are not placed in a college program, can take the exam next year without penalty.[17] After

[17]Students already in a program face a penalty.

students learn their scores, they make a decision about which programs/options to choose. Students are allocated to their preferences centrally by using the multi-category serial dictatorship allocation algorithm. Basically, starting from those with the highest allocation score, students are assigned to their highest ranked school that still has open seats. In this allocation process, the allocation scores of the students and their preferences are the only criteria. Students know their allocation scores and the previous years' cutoff scores while submitting their preference lists. The system has been stable since 1999.[18] Moreover, the number of seats in each major as well as the number of applicants was roughly stable from 2000-2002. As a result, program cutoffs were also relatively stable. These cutoffs are depicted in Figure 1. 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.[19]

3.1 Are Women's Preferences Different?

As in many other countries, there are significant gender gaps across college majors. Figure 2 presents the percentage of female and male students in each major[20] according to placement.[21]. 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. Part of the reason for this difference in

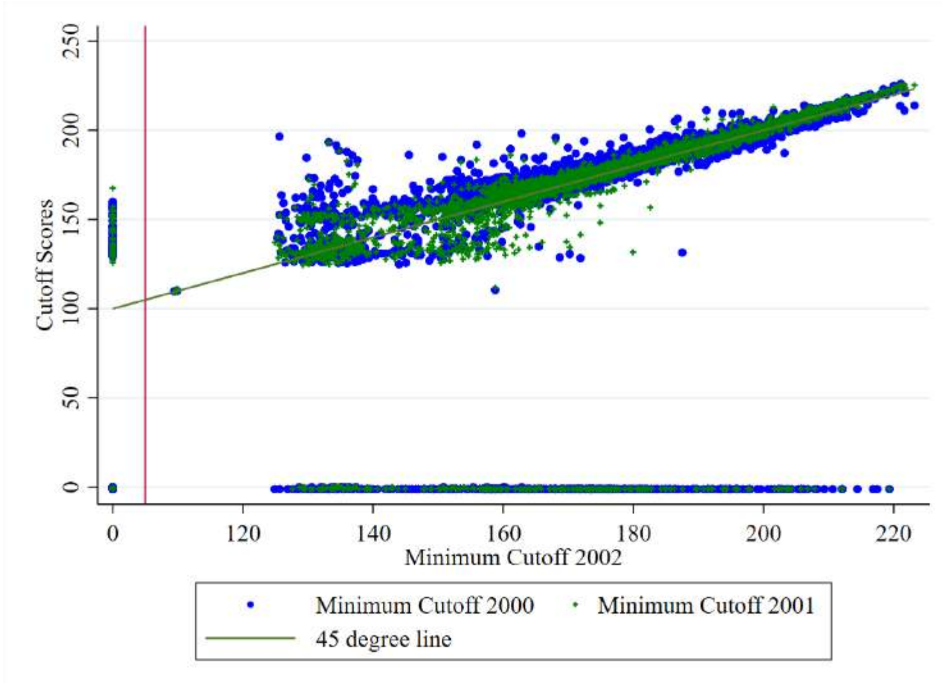
[18]In 1999 the rules changes so that instead of submitting preferences before knowing the score, they were submitted only after knowing them. This eliminated the effect of score uncertainty on the choice of preferences.

[19]The exception is the programs for which the cutoffs were not binding in some year. For these, the cutoffs could vary a lot across years. About 95 percent of these programs are located either in the Turkish Republic of Northern Cyprus or in Azerbaijan, which has an arrangement with Turkey which allows Turkish students to apply to their universities.

[20]The subjects that make up these majors are listed in Section D in the Appendix

[21]Figures A.14, A.15, A.16 in the Appendix present the percentage of male and female students in each college major for each of the three tracks separately.

Figure 1: Minimum Cutoff Scores



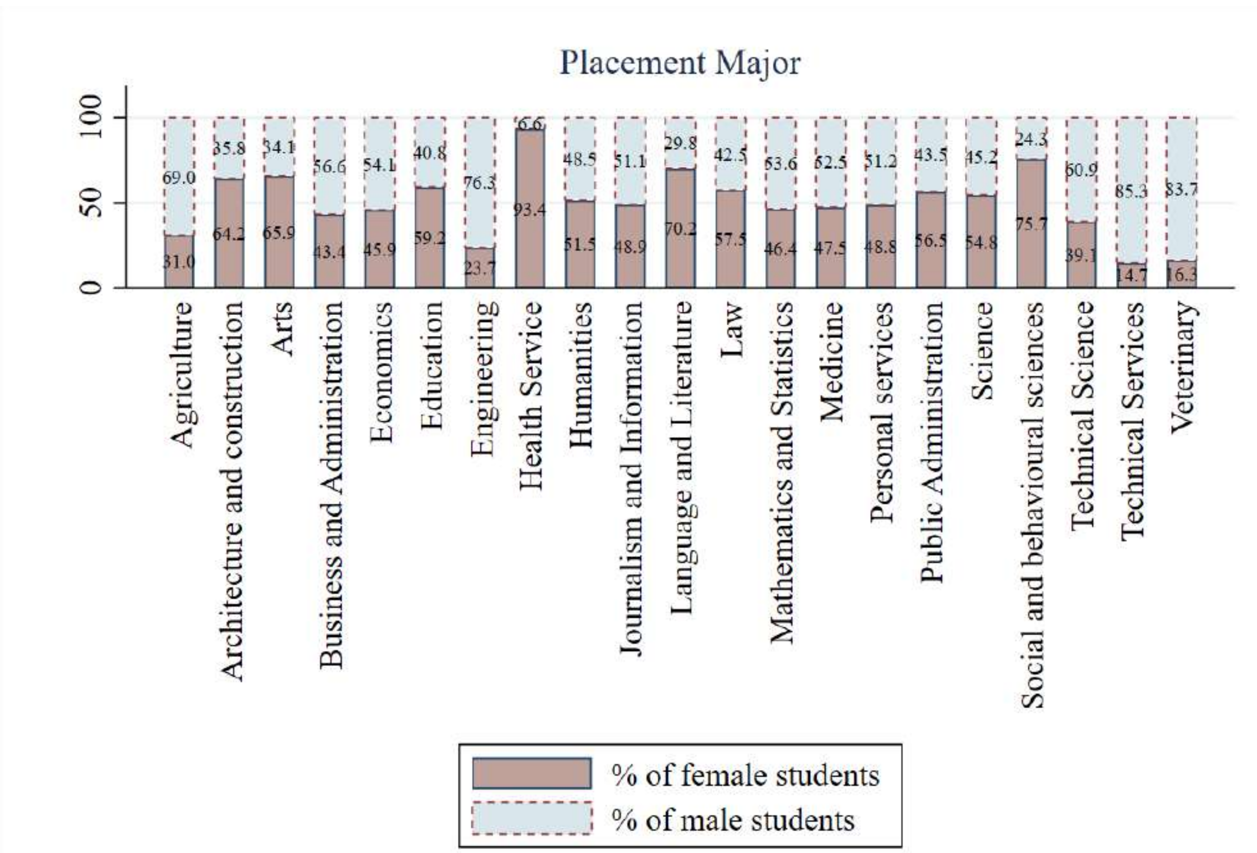
placement is that women seem have very different preferences.

Figure 3 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[22], while Figure A.17 does the same, but according to placement rather than preference.[23] 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 much more slowly with rank for men. This pattern is the result of systematic differences in the preferences of female and male students.

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

[23]The same graphs for the Social Studies and Turkish-Math track students are presented in Figures A.20 and A.21 in the Appendix.

Figure 2: Gender Differences in Major Choice (All Tracks)



Source: 2002 ÖSS applicants data

Figure 3: 1^{st} Preference Major (Science Track)

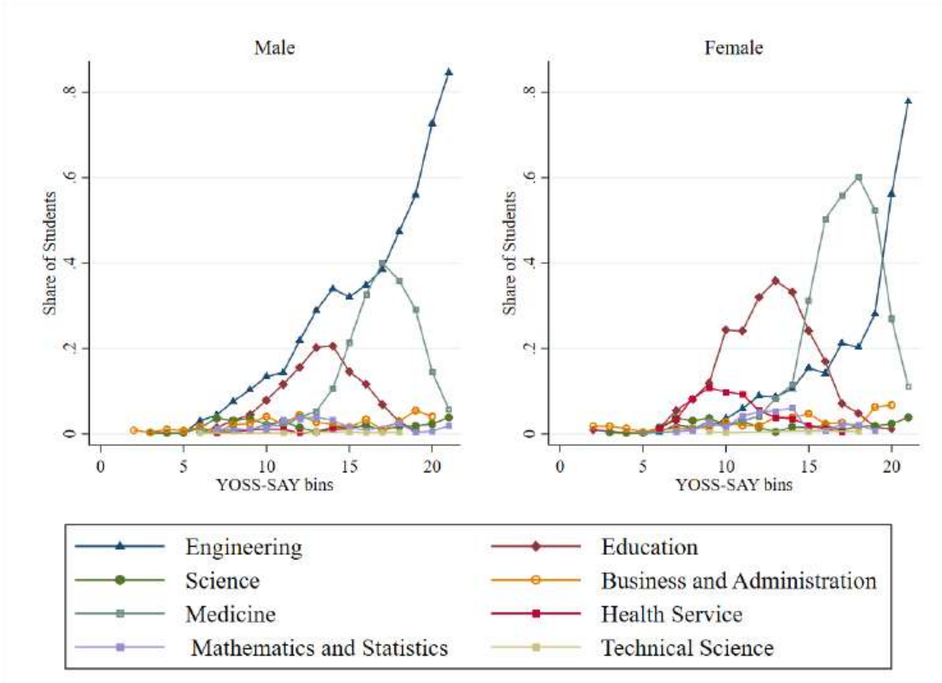
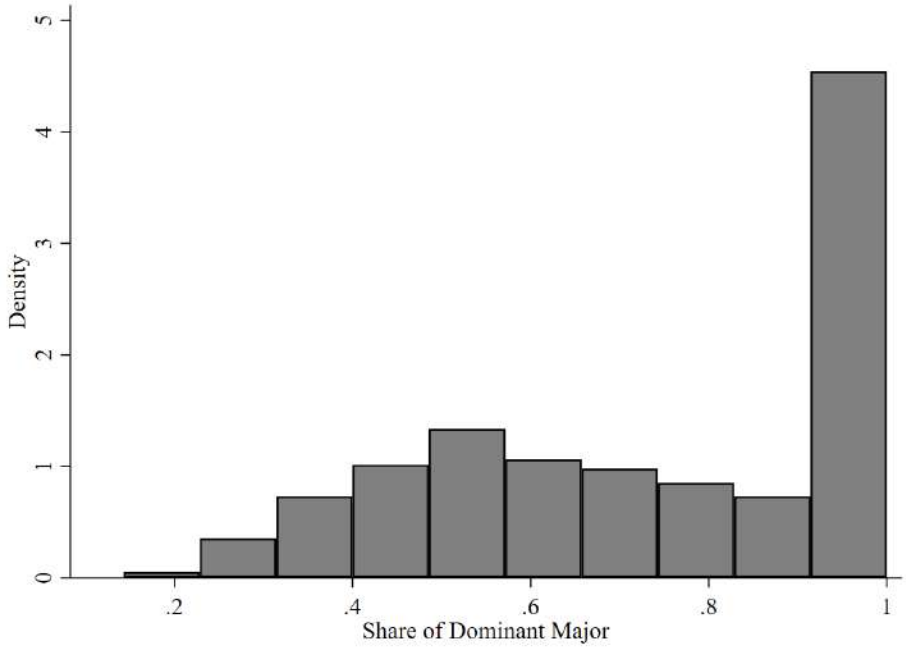


Figure 4 shows density of students according to the share of dominant major in their ranked preference list. It is clear from the figure that students fill their preference list with certain type of majors. This means that there are different types of students: some who say like medicine or engineering 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.

3.2 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? Saygin [2016], using data on Turkey for 2008, paints a picture that women aim lower than men, either because they are afraid of not being placed and having to retake, or because they are more risk averse and so drawn to safety schools. 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,

Figure 4: Distribution of Students according to Dominant Share of Major in Their Preference List



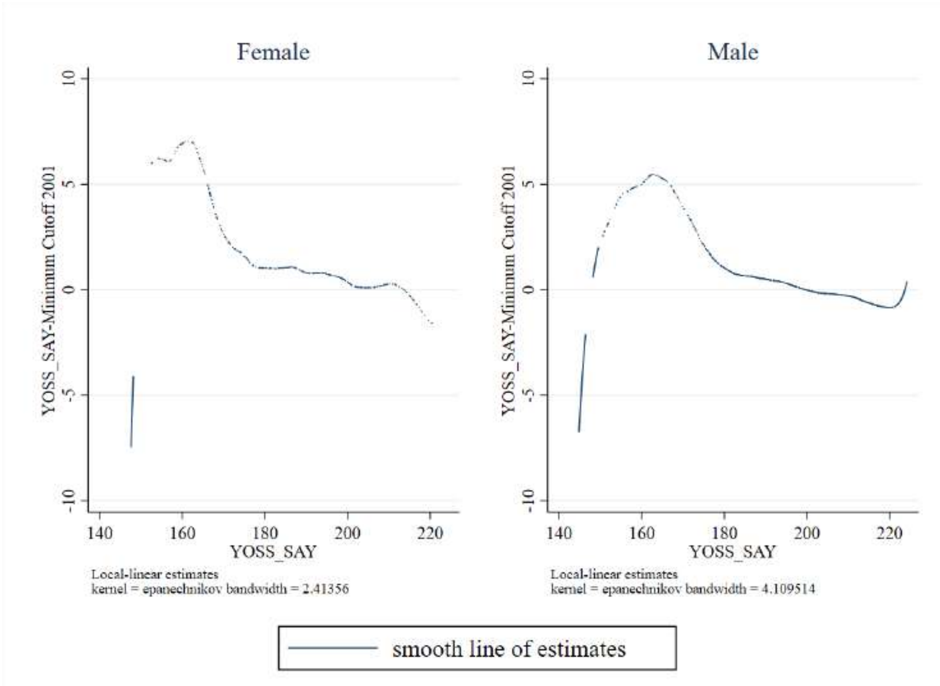
once we account for the majors students are placed in, women do not seem to aim lower than men. In other words, women tend to apply to majors where there is a larger dispersion of scores among students, rather than being less aggressive in their applications.

To begin with, we show that at least as far as applications go, women do not seem apply to less selective colleges *conditional* on score. Figure 5 plots non parametrically the difference in the placement score and the cutoff score by student for the program in which the student was placed. This is done for the Science Track.[24] Males are plotted on the right while females are on the left by gender. The figure suggests that weaker female students do seem to be less aggressive when applying, but that this difference vanishes as the placement score rises. As was described earlier, women and men have different patterns of applications and placements over the score interval. For example, men choose engineering even with a low score, while women do not.

In Table 2 we look deeper. We run the difference in the placement score and the cutoff

[24]The analogous figure for the Turkish Math track can be found in the appendix as Figure A.23. As there are not enough observations in the social science track, we do not plot this figure for this track.

Figure 5: Difference between Y-OSS Score and Minimum Cutoff of the Placement Department in 2001 w.r.t. Y-OSS score (Science Track)



score in 2001 for the program in which the student was placed on the variables in the data. In Column 1 we just run this difference on the male dummy. This gives a negative and significant coefficient of -0.58 suggesting that males on average are more aggressive in their applications. In Column 2, we control in addition for the city of the high school the student attended. This makes the coefficient slightly more negative. In Column 3, we add a number of controls. These controls have almost no effect of the mean gap though the coefficient on prep school expenditure is negative whenever such expenditure is positive suggesting the difference is lower for these subgroups. Finally, we add controls for the major in which the student was placed and the interaction of this major with the male dummy. The effect of this is startling. First, the male dummy we have been focusing on becomes insignificant and in fact becomes slightly positive suggesting that males on average are less aggressive in their applications. Second, the interaction of the male dummy and the major is never significant, though it is usually negative. This says that within each major of placement, males are not significantly more aggressive than females. Third, the major dummy is positive and significant for Health Service, Technical Science, Science and Vet Science. This says that all students applying to these majors tend to be less aggressive. In other words, these majors have a longer right tail in terms of the placement score 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 will look like men are applying more aggressively than women if we do not control as we do in Column 5. 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. Saygin [2016] in contrast, runs placement score on High School GPA and gender, along with other controls as well as controlling for the program placed. She does not does not interact gender with major of placement.

Table 2: 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***	1.020
	(0.167)	(0.172)	(0.168)	(1.178)
Income:				
250-500 TL			-0.194	-0.076
			(0.206)	(0.220)
More than 500 TL			0.135	0.453**
			(0.255)	(0.228)
Prep School Expenditure:				
No prep school			1.027	0.736
			(1.362)	(1.394)
Low			-1.622***	-1.531***
			(0.519)	(0.529)
Medium			-1.415***	-1.349**
			(0.523)	(0.518)
High			-0.265	-0.210
			(0.581)	(0.581)
Scholarship			-2.570***	-2.435***
			(0.522)	(0.550)
Parental Education				
Literate			-0.468	-0.980*
			(0.551)	(0.587)
Primary School			-0.099	-0.480
			(0.501)	(0.522)
Middle or High School			-0.238	-0.535
			(0.510)	(0.553)
College/Master/PhD			-0.443	-0.594
			(0.534)	(0.563)
Subject of Major				
Architecture and construction				0.712
				(1.333)
Education				1.484

(continued on next page)

VARIABLES	(1) YOSSSAY-Min	(2) YOSSSAY-Min	(3) YOSSSAY-Min	(4) YOSSSAY-Min
				(1.010)
Engineering				-0.305
				(1.026)
Health Service				3.527***
				(0.895)
Mathematics and Statistics				-0.109
				(0.987)
Medicine				1.375
				(0.931)
Science				3.089**
				(1.272)
Technical Science				11.205**
				(4.985)
Technical Services				2.241
				(1.536)
Veterinary				2.770*
				(1.465)
Male*Architecture and construction				-2.705
				(1.804)
Male*Education				-1.881
				(1.218)
Male*Engineering				-0.765
				(1.209)
Male*Health Service				-1.378
				(1.499)
Male*Mathematics and Statistics				-0.919
				(1.244)
Male*Medicine				-0.666
				(1.258)
Male*Science				-0.612
				(1.382)
Male*Technical Science				1.722
				(6.288)
Male*Technical Services				-1.031
				(1.951)
Male*Veterinary				0.821
				(1.797)
Observations	3,878	3,878	3,878	3,878
High School City FE	NO	YES	YES	YES

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

3.3 Do Women do Worse in the Entrance Exam?

Figures 6 and 7 present the cumulative distribution of exam scores (OSS) and placement scores (Y-OSS) by gender for first time takers and do so separately for each of the tracks. For students in each track, the weights used to calculate the placement score are those corresponding to the home track.[25] For the score used in the Science track programs (OSS-SAY) the male students' score distribution (in red) first order stochastically dominates that of

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

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). There are many explanations for the gender gap in such high stakes exams. The primary one seems to be that women perform worse under pressure than men, and/or that women do worse in high stakes multiple choice exams because they tend to not guess when it would be optimal for them to guess. Akyol et al. [2016], 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.

When it comes to the placement score, not the exam score (OSS vs. Y-OSS) the distribution for males first order stochastically dominates that for females (though less so than OSS-SAY), when we consider Y-OSS-SAY which is relevant for Science, but this is reversed when we consider Y-OSS-EA , which is relevant for Turkish-Math, and there is no apparent difference in the Y-OSS-EA distributions by gender. The reason the distributions for Y-OSS and OSS differ is because the Normalized High School GPA (AOBP) is part of the placement score and women do better then men in High School.[26] The distributions of AOBP are presented in Figure 8. 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

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

Figure 6: OSS Score Distributions by Gender

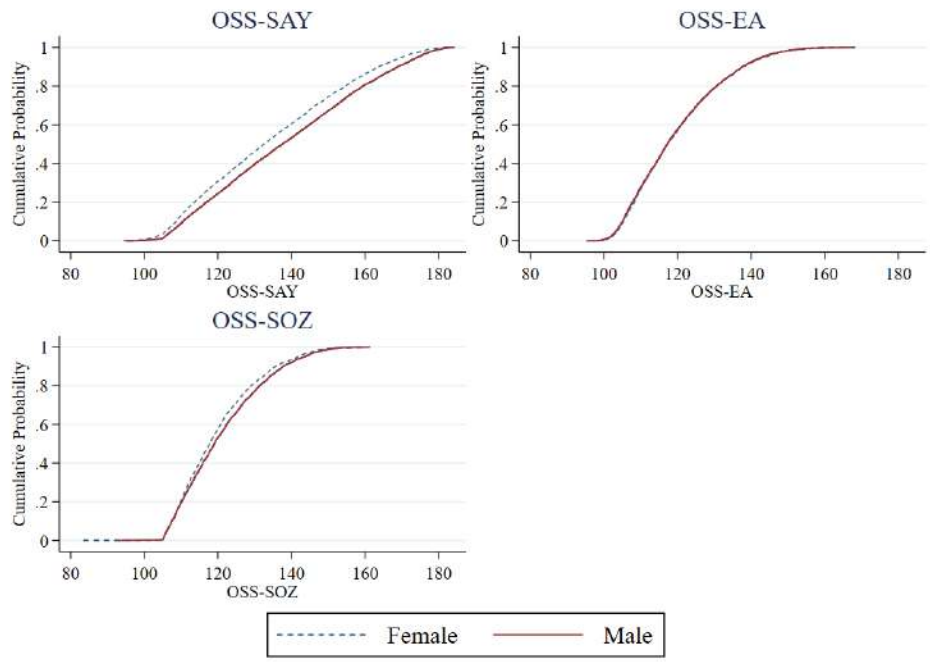


Figure 7: Y-OSS Score Distributions by Gender

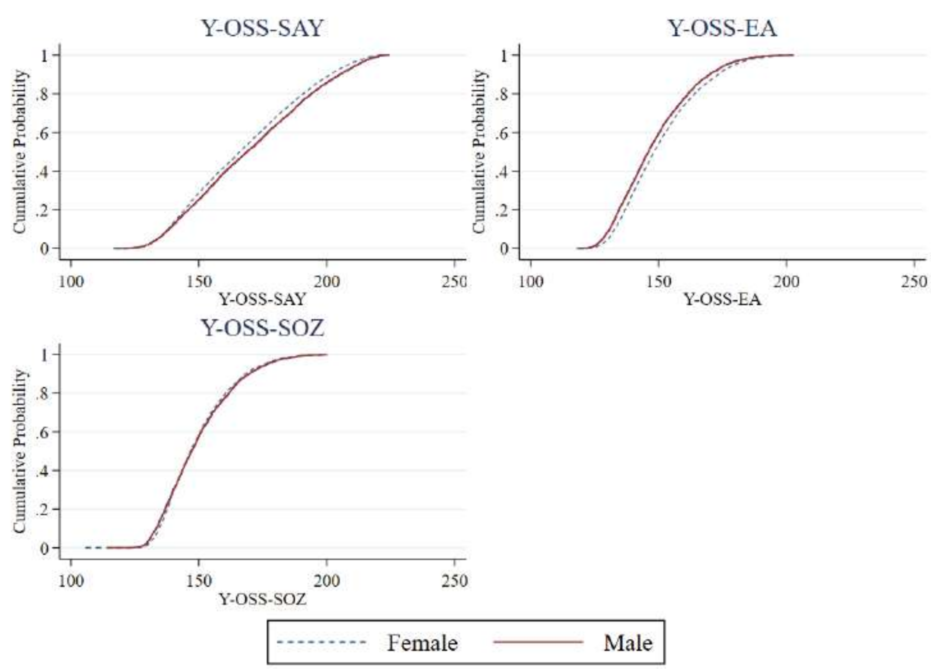
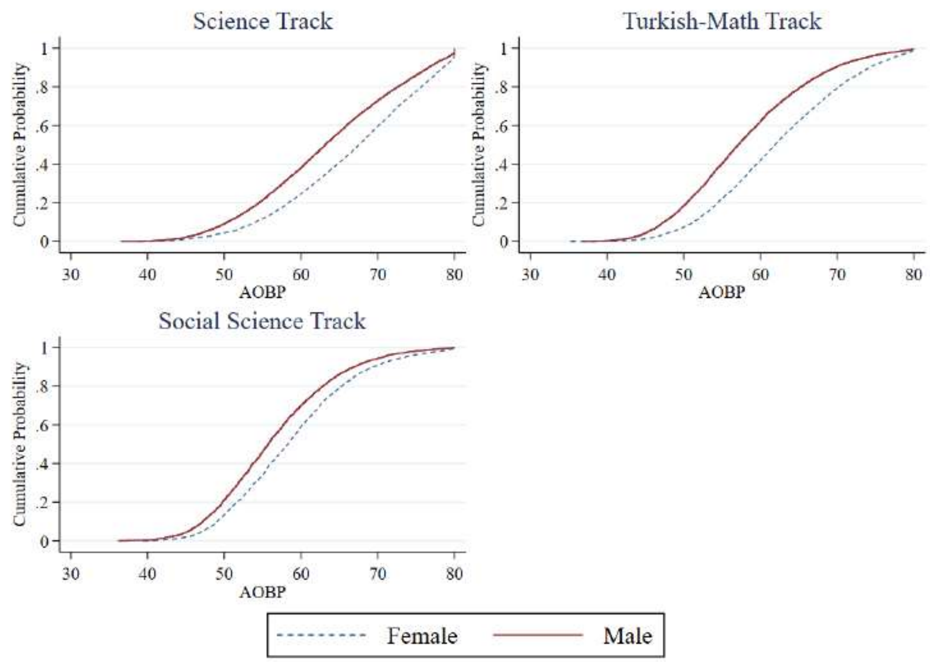


Figure 8: AOBP Distributions by Gender



often attributed to women maturing earlier then men.

4 The Score Gender Gap Function

Here we look at the gender gap in terms of scores so as to better understand what might be driving it. The gender gap is the difference in the score for males versus females, when they are otherwise equivalent. By definition, the gap would be zero if there was no intrinsic difference by gender and we were truly controlling for everything. In this case, the distributions of scores by gender, as a function of ability (proxied by say the performance in a standardized exam), should be the same. In practice we cannot control for everything. Nevertheless, we can learn a lot by controlling for whatever we can observe and then dropping controls to see how this affects the gender gap. A persistent and robust gender gap function even when we add controls suggests that there is a very limited scope for omitted variables to overturn the result. See Arcidiacono et al. [2020c] for a similar approach in examining racial bias against Asian applicants to Harvard.

In this spirit we run the following regression

$$OSS_{ij} = \alpha_j I_i(MALE) + \beta_j I_i(MALE)\,[AOBP] + \gamma_j I_i(MALE)\,[AOBP]^2 + \delta_j X_i + \varepsilon_{ij} \quad (1)$$

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.) X_i are the other individual level controls which can be thought of as the inputs into the production function of the subject score. Note that we allow for only AOBP to differ by gender in terms of its impact on score. We do so because the results from running the regressions with all the observables interacted with gender give the interaction with gender being significant only for AOBP and its square.

The observables (X_i) are the family background (the parents income group, the level of parents' education) the level of preparedness for the exam (the normalized high-school GPA and its square, 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. Using these controls controls for explanations based on parental investment in terms of high school choice, (controlled for by school fixed effects), preparation for the entrance exam (controlled for by prep course expenses), learning while in high school (controlled for by high school GPA[27]), selection on parental background (parents' education level). This is what we call Scenario 1[28].

A positive estimated level for this gender gap function, even with these controls may be driven by several other mechanisms. For instance, a gender gap in the exam scores could arise from a student's own effort. If for example, males do better in the exam, perhaps because the consequences of not doing well are worse for them, while they are slower to develop in high school and so do worse than girls there, then a positive gender gap would

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

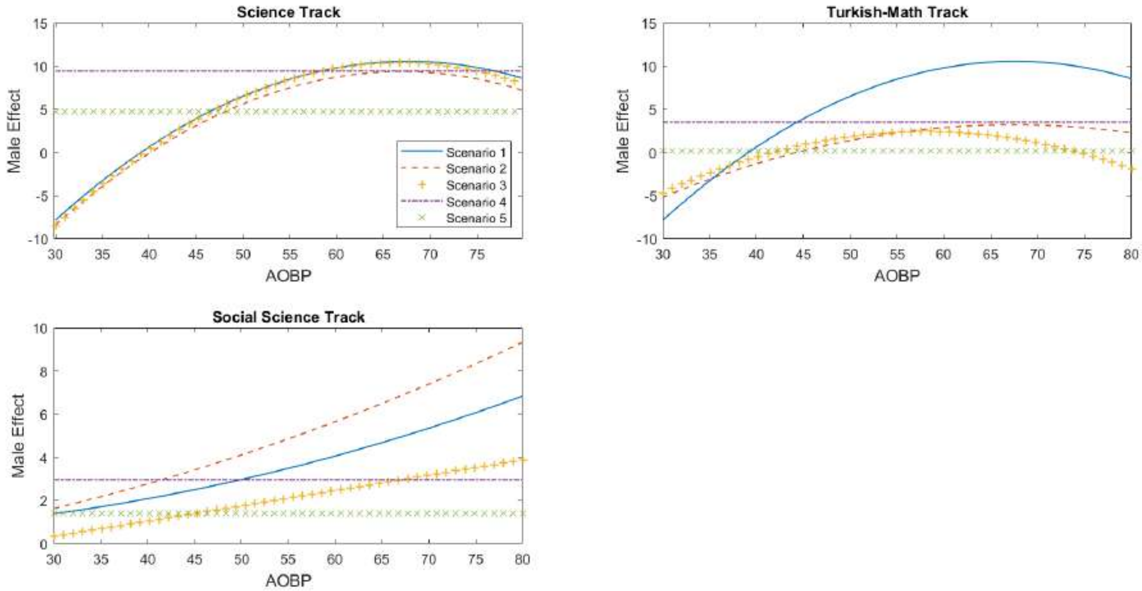
[28]In the appendix, see Figure A.24, we show that the quadratic form fits about as well as allowing the shape of the gender gap to be fully flexible. To do so, we first regress the score on the AOBP, its interaction with a male dummy, and observables in each AOBP score bin. We plot the estimates for the male dummy as a function of the AOBP score. For this reason, we feel safe using the quadratic specification.

arise in scores even after controlling for all observables. Or if girls are asked to do more chores than boys, they would have less time to spend on their studies and this would raise the gender gap. A negative estimate of the gender gap would arise if, for example, women did better in High School, but performed as well as males in the entrance exam. Recall that women do perform significantly better in High School, so that a negative and significant estimate for this function is what we would expect if other things given, women did as well as males.

The estimates are presented in Tables 3, A.7 and A.8 for the Science, Turkish Math and Social Science Tracks, respectively. The estimates for the Science Track are presented in the body of the paper. Estimates for other tracks are in the Appendix. Column 5 of Table 3 gives the mean gender gap in the science track to be 4.734 (and significant) when the only controls are for parental income and education, and where we have province fixed effects to control for differences between regions. This is positive suggesting males do better than females in the exam on average, given income. The estimate for α_j is negative for all j, but not significantly so for students in the Social Science tracks. It is slightly less negative for students in the Turkish Math track (-21.2 versus -48.9 in the Science Track) and highly significant. Column 4 adds the controls for $AOBP$ and $AOBP^2$ which allows performance in the exam to depend on performance in High School. This raises the mean gender gap to 9.425. Column 3 allows the gender gap to vary by AOBP and its square. The estimates show that the gender gap first rises and then falls with AOBP. Column 2 adds school fixed effects, thereby controlling for the school the student goes to. This could make a difference if school value added differs greatly. As is evident, there is little difference in Columns 2,3 and 4 in terms of the gender gap function. Column 1 further adds expenditure on prep school and whether the student obtained a scholarship for prep school. The estimates for the gender gap function are again relatively unchanged.

Once we include all observables as controls, as in Column 1, the shape of the function $\alpha_j + \beta_j [AOBP] + \gamma_j [AOBP]^2$ would give the level of the gender gap, as a function of AOBP, that remains once all observables have been accounted for. If our prior is that women are disadvantaged in general, then adding controls should reduce the gender gap, and *removing them should raise the gender gap.* By and large, except for when we remove AOBP, removing

Figure 9: Effect of being Male on OSS score



controls does not affect the gender gap in the Science track, suggesting that women are not advantaged or disadvantaged in general in the Science Track. When we remove the AOBP as a control, the gender gap *decreases*, because women are actually advantaged in terms of the AOBP.

We plot the gender gap function as for each track for each of the scenarios described below. This is depicted in Figure 9 which depicts the gender gap in each scenario over the range 30-80. In the first scenario, we include all the controls. The estimates are given in Table 3.The gender gap function for this scenario is drawn in Figure 9 as a blue solid line. Note that the gender gap is actually negative for low ability (AOBP) men suggesting that women actually score more than men do, and rises with ability peaking at a score of 65 and falling slightly after that. It has roughly the same shape in Scenario 1 in the Turkish Math track, and is roughly linear and increasing in the Social Science track. This shows that it is the male students with a higher AOBP who tend to do better than females in the entrance exam and highlights the importance of making the gender gap a function of AOBP.

In the second scenario, we include the same controls *except* for expenses on preparatory courses. If parents are less willing to cover preparatory school fees for their daughters than for their sons, the effect this investment asymmetry on scores should be picked up by the new estimate of the gender gap function. If this estimate is more negative than under scenario 1, it suggests that part of the reason boys do better in the exam is that they are given more resources to study. As we see, this function, drawn in orange dashed line in Figure 9 is roughly the same as in Scenario 2 for the Science track. This is entirely in line with what we described in terms of data patterns in the previous section where if anything, girls have slightly more spent on them in terms of such expenditures.[29] In the Turkish Math track, the curve for Scenario 2 lies below that for Scenario 1, suggesting that girls may actually be getting more spent on them in this track. In Social Studies, the curve for Scenario 2 lies above that for Scenario 1, suggesting that girls may be getting less spent on them in this track.

The third scenario replaces school fixed effects with province dummies. Thus, if parents are more likely to send their sons to elite schools, for example because they are far away, and these schools have greater value added, not controlling for schools would raise the gender gap. The function does not seem to change much here compared to Scenario 2 in either the science or Turkish Math Tracks, but is lower for the Social Studies track suggesting that here again, if anything, women are favored by their parents.

In the fourth scenario, we exclude the interaction of the male dummy with AOBP and its square in the regression, leaving just the dummy for being male. This restricts the regression to giving a uniform gender gap (no matter what be the AOBP), while controlling for AOBP in the regression. the estimate of α_j is positive and significant for all tracks, 9.43, 3.65 and 3.2 in the Science, Turkish Math and Social Science tracks respectively. This highlights the importance of allowing for the gender gap to vary by AOBP. Not all males do better than females, the better ones do, but the worse ones do not as is evident from comparing the gender gap function from Column 1 and Column 4. Note that we still control for parents' education, income and geography.

In the fifth scenario we drop AOBP and its square from the controls as well. Recall that when we look at the distributions of exam scores for males and females, the men do slightly

[29]In part, this may be due to selection against girls finishing high school and taking the university entrance exam as shown in the data patterns.

better. Since women do better than men in AOBP score, controlling for the AOBP score will make the gender gap rise. Thus, dropping it will make the gender gap fall. The results are reported in column 5. The gender gap falls to 4.73, 0.3 and 1.6 in the Science, Turkish Math and Social Science tracks respectively.

Table 3: Gender Gap in OSS Score: Science Track

	(1)	(2)	(3)	(4)	(5)
VARIABLES	OSS-SAY	OSS-SAY	OSS-SAY	OSS-SAY	OSS-SAY
Male	-48.880***	-48.924***	-51.990***	9.425***	4.734***
	(9.110)	(9.226)	(10.061)	(0.303)	(0.368)
AOBP	-1.719***	-1.795***	-1.624***	-0.378**	
	(0.223)	(0.226)	(0.263)	(0.182)	
$AOBP^2$	0.022***	0.023***	0.023***	0.014***	
	(0.002)	(0.002)	(0.002)	(0.001)	
Male*AOBP	1.758***	1.741***	1.869***		
	(0.291)	(0.296)	(0.326)		
Male*$AOBP^2$	-0.013***	-0.013***	-0.014***		
	(0.002)	(0.002)	(0.003)		
Income(base: Less than 250 TL)					
250-500 TL	-0.590*	-0.216	1.284***	1.270***	0.856*
	(0.302)	(0.306)	(0.367)	(0.367)	(0.450)
More than 500 TL	-0.881**	-0.512	3.703***	3.686***	3.037***
	(0.384)	(0.382)	(0.453)	(0.454)	(0.597)
Prep Expenditure(base: Missing)					
No prep school	-1.982***				
	(0.570)				
Low	5.157***				
	(0.495)				
Medium	4.544***				
	(0.544)				
High	4.505***				
	(0.638)				
Scholarship	7.389***				
	(0.727)				
Parental Education					
Literate	-0.886	-1.456*	-5.493***	-5.524***	-4.658***
	(0.760)	(0.780)	(0.902)	(0.907)	(1.094)
Primary School	0.792	0.428	-2.775***	-2.793***	-3.147***
	(0.524)	(0.530)	(0.645)	(0.648)	(0.830)
Middle or High School	0.730	0.663	0.225	0.169	-0.257
	(0.490)	(0.497)	(0.607)	(0.609)	(0.800)
College/Master/PhD	1.351***	1.341***	6.149***	6.113***	8.910***
	(0.507)	(0.512)	(0.620)	(0.622)	(0.850)
Observations	13,505	13,505	13,505	13,505	13,505
School FE	YES	YES	NO	NO	NO
Province FE	NO	NO	YES	YES	YES

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

5 Modeling of College Preferences

5.1 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, X_{ij}.

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

By choosing program j, the student obtains utility

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

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 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:*

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

If students knew the cutoffs when they made their lists, or had rational expectations about them, then the first assumption would be true. If students used last years cutoffs as their best guess about the next years cutoffs, then the second assumption would be true. If cutoffs fluctuate but their ranking is unchanged over the years, then both assumption 1 and 2 would be true if either was true. Chade and Smith [2006] show that if there is uncertainty in the cutoffs, and this uncertainty is not correlated across schools, then the optimal list may have holes in it, but whatever is on the list must be in the order they have in preferences. Shorrer [2019] shows this is also so when there is perfect correlation between schools in terms of the uncertainty. A rank order list that is inconsistent with one's preferences (that does not respect their true order) is weakly dominated, see Haeringer and Klijn [2009]. This is what motivates the third assumption. Note that we only assume that this is true for the placement under 2001 and 2002 cutoffs, not for everything on the list as these element are more salient for the student. Programs that the student does not expect to placed in are less salient and we do not assume they are ranked correctly relative to these two.

We use the above three conditions to implement a maximum likelihood estimator for the key preference parameters: $\beta, \gamma_t, \sigma_t, t = 1, \ldots, T$. The likelihood function is derived in

Appendix B.

We estimate the model independently for male and female applicants in three major high school tracks (Science, Turkish-math and Social Science). To avoid selection issues caused by exam retaking, we only include those applicants who never took the college entrance exam in the past. 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.C.

5.2 Demand Estimates

Table 5 presents the estimates of the common parameters, β, by high school track and gender. The first set of variables are dummies for the program being a distance program interacted by the income level bin. This assigns a utility to distance programs for applicants in each income bin. If a program is a distance program, then the program also had a dummy for being offered in this manner. Distance programs are liked a lot less than regular programs.

The next variable is distance which is the distance between the program's campus and the high school attended by the student. Programs geographically remote from the applicant's high school tend to be valued less. In all three high school tracks, females have 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.[30] A female applicant with the same background would pay 2,291 Turkish liras.[31] 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..

[30]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.

[31]The above numbers are roughly similar to 958 US dollars for males and 1,527 US dollars for females in 2002.

The next set of variables in Table 5 are dummies for the program being an evening program interacted with the income bin. This assigns a utility to evening programs for applicants in each income bin. Evening programs seem to be less disliked than distance ones. These are the same programs offered in the day, but as they have lower cutoffs typically, they may be preferred by some students. Typically, students do not work while attending school in Turkey.

The next set of variables involve the variable Placement score SAY, EA or SOZ. These variables are dummies that are equal to 1 if the program uses YOSSSAY, YOSSEA or YOSSSOZ respectively as the placement score. It is interacted with the variable no rank, which indicates the program was unranked, probably because it was new, and with the variables rank and rank squared. This gives a higher base utility to higher ranked programs and allows utility to vary in a quadratic manner with rank. As can be seen, higher rank is associated with greater utility. Even when base utility falls to begin with, the turning point is quite low.

Applicants have strong distaste for high tuition. In line with common wisdom, applicants from more well-off families tend to be less sensitive to tuition. The next variable is the same province dummy. As this is positive, applicants also prefer to stay in the same province even after controlling for distance.

The estimates for γ_t are reported in the Appendix in Tables A.9 to A.14. Consider Table A.9 which reports the estimates for women in the Science Track. There are eight unobserved types for this regression. For the Science track, students place into more than eight majors. Increasing the number of types is computationally costly, we limit the number of types to the number of majors the students place into or eight, whichever is lower. The estimated coefficients for each type are reported in each of the 8 columns. The probability that an agent is of a particular type is reported in the last row. Some types are more likely than others: type 4 and 8 are 6 to 9 times more likely than type 1 or 2. The variable placement score SAY (EA) is a dummy that is 1 if the program uses YOSSSAY (YOSSEA) as the placement score. This raises the utility for such programs so that such students in the science track also consider such programs. The next set of variables are the program dummies. The omitted program is education, so a positive coefficient on a program dummy means that for this type, such programs are better than education. Type 1 students have a positive coefficient on Medicine and Health Service which mostly accept YOSSSAY, so we can think of science track students of type 1 as being those who like medical or nursing schools. The third set of estimates captures the value of the outside option for the student as a function of student's three scores (YOSSSAY, YOSSEA or YOSSSOZ) and the high school GPA. Reported in the table are the estimates of the interaction of the dummy for not being placed with these three scores and the GPA. The outside option is retaking. This is more likely the higher is the difference in your predicted score and your actual score. In other words, retakers tend to be those who did worse than they should have. This is captured in a flexible way by interacting not being placed (recall there is a penalty for being placed if you retake) and the three scores and GPA.

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, 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 assumptions 1 only. In each case for the models in Column 1-4, we estimate the model and then generate the placement based on the estimates. In column 5, we assume preferences are as given by the students list and generate 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 say 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

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	8.1%	12.2%	11.5%	11.1%	8.6%

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

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, 8.6% should be the fraction of switches in major that a model that represents preferences well should obtain. Then 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. 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 to better understand how well our chosen specification does, we create 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 the Appendix in Figures A.25 to A.29.

In Figure A.25 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.26. 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.27 to A.29 in the Appendix.

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 91.9% when we use the fitted model and 91.4% when we 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.[32] 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 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

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

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.25 to A.29 differ from one another. To make the difference easier to see we present a heat map of the differences in Figures **??** to A.29 and Figure A.25. 10 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.

Table 5: Estimated demand parameters, common coefficients β

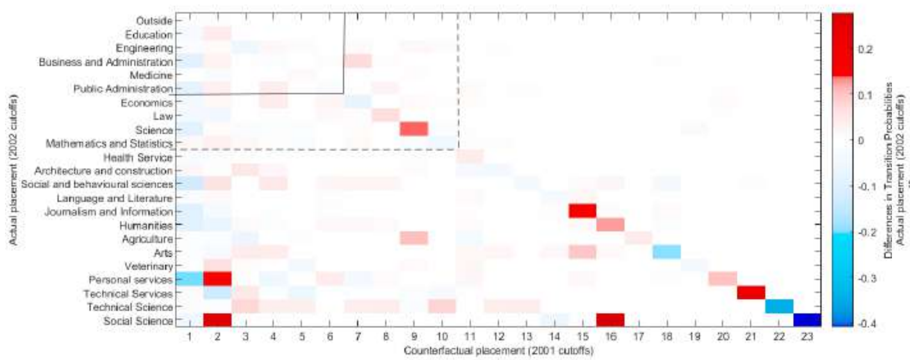
Track	Science		Turkish-Math		Social Science	
Gender	Female	Male	Female	Male	Female	Male
Variable						
Dist program *Income=1	-6.11	-5.36	-3.09	-0.21	-0.62	-6.99
Dist program *Income=2	-6.53	-5.68	-2.63	-0.50	-1.37	-3.92
Dist program *Income=3	-5.22	-4.31	-2.52	0.71	-0.51	-2.63
Distance	-2.86	-1.85	-2.48	-1.60	-2.25	-1.71
Evening program *Income=1	-0.02	-0.02	-0.38	-0.12	-0.37	0.02
Evening program *Income=2	-0.01	0.21	-0.15	0.03	0.55	0.95
Evening program *Income=3	0.32	0.48	-0.07	0.10	0.14	-0.09
Placement score: EA * no rank	-2.69	-3.87	0.61	1.01	1.66	-7.06
Placement score: EA * rank	-17.77	-22.67	-1.51	-1.98	-0.92	-2.13
Placement score: EA * rank^2	25.87	30.45	7.75	9.48	7.89	7.24
Placement score: SAY * no rank	3.01	1.48				
Placement score: SAY * rank	0.29	-6.29				
Placement score: SAY * rank^2	14.51	20.47				
Placement score: SOZ * no rank			7.39	3.32	5.53	0.11
Placement score: SOZ * rank			2.31	-6.05	6.52	-13.15
Placement score: SOZ * rank^2			12.25	17.70	5.08	22.73
Same province	1.24	1.09	1.47	1.49	1.68	1.42
Tuition *Income=1	-12.48	-12.86	-8.70	-9.52	-10.57	-10.08
Tuition *Income=2	-10.99	-10.40	-8.36	-7.97	-10.41	-10.17
Tuition *Income=3	-6.98	-6.58	-4.56	-4.49	-6.31	-6.59

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.

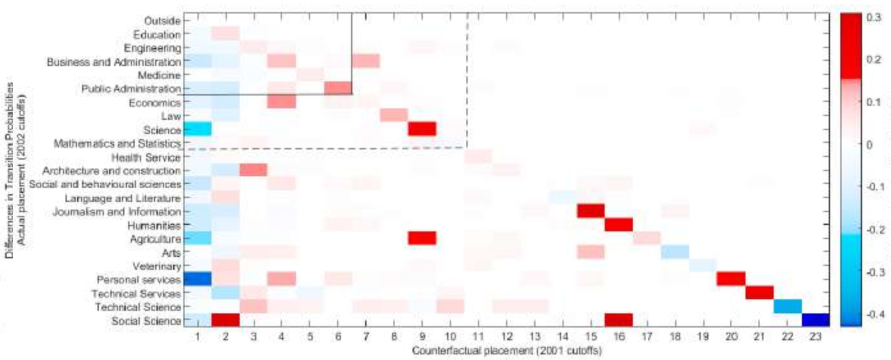
Units: Tuition — 10,000 TL, distance — 1,000 km, rank — varies from 0 (lowest cutoff among programs accepting the same type of score) and 1 (highest cutoff). No rank – an indicator variable for programs not included in the 2001 ranking.

Figure 10: Differences in Transition Probabilities between Model and Data for Different Model Specifications

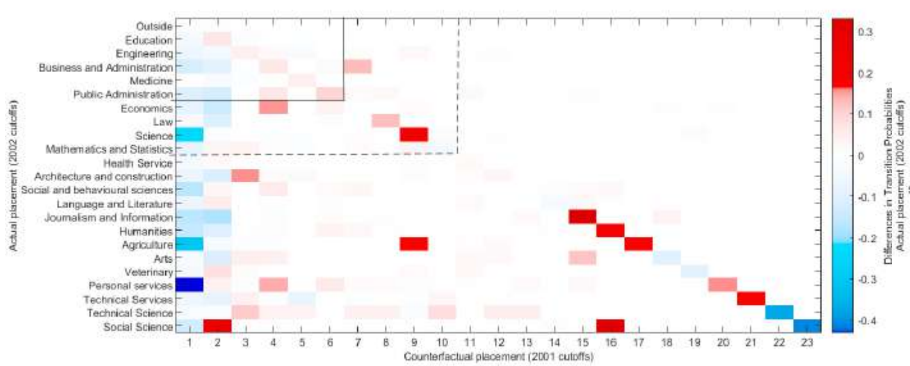
(a) Main Model



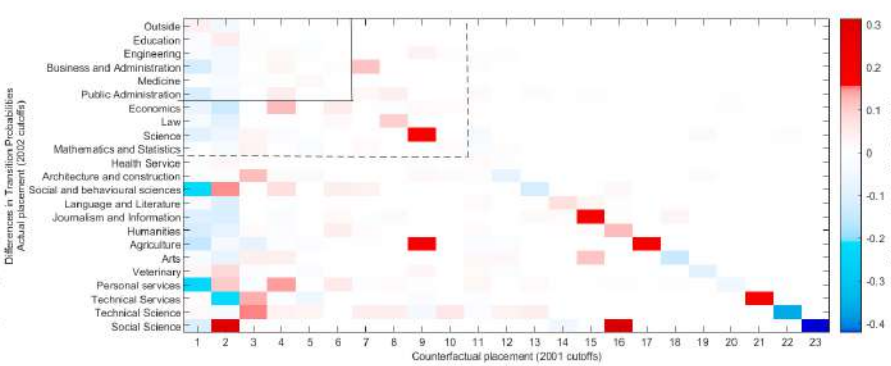
(b) Alternative 1



(c) Alternative 2



(d) Alternative 3



6 Policies Targeting the Gender Gap in Scores

In this section we look at placements by gender across major tracks under various counter-factual scenarios for students in the three tracks.

Table 11, 12 and 13 show what these placements are and how these placements change across the scenarios considered for first time takers in the Science, Turkish Math and Social Studies tracks respectively. 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. But for students from the Social Studies track, there is no real difference. We look at scenarios as in Table 3 for students in the Science track, Table A.7 and A.8 for students in the Turkish Math and Social Science tracks respectively. The status quo placements and the counter-factual placements are shown in Figure 11.Note that counterfactuals are across all 3 tracks.

The brightest red and blue bars show the placements for females and males respectively under the status quo. As can be seem men dominate in Engineering while women dominate in Education and Health Services. In each scenario that follows, we give each woman in each track, the points (negative or positive) that she would have obtained had she been male in the same track with the same AOBP score. We then compute her use her allocation score in each program. Recall that different programs put different weights on the different component exams and have penalties for applcants outside the main track so that each student has three allocation scores, with only one of them being relevant for any given program. We then use our estimated preferences and these allocation scores to allocate her to the program which she most prefers among those that are feasible for her, given equilibrium cutoffs. These equilibrium cutoffs are generated endogenously. We deal with repeat takers and students from the language track as a fringe in these simulations. They obtain the same points as a first time taker would have. However, since their preferences are not estimated, we use theit lists as their preferences and allocate them to the program that is highest on their preference list and is feasible. This is why, for example, despite the fact that only Science track students can go to Engineering, the increase in female presence is larger than the decrease in the male presence. Corresponding changes among repeat takers are occurring since the total number of seat in Engineering is fixed.

In Scenario 1, the adjoining bar, we give females the exam score that they would have obtained, given characteristics, had she been male as we control for all observables. We then compute the allocations using the estimated preferences, assuming that cutoffs adjust as needed to equate allocation to seats available, i.e., we look at general equilibrium. Consider for example, Engineering placements for students in the Science track. Table 11 shows that female presence in Engineering programs rises from .07 to .095, while that of males falls from .17 to .16. Recall also that women get a different bonus in terms of points depending on characteristics. A woman with a score of 180 will get a relatively high bonus and as such women are more likely to apply for Engineering programs (recall good women considered these programs, while less able women did not) the impact on Engineering will be relatively large. Only students from the Science track are in practice able to do Engineering[33] so that there are no placements in Engineering from other tracks. In contrast, consider Economics or Business Administration. In Scenario 1, students from the Science track decrease their presence in these majors, probably because they switch to Engineering. However, students in Turkish Math increase their presence in Economics and Business Administration.

Scenario 2 is labeled prep nl. This gives females the exam score males would have obtained, without controlling for *prep school* expenditure. If men were favored in terms of prep school expenditure, and this was why women got lower scores and so could not choose competitive majors like Engineering, this scenario would reduce the presence of women in Engineering. As is evident, this slightly raises the presence of women in Engineering relative to the first counter-factual, and does little elsewhere, so we conclude that this is not the main driver of placements.

Scenario 3 is labeled hschool nl gives the counter-factual placement if women had the score men would have had without controlling for both prep school expenditure and high school fixed effects. If men were favored in terms of sending them to better high schools, and this was why women got lower scores and so could not choose competitive majors like

[33]The penalty for students from other tracks in Engineering is high enough that they are in effect banned.

Engineering, this counterfactual should increase the presence of women in Engineering. This scenario actually reduces the presence of women in Engineering and raises that of men. This is expected as we saw earlier that women do not seem to be going to worse schools, see Table 1.

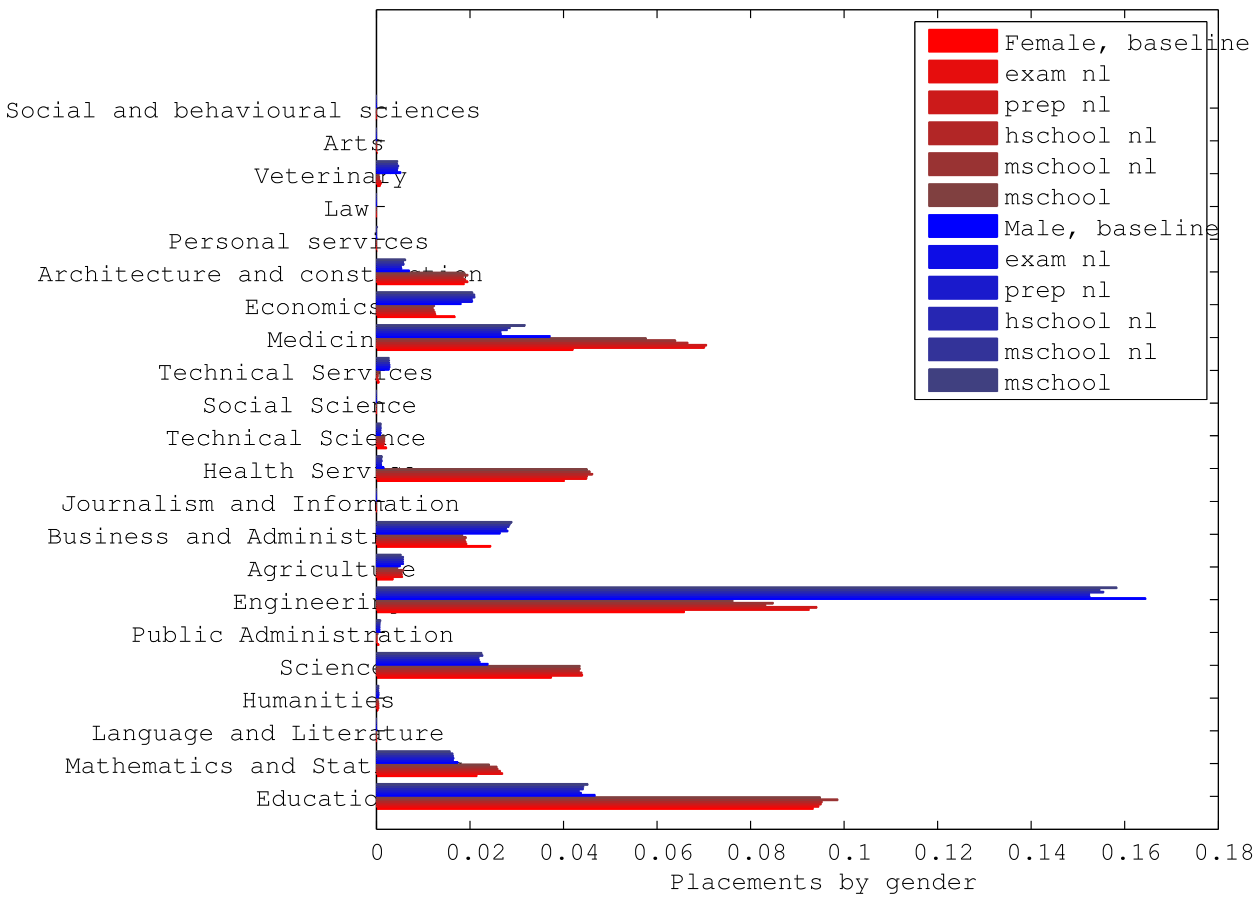
Scenario 4, labeled mschool nl drops the interaction of the male dummy with AOBP while controlling for AOBP and its square, so that it only gives women the average gender gap increase in score. This means women with AOBP scores around 70 get a bit less, while other women get more points than in scenario 3, see Figure **??**. This slightly increases the presence of women in Engineering as the average gender gap estimated is quite high, over 9 points.

The fifth scenario also drops the controls for AOBP and its square. In this scenario, the average gender gap is lower than that in Scenario 4. This quite dramatically reduces the presence of women in Engineering. Giving all women a bonus raises competition in programs women like more than in programs women tend to dislike. As a result, the fraction of women in Health Services, at least among first time takers) does not change much in scenarios 4 and 5, and actually falls in education in scenario 5 relative to 4.

We perform four more counterfactuals. In the first, we just give women the preferences of men and see how their placement changes in general equilibrium. In the second, we give women in the East, which is more conservative, the preferences of the women in the West. Third, we give all women a 5 point bonus and see how placements change. In the fourth, we give women a five point bonus, but only in majors where males dominate in placement. [These counterfactuals are coming.]

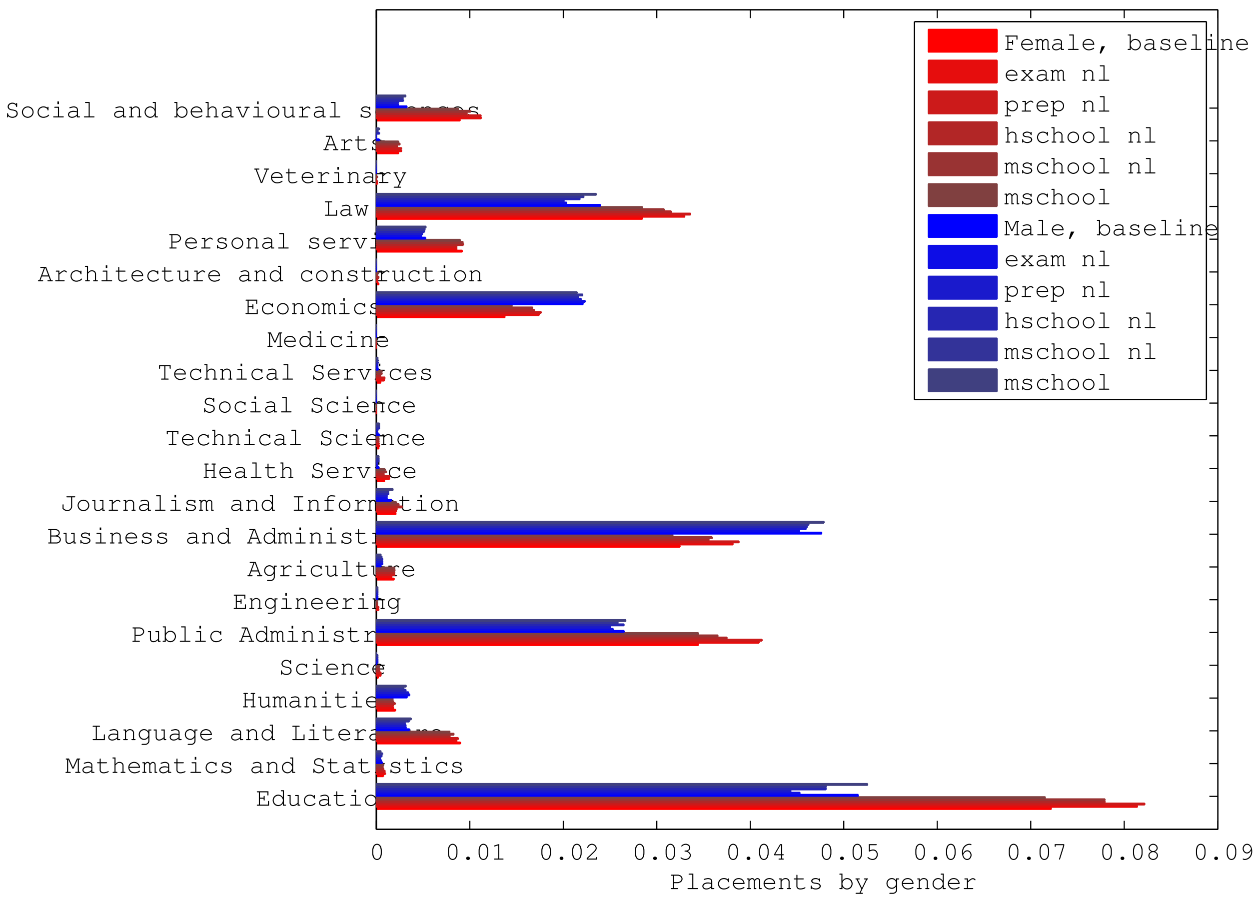
7 Conclusion

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 the score gender gap is real: women do perform worse than men, but this is more so in the middle of the ability distribution. At the low end, they actually perform better than men and at the top they do less badly than in the middle.



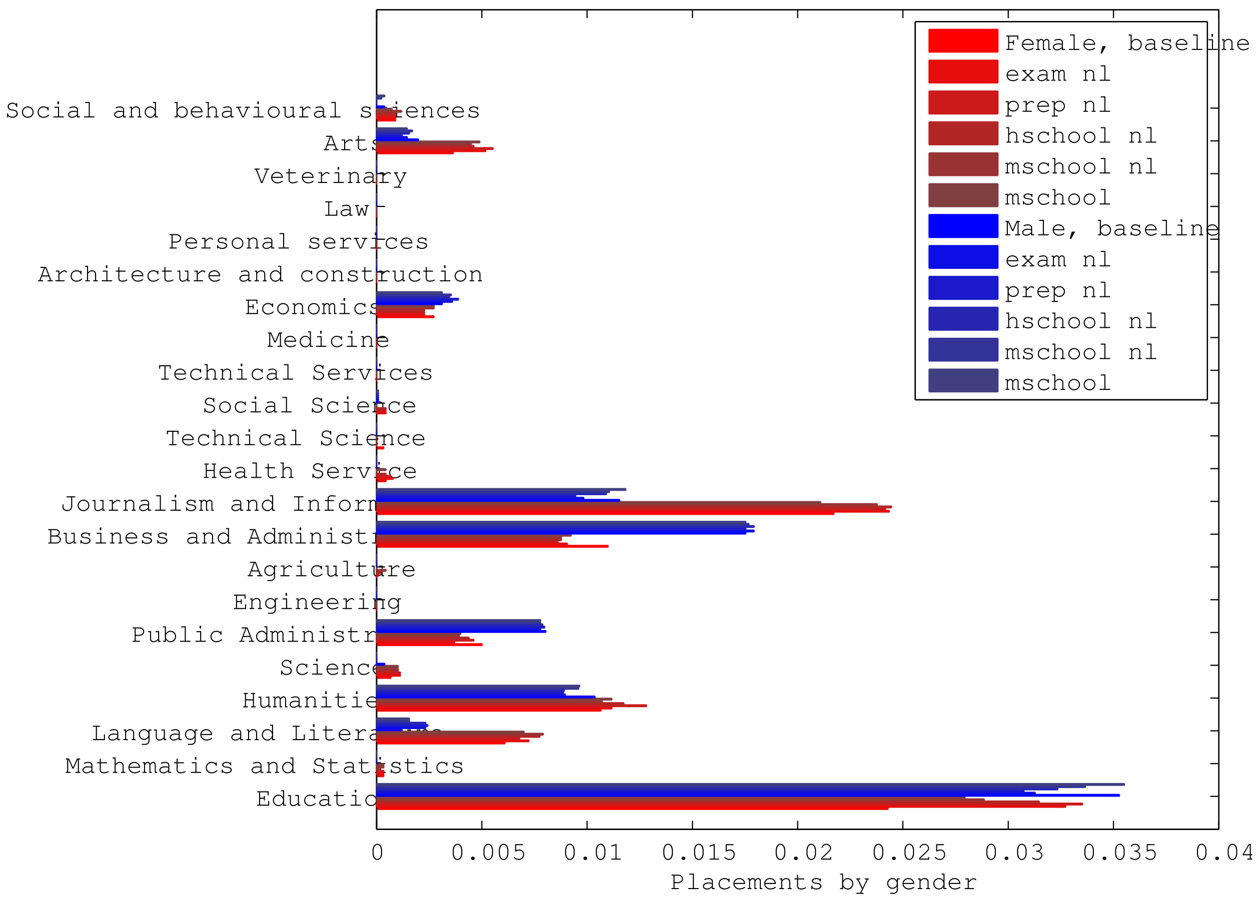
Notes: The counterfactual policies adjust female scores according to the gender gap regressions similar to those in Table 3, following the same order, but using HGPANORM instead of AOBP. Placements to the outside option is omitted. Using first time takers from the estimation sample only.

Figure 11: Predicted placements by gender, science track



Notes: The counterfactual policies adjust female scores according to the gender gap regressions similar to those in Table A.7, following the same order, but using HGPANORM instead of AOBP. Placements to the outside option is omitted. Using first time takers from the estimation sample only.

Figure 12: Predicted placements by gender, Turkish-Math track



Notes: The counterfactual policies adjust female scores according to the gender gap regressions similar to those in Table A.8, following the same order, but using HGPANORM instead of AOBP. Placements to the outside option is omitted. Using first time takers from the estimation sample only.

Figure 13: Predicted placements by gender, Social Studies track

Third, that while differences in performance can account for a small part of the placement gender gap, differences in preferences are more important. Fourth, 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.

References

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

Ş Pelin Akyol, James Key, and Kala Krishna. Hit or miss? test taking behavior in multiple choice exams. Technical report, National Bureau of Economic Research, 2016.

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, Josh Kinsler, and Tyler Ransom. Recruit to reject? harvard and african american applicants. Technical report, National Bureau of Economic Research, 2020b.

Peter Arcidiacono, Josh Kinsler, and Tyler Ransom. Asian american discirimination in harvard admissions. Technical report, National Bureau of Economic Research, 2020c.

Peter Arcidiacono, Josh Kinsler, and Tyler Ransom. Legacy and athlete preferences at harvard. *Journal of Labor Economics*, forthcoming.

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

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.

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.

Yan Chen and Tayfun Sönmez. School choice: an experimental study. *Journal of Economic theory*, 127(1):202–231, 2006.

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.

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.

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. 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.

Linda Datcher Loury. The gender earnings gap among college-educated workers. *Industrial and Labor Relations Review*, 50(4):pp. 580–593, 1997.

Linda Datcher Loury and David Garman. College selectivity and earnings. *Journal of Labor Economics*, 13(2):pp. 289–308, 1995.

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.

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.

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.

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

A Appendix

A.1 Additional Tables and Figures

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

	(1)	(2)	(3)	(4)
VARIABLES	YOSSEA-Min	YOSSEA-Min	YOSSEA-Min	YOSSEA-Min
Male	-0.924***	-1.126***	-1.240***	0.795
	(0.263)	(0.294)	(0.340)	(0.568)
Income:				
250-500 TL			-0.290	-0.032
			(0.291)	(0.236)
More than 500 TL			0.763	1.280**
			(0.588)	(0.531)
Prep School Expenditure:				
No prep school			-0.741	0.028
			(1.113)	(0.730)
Low			-1.205	-0.465
			(0.808)	(0.510)
Medium			-2.517***	-1.289**
			(0.851)	(0.595)
High			-1.788*	-0.551
			(0.955)	(0.588)
Scholarship			-0.643	0.338
			(1.116)	(0.725)
Parental Education				
Literate			0.724	0.080
			(1.005)	(0.934)
Primary School			0.017	-0.392
			(0.699)	(0.610)
Middle or High School			-0.010	-0.230
			(0.451)	(0.518)
College/Master/PhD			-0.142	-0.037
			(0.512)	(0.674)
Subject of Major				
Economics				-0.854**
				(0.324)
Education				6.770***
				(0.565)
Humanities				0.127
				(0.984)
Journalism and Information				5.444
				(4.034)
Law				4.117***
				(0.533)
Personal services				0.059
				(0.572)
Public Administration				0.799
				(0.569)
Social and behavioural sciences				1.632***
				(0.604)
Other				14.876***
				(5.105)
Male*Economics				-0.819
				(0.673)
Male*Education				-1.368**

(continued on next page)

	(1)	(2)	(3)	(4)
VARIABLES	YOSSEA-Min	YOSSEA-Min	YOSSEA-Min	YOSSEA-Min
				(0.677)
Male*Humanities				0.687
				(1.841)
Male*Journalism and Information				0.856
				(5.078)
Male*Law				-1.159
				(0.790)
Male*Personal services				-0.992
				(1.119)
Male*Public Administration				-0.453
				(0.694)
Male*Social and behavioural sciences				-7.295**
				(3.157)
Male*Other				-4.740
				(7.267)
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.
Standard errors are clustered at the high school city level

Table A.7: Gender Gap in OSS Score:: Turkish-Math Track

	(1)	(2)	(3)	(4)	(5)
VARIABLES	OSS-EA	OSS-EA	OSS-EA	OSS-EA	OSS-EA
Male	-21.181***	-24.054***	-27.001***	3.652***	0.304
	(6.454)	(6.564)	(7.451)	(0.207)	(0.233)
AOBP	-0.738***	-0.819***	-0.468***	0.058	
	(0.153)	(0.157)	(0.176)	(0.124)	
$AOBP^2$	0.011***	0.012***	0.010***	0.006***	
	(0.001)	(0.001)	(0.001)	(0.001)	
Male*AOBP	0.712***	0.809***	1.009***		
	(0.218)	(0.221)	(0.255)		
Male*$AOBP^2$	-0.005***	-0.006***	-0.008***		
	(0.002)	(0.002)	(0.002)		
Income(base: Less than 250 TL)					
250-500 TL	-0.159	0.318*	0.888***	0.903***	0.218
	(0.187)	(0.187)	(0.216)	(0.217)	(0.250)
More than 500 TL	-0.212	0.821***	3.826***	3.827***	2.975***
	(0.267)	(0.265)	(0.307)	(0.308)	(0.361)
Prep Expenditure(base: Missing)					
No prep school	-0.386				
	(0.263)				
Low	3.607***				
	(0.251)				
Medium	4.524***				
	(0.299)				
High	4.982***				
	(0.442)				
Scholarship	3.827***				
	(0.658)				
Parental Education					
Literate	-0.823*	-1.172***	-3.193***	-3.201***	-1.936***
	(0.430)	(0.441)	(0.487)	(0.488)	(0.564)
Primary School	0.076	-0.294	-1.778***	-1.796***	-1.318***
	(0.324)	(0.332)	(0.387)	(0.388)	(0.464)
Middle or High School	0.199	0.154	0.298	0.280	0.451
	(0.309)	(0.313)	(0.374)	(0.375)	(0.456)
College/Master/PhD	0.580	0.796**	4.121***	4.112***	5.374***
	(0.354)	(0.360)	(0.461)	(0.463)	(0.566)
Observations	12,664	12,664	12,664	12,664	12,664
School FE	YES	YES	NO	NO	NO
Province FE	NO	NO	YES	YES	YES

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

Table A.8: Gender Gap in OSS Score:: Social Science

	(1)	(2)	(3)	(4)	(5)
VARIABLES	OSS-SOZ	OSS-SOZ	OSS-SOZ	OSS-SOZ	OSS-SOZ
Male	0.524	-0.585	-1.351	3.197***	1.592***
	(14.905)	(16.296)	(13.038)	(0.326)	(0.351)
AOBP	0.322	0.239	0.721**	0.869***	
	(0.392)	(0.427)	(0.345)	(0.217)	
$AOBP^2$	0.003	0.003	-0.001	-0.002	
	(0.003)	(0.004)	(0.003)	(0.002)	
Male*AOBP	-0.001	0.044	0.058		
	(0.518)	(0.566)	(0.455)		
Male*$AOBP^2$	0.001	0.001	0.000		
	(0.004)	(0.005)	(0.004)		
Income(base: Less than 250 TL)					
250-500 TL	-0.047	0.520	0.466	0.443	-0.078
	(0.393)	(0.398)	(0.336)	(0.336)	(0.371)
More than 500 TL	-0.353	0.840	2.155***	2.139***	1.047*
	(0.584)	(0.589)	(0.494)	(0.494)	(0.536)
Prep Expenditure(base: Missing)					
No prep school	-0.299				
	(0.449)				
Low	4.281***				
	(0.455)				
Medium	5.406***				
	(0.662)				
High	5.931***				
	(1.340)				
Scholarship	6.283***				
	(1.529)				
Parental Education					
Literate	-0.702	-1.062	-2.446***	-2.436***	-1.593*
	(0.927)	(0.985)	(0.844)	(0.844)	(0.930)
Primary School	0.449	0.075	-1.059	-1.074	-0.732
	(0.688)	(0.743)	(0.665)	(0.663)	(0.744)
Middle or High School	0.671	0.534	0.502	0.483	0.397
	(0.701)	(0.760)	(0.679)	(0.679)	(0.756)
College/Master/PhD	0.773	1.387	3.484***	3.436***	3.758***
	(0.875)	(0.926)	(0.871)	(0.869)	(0.992)
Observations	4,765	4,765	4,765	4,765	4,765
School FE	YES	YES	NO	NO	NO
Province FE	NO	NO	YES	YES	YES

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

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

Placement score: SAY	100.82	-5.16	-6.69	-6.76	-1.58	-2.41	33.18	-3.95
Placement score: EA	77.52	3.41	-1.87	-6.41	2.07	-3.92	0.00	1.08
Major: Agriculture	4.98	-0.84	-1.16	-1.91	-4.55	-5.77	-6.22	-7.85
Major: Architecture and construction	6.91	-0.95	1.01	1.74	-5.29	0.11	-5.32	-3.67
Major: Business and Administration	-4.28	3.62	2.73	-0.69	-0.97	4.40	-0.55	-6.45
Major: Economics	-3.29	3.70	2.29	4.01	-1.66	4.32	0.00	-7.17
Major: Engineering	-1.24	2.85	1.14	1.78	-4.45	0.27	-5.32	-5.25
Major: Health Service	6.46	-0.60	0.56	1.53	-5.11	-5.52	1.75	-2.87
Major: Mathematics and Statistics	3.79	4.23	-1.77	1.71	0.16	-1.27	-5.04	-6.43
Major: Medicine	5.18	4.76	-2.68	3.54	2.33	-3.49	2.90	2.52
Major: Science	2.57	-0.55	-4.26	2.23	-0.03	-0.88	-1.30	-5.62
Major: Other	0.92	1.19	1.53	-0.12	-7.29	0.32	-4.81	-10.76
Outside: SAY	72.72	65.26	47.86	10.02	-29.08	5.16	-25.27	16.48
Outside: EA	20.75	-38.76	-53.97	-25.41	31.42	28.58	125.51	-17.08
Outside: SOZ	49.10	-39.04	4.26	15.42	-5.67	-36.08	-30.66	0.67
Outside: HGPANORM	-10.27	73.27	47.06	43.32	32.97	12.84	-33.97	-3.73
Type share	0.03	0.04	0.15	0.23	0.09	0.08	0.12	0.26

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

Placement score: SAY	73.72	-4.52	-5.13	5.81	1.00	-4.70	-4.63	-1.59
Placement score: EA	73.92	-9.16	-4.34	9.36	-3.95	-1.71	-0.14	2.83
Major: Agriculture	2.13	1.86	0.39	-1.01	-2.24	-2.92	-5.17	-7.00
Major: Architecture and construction	-0.34	4.48	-3.81	0.71	-3.62	-3.83	-1.63	-4.56
Major: Business and Administration	2.38	-0.36	2.92	3.25	-0.07	4.61	0.01	-5.98
Major: Economics	0.17	-0.01	3.03	2.56	-0.01	4.52	-0.60	-6.56
Major: Engineering	-0.33	3.52	2.14	2.57	1.11	0.16	0.22	-4.30
Major: Health Service	-3.49	-3.62	-4.49	-2.38	2.54	-3.66	-2.87	-3.95
Major: Mathematics and Statistics	-3.02	-2.60	-1.56	2.06	-0.71	-1.75	1.89	-3.70
Major: Medicine	17.42	-1.55	1.01	2.07	3.60	0.58	4.05	1.95
Major: Science	-1.22	0.48	-3.48	0.63	3.51	-1.31	0.40	-4.75
Major: Technical Science	-1.94	-2.53	-4.07	2.20	-1.95	-2.62	-1.29	-5.99
Major: Technical Services	-0.23	-0.85	0.54	-2.09	-1.74	-2.84	3.46	-3.43
Major: Veterinary	6.94	1.66	2.05	4.73	4.27	-1.70	-3.48	-1.38
Major: Other	3.14	-1.72	-3.26	3.15	-0.84	-3.02	-6.01	-7.50
Outside: SAY	69.65	8.62	-2.57	12.75	-31.46	69.22	3.56	9.76
Outside: EA	45.36	-18.50	6.27	7.23	47.65	-89.88	-3.83	-3.27
Outside: SOZ	-22.86	14.02	-5.98	-0.75	-6.98	18.75	1.28	1.69
Outside: HGPANORM	131.77	-24.37	32.13	41.24	30.27	22.18	-13.90	-25.38
Type share	0.01	0.05	0.25	0.32	0.03	0.10	0.08	0.16

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

Placement score: EA	-3.57	0.97	-0.04	-8.92	6.65	0.47	4.30	3.75
Placement score: SOZ	-18.71	-4.22	-9.97	-9.82	0.67	-6.85	-0.50	-5.73
Major: Arts	4.31	3.60	3.53	-0.02	-0.19	-1.46	-2.46	-5.81
Major: Business and Administration	1.53	-3.81	-2.66	-0.16	2.90	-2.97	-6.35	-7.96
Major: Economics	1.71	-1.19	-5.86	0.02	2.12	-2.93	-6.45	-6.64
Major: Humanities	-3.92	-0.79	-2.71	3.30	-5.97	-4.76	-10.31	-9.96
Major: Journalism and Information	-5.16	-5.64	-2.54	6.57	-0.76	-7.59	-10.63	-10.91
Major: Language and Literature	-0.20	2.57	-4.20	0.26	-0.14	-5.60	-3.37	1.36
Major: Law	2.57	1.92	-0.27	11.14	-0.82	3.89	-6.16	-6.30
Major: Personal services	0.90	-5.41	-5.03	-0.10	0.40	-3.08	-6.87	-6.96
Major: Public Administration	1.66	0.12	-0.58	10.99	3.17	-2.68	-5.25	-9.18
Major: Social and behavioural sciences	-0.23	-2.28	0.24	-0.04	-3.11	-5.22	-8.12	-7.11
Major: Other	-7.61	-4.94	-1.75	-5.14	3.25	-6.38	-2.91	-3.98
Outside: SAY	-42.10	65.99	-17.90	1.52	55.23	11.74	22.31	99.57
Outside: EA	-36.55	-59.94	34.57	14.41	-11.58	3.34	-28.82	-82.21
Outside: SOZ	75.76	3.62	-12.44	-10.43	-33.52	-8.54	12.95	7.51
Outside: HGPANORM	17.24	10.51	19.94	13.90	41.27	18.81	21.67	-34.85
Type share	0.14	0.03	0.12	0.05	0.04	0.27	0.19	0.16

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

Placement score: EA	23.00	-2.52	-4.29	13.37	1.84	3.08	0.74	0.18
Placement score: SOZ	-0.00	-12.55	-6.47	15.24	-11.44	-1.13	-1.27	-3.41
Major: Business and Administration	4.21	3.95	2.46	0.38	-2.11	-4.30	-5.40	-6.01
Major: Economics	3.66	4.17	1.83	-0.37	-4.48	0.69	-5.36	-7.20
Major: Journalism and Information	-1.74	-3.37	-2.87	7.52	-4.18	-1.38	-8.41	-7.67
Major: Language and Literature	-0.00	-0.08	-1.46	-1.96	-0.03	0.59	-1.59	-4.23
Major: Law	4.42	3.23	4.35	-0.03	-0.84	2.13	-4.88	3.06
Major: Personal services	6.69	-3.08	-2.98	6.44	-6.84	-5.57	-7.75	-3.97
Major: Public Administration	3.28	1.89	2.95	5.49	-4.03	3.01	-4.39	-3.15
Major: Social and behavioural sciences	-1.94	1.43	-3.53	7.92	-2.79	-5.22	-7.99	-4.03
Major: Other	-0.49	-5.36	-7.30	-1.48	-1.18	0.79	-5.44	-7.31
Outside: SAY	54.80	5.39	29.85	-98.81	1.06	35.47	13.33	17.91
Outside: EA	35.35	-15.97	-36.08	-2.05	-5.69	-17.26	11.35	-8.83
Outside: SOZ	-64.54	19.34	4.52	105.74	10.85	-2.95	-11.60	5.92
Outside: HGPANORM	124.88	-0.16	65.74	100.63	21.21	7.86	-14.52	-20.70
Type share	0.02	0.11	0.20	0.01	0.10	0.07	0.22	0.29

Table A.13: Estimated demand parameters, type-specific coefficients and type shares, Social Science track, female

Placement score: EA	1.64	-3.69	-8.95	-0.20	-0.00	-99.12	-12.09	36.91
Placement score: SOZ	5.90	-9.53	-6.68	-5.90	0.00	-100.90	-3.19	42.17
Major: Arts	7.23	5.24	1.59	0.27	0.00	-2.73	-4.82	-6.49
Major: Business and Administration	-1.58	-1.78	-1.50	-7.46	0.00	-3.39	-1.53	4.11
Major: Humanities	-2.58	-1.33	-6.76	-5.97	0.00	-0.26	-2.40	-9.27
Major: Journalism and Information	3.54	-1.54	1.73	-5.78	0.00	-0.33	-7.29	-8.99
Major: Language and Literature	1.37	-1.04	-5.98	-5.12	-0.00	1.85	-6.74	-9.29
Major: Public Administration	-3.92	-1.08	6.04	-5.74	-0.00	-2.02	-0.05	4.29
Major: Other	-0.57	3.22	-5.25	-8.66	0.00	0.30	-6.49	-8.23
Outside: SAY	-3.71	59.88	32.18	56.90	-0.00	-10.77	18.32	85.18
Outside: EA	-3.91	29.37	-79.45	-71.94	-0.00	-54.36	12.08	-99.36
Outside: SOZ	-4.15	-71.93	40.55	21.98	-0.00	-64.26	-25.44	63.68
Outside: HGPANORM	-0.69	2.96	38.33	3.60	-0.00	-52.83	9.92	100.77
Type share	0.01	0.02	0.32	0.17	0.00	0.20	0.25	0.04

Table A.14: Estimated demand parameters, type-specific coefficients and type shares, Social Science track, male

Placement score: EA	-2.60	-8.43	29.49	-4.37	82.65	-8.59	-0.92
Placement score: SOZ	-1.47	12.44	26.37	7.17	95.25	4.37	-3.38
Major: Arts	2.39	-0.55	-0.56	-1.15	-1.35	-4.73	-5.45
Major: Business and Administration	-3.55	-0.01	3.92	-0.01	3.76	-0.07	-7.56
Major: Humanities	-4.50	-3.67	3.22	2.71	-8.64	-2.20	-5.05
Major: Journalism and Information	2.91	-6.92	-1.07	1.96	-5.11	-5.92	-4.51
Major: Public Administration	2.02	-0.02	3.62	-0.00	4.66	0.01	-1.78
Major: Other	-5.07	-7.17	3.41	0.84	-7.76	-2.05	-4.94
Outside: SAY	50.25	-2.21	-52.48	-16.97	109.12	96.27	24.18
Outside: EA	-7.96	-2.57	-15.82	-9.54	32.47	8.05	37.76
Outside: SOZ	-37.33	-3.20	81.23	36.72	1.33	-76.38	-47.42
Outside: HGPANORM	40.82	-1.49	104.84	-5.44	106.93	-28.24	-22.08
Type share	0.19	0.00	0.04	0.09	0.11	0.16	0.41

Table A.15: 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.14: Gender Differences in Major Choice (Science Track)

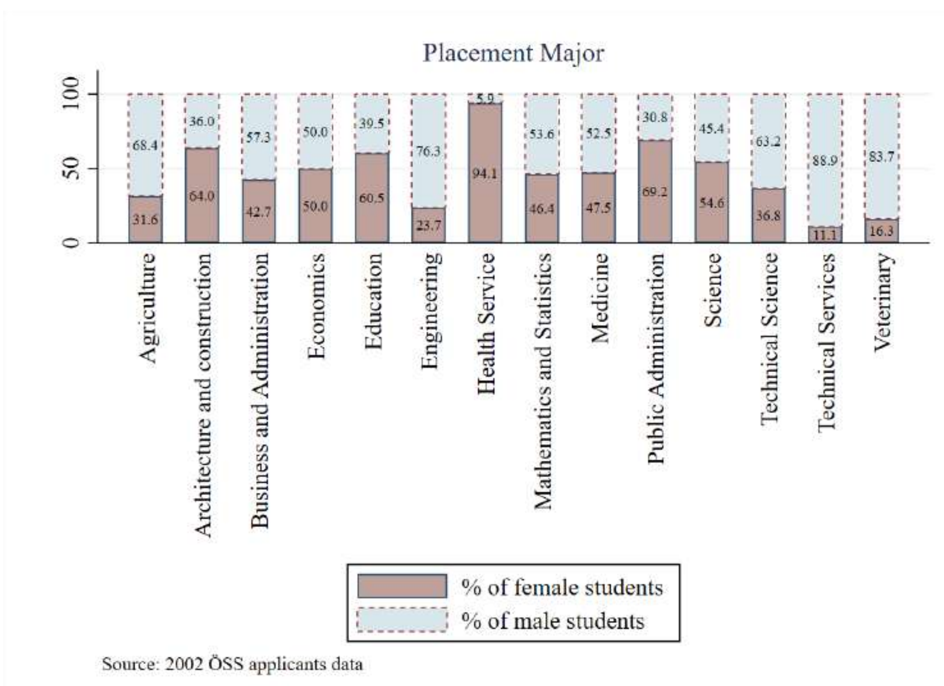


Figure A.15: Gender Differences in Major Choice (Turkish-Math Track)

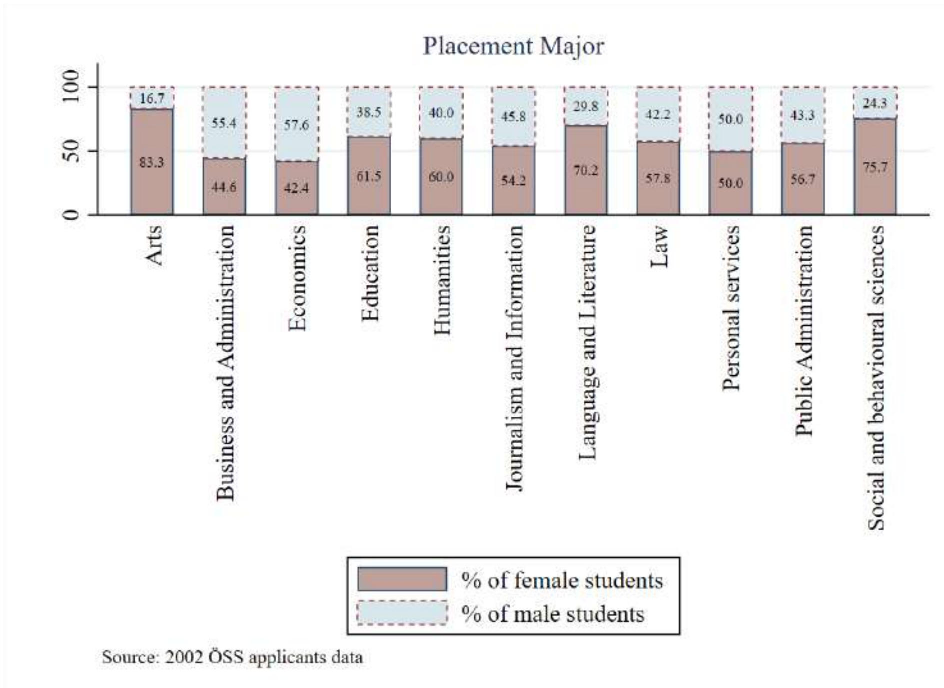


Figure A.16: Gender Differences in Major Choice (Social Science Track)

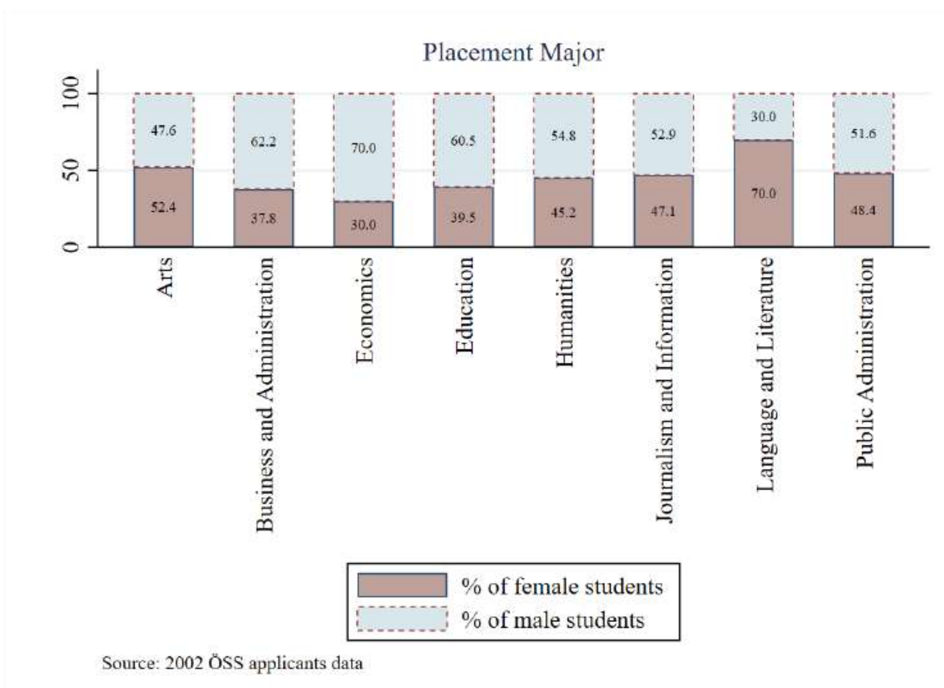


Figure A.17: Preferred Majors (Science Track)

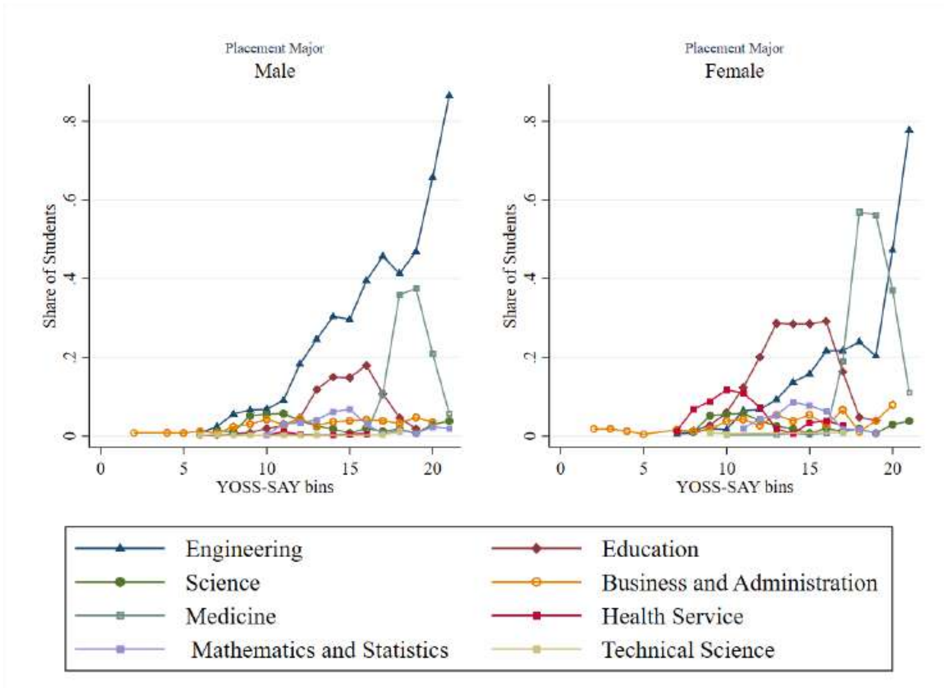


Figure A.18: Preferred Majors (Turkish-Math Track)

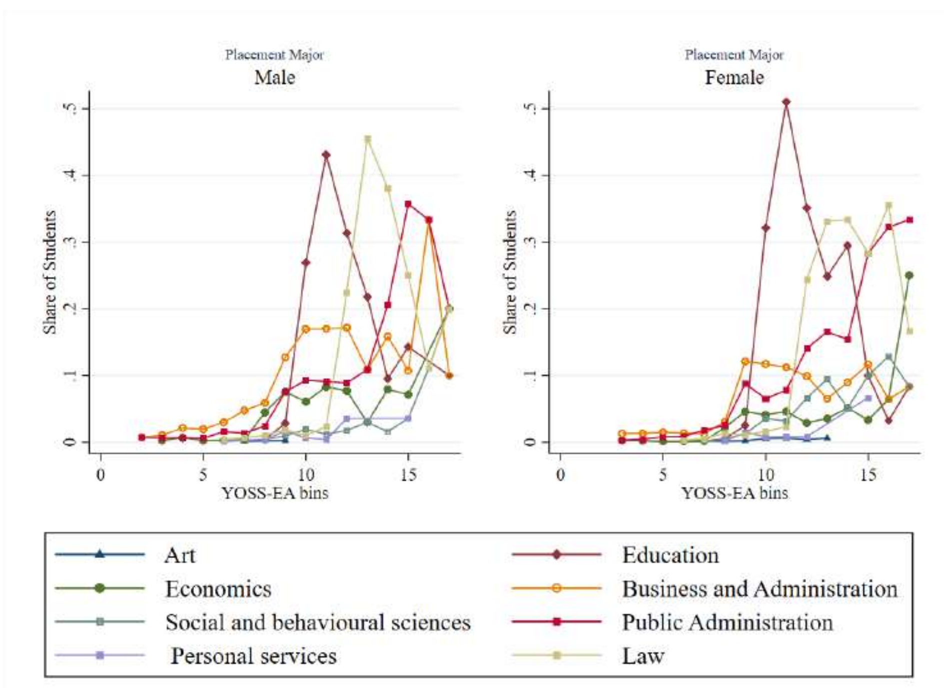


Figure A.19: Preferred Majors (Social Science Track)

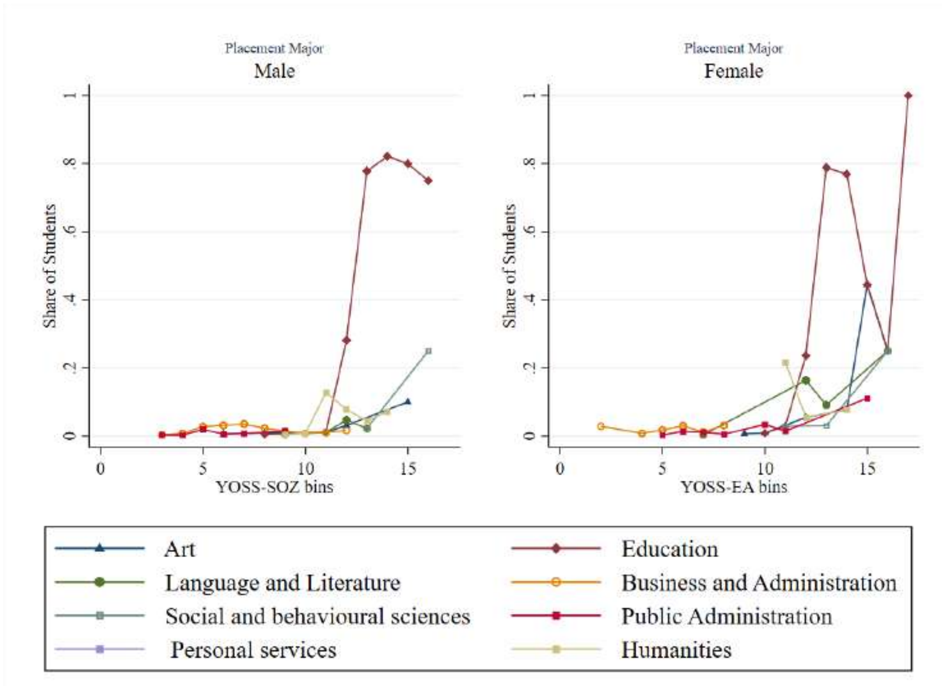


Figure A.20: 1st Preference Major(Turkish-Math Track)

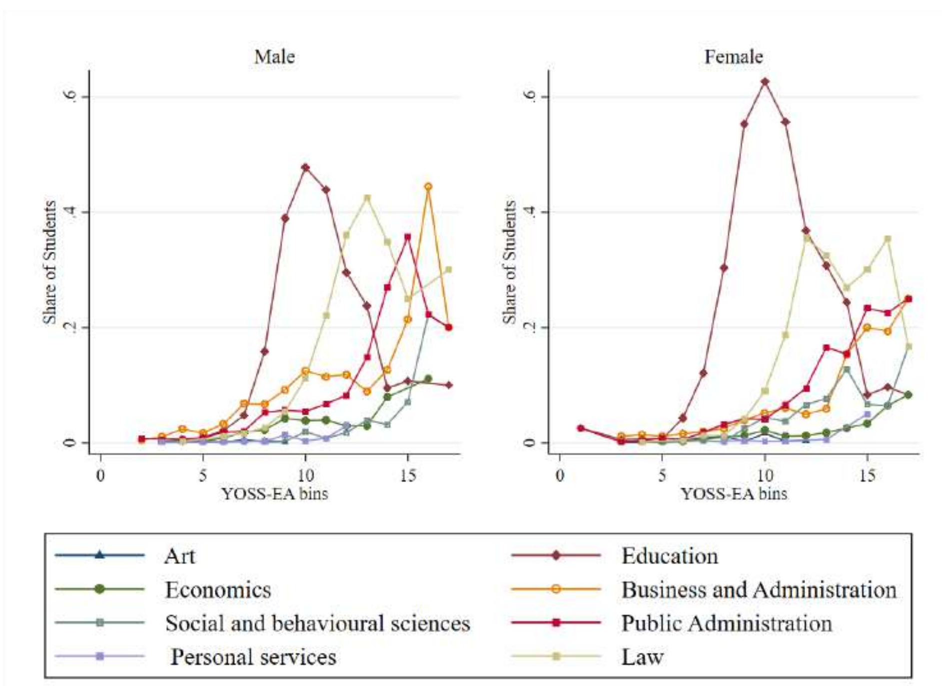


Figure A.21: 1st Preference Major (Social Science Track)

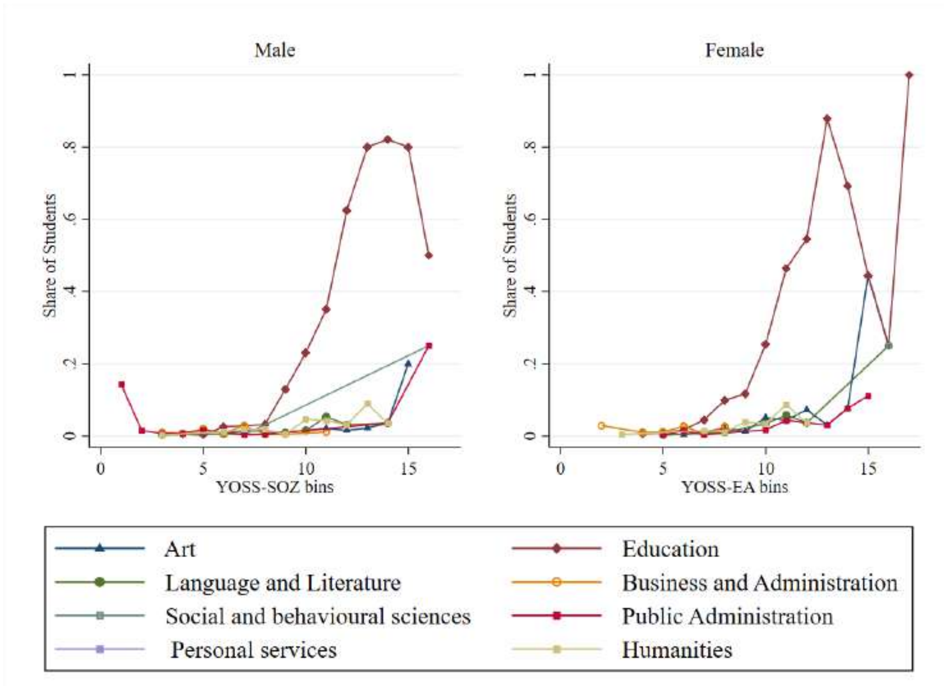


Figure A.22: Difference between Y-OSS Score and Minimum Cutoff of the 1^{st} Ranked Department in 2001 by Gender

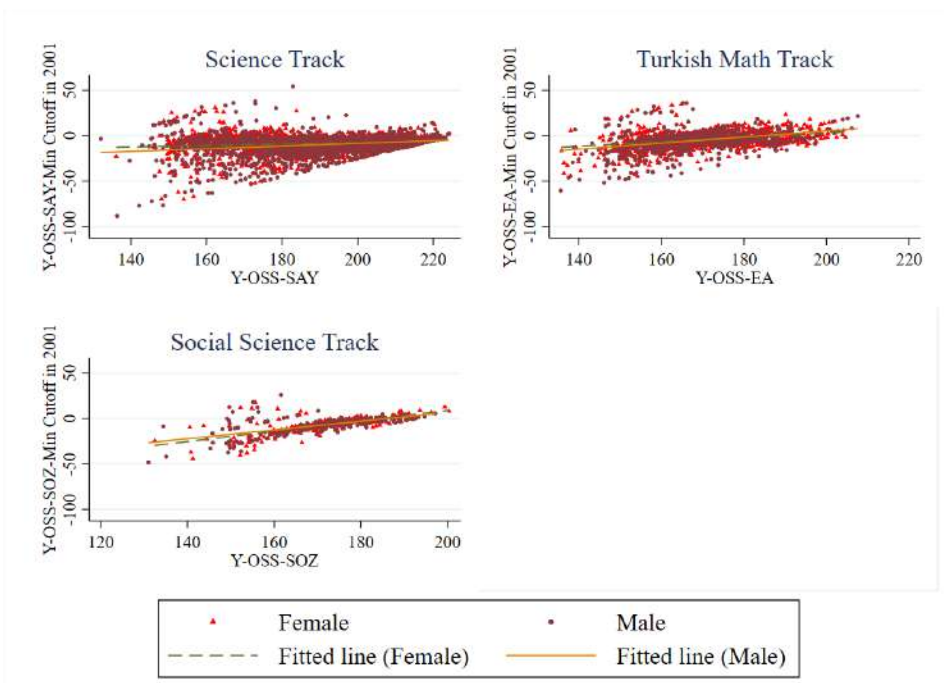


Figure A.23: Difference between Y-OSS Score and Minimum Cutoff of the Placement Department in 2001 w.r.t. Y-OSS score (Turkish-Math Track)

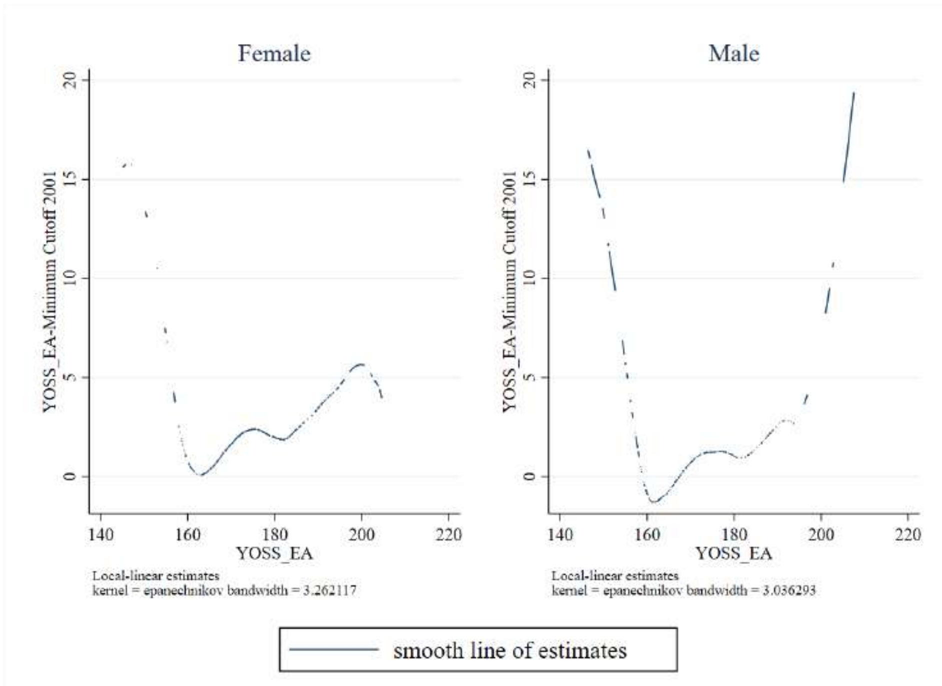


Figure A.24: Effect of being Male on OSS score (Science Track)

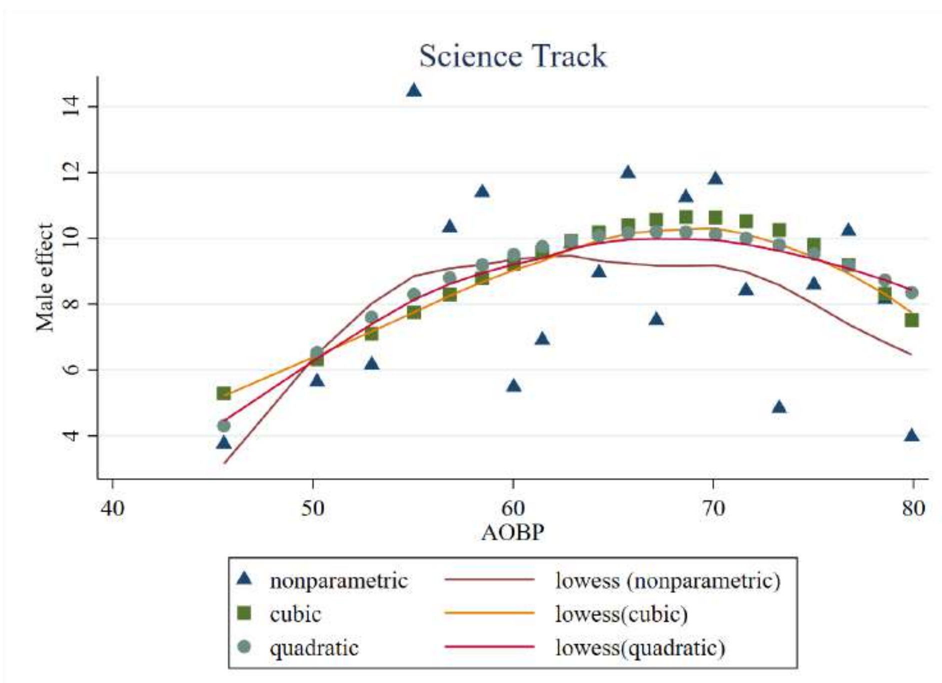
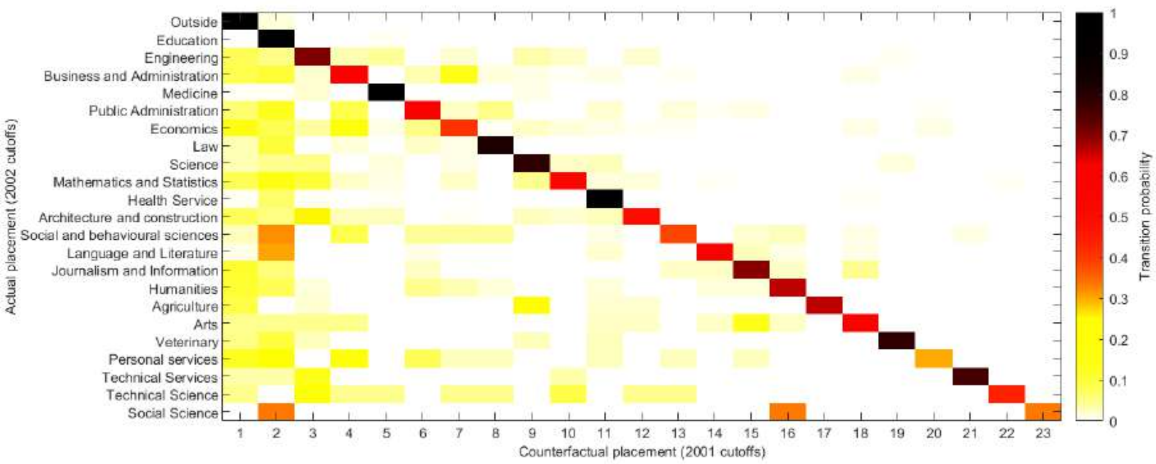
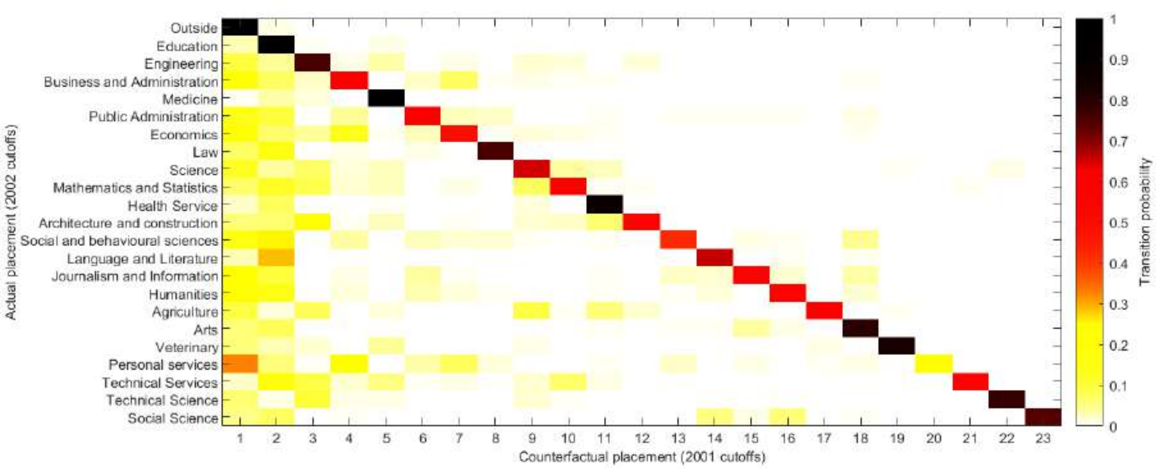


Figure A.25: Transition matrix for majors of placement, predicted using the preference data



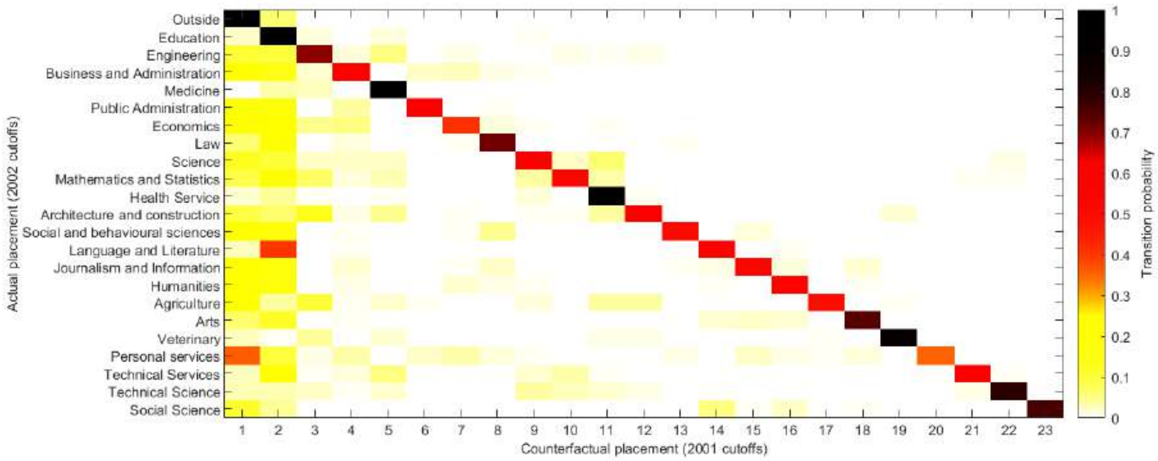
Notes: Actual major — major of placement in 2002. Counterfactual major — major of placement if the admission cutoffs are the same as in 2001. “Outside” corresponds to not being placed. Counterfactual majors follow the same order as the actual ones (e.g., the label 3 corresponds to Engineering). The value 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.26: Transition matrix for majors of placement, predicted using the estimated model



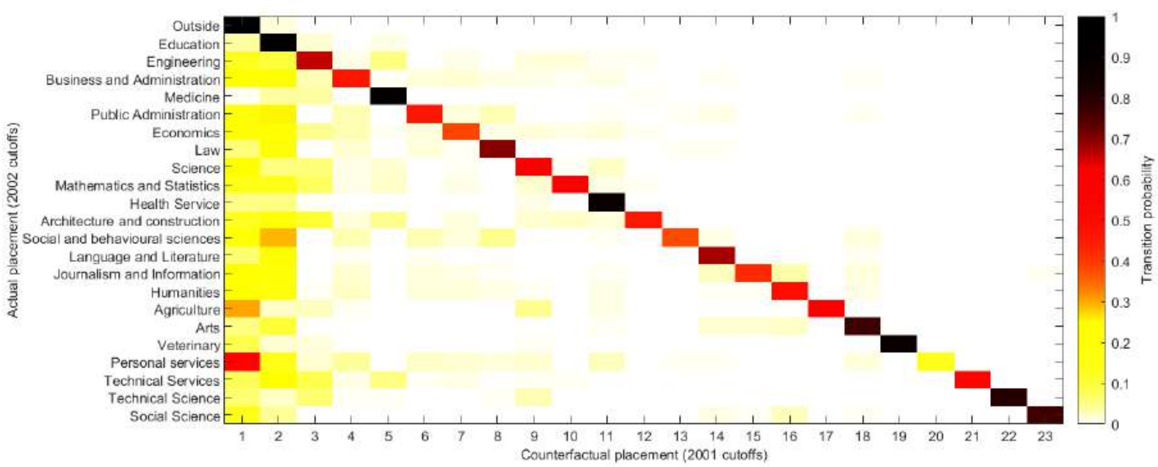
Notes: Actual major — major of placement in 2002. Counterfactual major — major of placement if the admission cutoffs are the same as in 2001. “Outside” corresponds to not being placed. Counterfactual majors follow the same order as the actual ones (e.g., the label 3 corresponds to Engineering). The value 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.27: Transition matrix for majors of placement, using latent class logit and ex-post stability in 2002 (alternative specification 1)



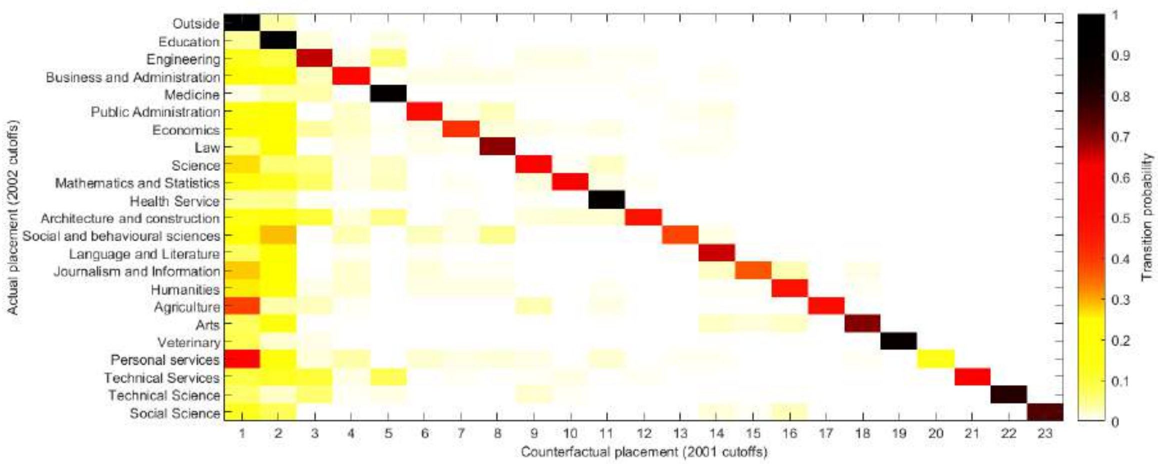
Notes: Actual major — major of placement in 2002. Counterfactual major — major of placement if the admission cutoffs are the same as in 2001. "Outside" corresponds to not being placed. Counterfactual majors follow the same order as the actual ones (e.g., the label 3 corresponds to Engineering). The value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

Figure A.28: 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.29: 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.

B Deriving the Likelihood Function

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

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

$$\mathcal{L}_i(\theta; j_{i1}, j_{i2}, L_i, C_{i1}, C_{i2}) = \sum_{t=1}^{T} \sigma_t \mathcal{L}_{it}(\theta; j_{i1}, j_{i2}, L_i, C_{i1}, C_{i2}, t) \tag{2}$$

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:

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

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

$$\times \prod_{k \in A_1 \setminus j_1} \exp(-\exp(-\varepsilon_{j_1} - \delta_{j_1} + \delta_k)) \exp(-\varepsilon_{j_1} - \exp(-\varepsilon_{j_1})) d\varepsilon_{j_1}$$

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

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

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

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

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

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

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

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

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

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

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

Case 3: $j_1 = j_2$

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

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

C Estimation Details

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

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

1. The highway distance between student's high school and the program's campus.[34] A dummy for the high school and the campus being in the same province.

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

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

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

4. Interactions of net tuition, 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.

D 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

E 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 (3). Students keep their AOBP over attempts made.

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

$$+ [(OBP_{ij} \times \frac{\ddot{O}SS_{jk}}{80}) - (\frac{\ddot{O}SS_{jk}}{80} \times \min_{i \in j}(OBP_i))]$$
$$\times \left[\frac{80 - [(\frac{\ddot{O}SS}{80} \times \min_{i \in j}(OBP_{ij})) - (\frac{\ddot{O}SS - 80}{10})]}{(\frac{\ddot{O}SS}{80} \times \max_{i \in j}(OBP_{ij})) - (\frac{\ddot{O}SS}{80} \times \min_{i \in j}(OBP_{ij}))} \right] \tag{3}$$

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[35] 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:

$$OBP_{ij} = 10 \frac{gpa_i - \mu_{gpa,j}}{\sigma_{gpa,j}} + 50 \tag{4}$$

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 (4) 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

[35]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 (4) 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.