The Gender Gap in College Major Choice. The case of Chile.*

Paola Bordón[1], Catalina Canals[2] and Alejandra Mizala[3]

[1]Universidad Alberto Hurtado and COES.

[2]CIAE, Institute of Education, Universidad de Chile.

[3]CIAE, Institute of Education and Department of Industrial Engineering, Universidad de Chile.

August 21, 2019

Abstract

This paper studies the gender differences in college applications in Chile. We developed a model for college-major application based on the Chilean institutional setting. We estimate a nested logit model to predict the first-preference college major choice of a random sample of applicants to the Chilean centralized college admission system. Using counterfactual exercises, we show that there would be a greater reduction in the gender gap in program applications if males applicants adopted female college major preference parameters than if female applicants adopted male college major preference parameters. Additionally, we find that males apply to selective programs even when they are marginal candidates while equally qualified women will tend not to apply to those same selective programs. These findings can be understood as a crowding out effect caused by prevailing male preferences being an important contributor to the gender gap in STEM fields, keeping out equally talented women from those majors.

JEL Classification: I20, I23, J16.

Keywords: Gender gap, gender preferences, college-major choice, crowding out, nested logit.

*Corresponding author Alejandra Mizala, Periodista José Carrasco Tapia Nº 75, Santiago, Chile, phone: +562 2978 4034, email: amizala@uchile.cl. Paola Bordón, Department of Economics, Erasmo Escala 1835, Santiago, Chile, email: pbordon@ahurtado.cl. Catalina Canals, email: catalina.canals@ciae.uchile.cl.

1 Introduction

There has been much focus on increasing female enrollment in science, technology, engineering and math (STEM) fields due to the empirical evidence showing female under-representation in STEM and it's contributing effect to the gender gap on earnings (Arcidiacono, 2004). STEM fields are associated with more valuable job-related human capital leading to higher monetary returns. While much of the attention has been placed on increasing female participation in STEM fields we see in our data more potential to address the gender gap by increasing males' willingness to consider non-STEM fields, diminishing the crowding out effect observed in college major applicants' decisions.

In this paper we examine what would happen if females pursued male preference strategies in college applications and vice versa, an under-explored topic in the literature. This exercise sheds light on the drivers of the gender gap in college major applications.

We are able to use the revealed preferences of students for college-major choices by taking advantage of Chile's Centralized Admission System (SUA) from which we have data on the entire cohort of applicants to enter Chilean universities in 2015. In the application process students must first take a series of admission tests and, when they receive their final scores must submit a ranked preference list of up to 10 college-major programs they wish to be considered for, based on their high school and test performance.

In order to estimate a college-major choice model we rely on a nested logit methodology because the university preference is correlated to the major preference. In fact, in Chile enrollment is universally a paired college-major program, unlike admissions in most universities in the United States where enrollment to a university allows students to change their majors throughout their college career.

Our results from the estimation are in line with what has been found in the literature with females applying exceptionally more to non-physician health majors and males applying exceptionally more to engineering and technology majors. These results also show that proficiency in math measured through both GPA and test scores has a large effect on both male and female preferences to apply to engineering majors.

Additionally, we find there are gender differences in the probabilities of applying to college majors that can be explained if the parents work on occupations related to those majors, and in the gender composition of high school classes.

We also run exercises to create different students profiles. First, we take the strategy of averaging all observable characteristics, maintaining only male and female differences to estimate the ceteris paribus effect of gender on the probability of application to each major. These results show that an average male applicant is much more likely to apply to engineering and technology majors than a female with the same average characteristics, and an average female applicant is much more likely to apply to non-physician health majors than a male with average characteristics. This strategy isolates the unobservable preferences in college-major choices that are tied specifically to gender.

Secondly, we introduce proficiency distributions tied to the probability of applying to the most selective majors and universities for males and females, maintaining the average across all other observable characteristics. We find that high achievement students (males and females) have a higher probability of application to the most selective areas (medicine & odontology, civil engineering and law) and the most selective universities. However, when we compare average males and females with the same academic results, this tendency is stronger for males, showing that males are more likely than females to bet on the most selective majors considering their subject grades in high school and test results. That is, we find that males apply to competitive programs even when they are marginal candidates while equally qualified women will tend not to apply to those same competitive programs. Thus, the gender gaps on college major choices are based on differentiated behaviors of males and females, particularly of average students. These findings reveal a crowding out effect caused by prevailing male preferences, which in turn contributes significantly to the gender gap in STEM fields, keep out equally talented women from those majors.

Based on these findings, we use the structural estimated model to construct a series of counterfactuals that shed light on the most effective ways to reduce the gender gap in college major choices. Our estimates contribute to the existing literature by showing that a greater reduction in the gender gap in major choice would occur if more males adopt female preference parameters than if more females adopt male preference parameters in female dominated areas. Until now, most of the discussion on gender gap advocates for increasing female participation in STEM fields, whereas little attention has been put on increasing male participation on health or education. Our results suggest that a larger decrease in the gender gap in college major choice can be accomplished increasing the number of males in female dominated areas instead of the other way around.

This paper is organized as follows. Section 2 reviews the literature, section 3 presents a description of the Chilean higher education system and the application process, section 4 describes the data, section 5 presents the model and the estimation procedure used in this paper, section 6 provides the estimation results, section 7 shows some profiles or different combinations of simulated choice probabilities for an average student differentiated by gender, section 8 provides the counterfactuals exercises, section 9 analyzes the gender differences in light of our findings, and section 10 concludes.

2 Review of the Literature

Over the past decades many countries have made relevant progress in narrowing or closing gender gaps in years of schooling, and secondary and post-secondary school attendance. However, the opposite is true for the content of education received; empirical evidence shows that in many countries women are under-represented in the fields of mathematics engineering and science, whereas they are over-represented in humanities, languages, education and the arts. In 2012, only 14 percent of women who entered for the first time to the university in the OECD countries chose science-related fields of study vis à vis 39 percent of men (OECD, 2015). Even more, there is gender segregation in science; Sikora and Pokropek (2012a) shows -using data from PISA 2006-, that in all the 50 countries included in the test science-oriented girls prefer biology, agriculture or health careers, whereas boys favor careers in computing, engineering or mathematics.

The proportion of female and male students across fields of study with over-representation of one gender over the other has been called horizontal sex segregation in education. This phenomenon can explain gender segregation in the labor market, because the content of schooling accounts for a substantial part of the gender gap in jobs and earnings, Arcidiacono (2004) estimates a dynamic discrete choice model in which he finds positive returns to STEM. In general, stereotypically male subjects create more valuable job-related human capital and generate a higher monetary return.

Therefore, it is relevant to understand why girls do not choose the most rewarding majors in terms of future wage and labor market opportunities. The literature on this issue has not yet come to definitive answers.

The basic question is if these differences are explained by nature or nurture. Research in Psychology and Medicine has explained gender segregation by the presence of biological and neurological gender differences. According to this approach, boys use more cortical areas dedicated to spatial and mechanical functioning (Kimura, 2000). On the contrary, girls develop more the part of the brain devoted to verbal and emotional functioning. For this reason, girls may underperform relatively in technical and quantitative subjects from childhood and gradually disengage from these subjects (Killgore and Yurgelun-Todd, 2004; Lenroot, et al. 2007).

However, in some of the top-performing countries and economies in PISA girls perform equal than their male classmates in mathematics and attain higher scores than all boys in most other countries around the world, this is for instance the case of Hong Kong, Shanghai, Singapore and Taipei-China. These results suggest that the gender gap in mathematics is not determined by innate differences in ability. Furthermore, Favara (2012) finds that the belief that men are naturally more skilled at technical/quantitative domains is empirically unfounded and attainments, such us performance and grades, are not able to explain alone subject choices; Turner and Bowen (1999) and Dickson (2010) find that SAT scores play a small role in major gap, also, Justman and Méndez (2018) find that female students require stronger prior signals of mathematical ability to choose male-dominated subjects. Thus, girls and boys performing equally in the same subjects choose differently and according to their own gender stereotype; nonetheless, girls and boys at the top of the grade distribution behave similarly in terms of subject choice. However, there is no agreement on the role of pre-college factors, Speer (2017) using a broader array of pre-college test scores (the ASVAB), shows that differences in college preparation can actually account for a large portion of most gender gaps in college major content. Delaney and Devereux (2019) use a preference ranking for all secondary school students who apply for college in Ireland. They find that, of the 22 percentage points raw gender gap in STEM, about 13 percentage points is explained by differential subject choices and grades in secondary school. They also find that subject choices are more important than grades.

Other literature associates career choices with gender stereotypes.[1] The argument is that gender specific attributes such as risk aversion, self-confidence and attitude towards competition contribute to sketch the gender identity. These attributes develop during childhood and affect boys' and girls' choices along their lifecycles (Sutter and Rützler, 2010; Gneezy and Rustichini, 2004).

The literature shows that women are generally more risk averse than men and more likely to shy away from competition. Some authors suggest that these characteristics are related to a gender gap in self-confidence (Booth and Nolen, 2011 and 2012; Croson and Gneezy, 2009; Gneezy et al., 2003; Niederle and Vesterlund, 2007 and 2010; Datta Gupta et al., 2005). Niederle and Yestrumskas (2008) find that females are less likely to seek challenges than men with the same abilities and this is because they have a higher risk aversion or higher uncertainty about their own abilities than men do. Saltiel (2019) focuses on the role of mathematical self-efficacy, which measures an individual's perceived ability to perform math-related tasks, in explaining gender gaps in math-intensive majors. He finds that math problem solving ability and self-efficacy are strong predictors of STEM enrollment for both men and women; he also finds that the correlation between them is higher for men than for women, indicating a relative shortfall of high-achieving women who are confident in their math ability. Kutz-Costes et al (2008) also suggests that girls' perception of their own mathematics and sciences abilities is lower than boys´ perception. Thus, if advanced math courses are relatively challenging, this explains why females often opt out of those courses.

Gneezy et al (2003) presents experimental evidence supporting that women may be less effective than men in mixed-sex competitive environment, although they are able to perform similarly in non-competitive environment and better in single-sex environment. Niederle and Vesterlund (2007) argue that lower self-confidence and less taste for competition play a substantial role in mathematics – and more so at the right tail of distribution. Örs, Palomino and Peyrache (2013) and Jurajda and Münich (2011), find that females underperform in high-stake tests relative to males with similar abilities, and Niederle and Vesterlund (2010), conclude that the gap at the top of the math test score distribution does not necessarily imply a gap in the underlying math abilities. One reason is that girls have particularly little faith in their own math abilities – conditional on actual abilities – due to extensive gender-stereotyping in math working. Another reason is that math tends to be a very competitive discipline because the answers to exercises are either right or wrong.

[1] Another possibility is that gender discrimination in the labor market generates sex-differences in subject choice. Female students anticipate potential gender discrimination in the labor market avoiding those majors which offer higher rewards for men than for women.

Alternatively, experimental evidence suggests that the gender gap in the choice of the college major is mainly due to differences in non-pecuniary preferences and tastes, pointing out the role of preferences and social interactions in explaining how individuals make decisions; several papers highlight the importance of preferences in driving STEM gaps (Bartolj and Polanec, 2012; Zafar, 2013; Wiswall and Zafar, 2014, 2017). The gender specific attributes discussed above might explain why boys and girls have different educational preferences. Differences in attitudes and preferences might affect the relative importance of pecuniary versus non-pecuniary benefits (Turner and Bowen, 1999), that is, economic incentives are not sufficient for girls to enroll and stay in male traditional fields of study (Noe, 2010).

Favara (2012) studies the role of social identity in determining individual behaviors and gender differences in economic outcomes. She integrates the concept of gender identity into an economic model of educational choices, to test the hypothesis that students' preferences are shaped by notions of gender identity congruence. Students choose their subject according to both their expected monetary returns and the pay-offs in terms of identity. If a student conforms to the social norms of the reference group, here identified by gender, she receives an indirect utility (non-pecuniary pay-off) due to a more rewarding self-image. Conversely, violating the prescriptions of gender identity generates a loss of utility. She finds that gender stereotyping affects educational choices from the age of 14, and its effect is larger for girls than for boys. She also finds evidence that gender preferences can be modified by the environment, single-sex schools lead students to a less stereotyped educational choice, after controlling for endogenous self-selection into single-sex schools. However, Park, Behrman and Choi (2018) assess causal effects of single-sex schools on different STEM outcomes in Seoul, where assignment to single-sex or coeducational high schools is random. They find significantly positive effects of all-boys schools but not for girls.

In the same line, Humlum et al. (2012), using Danish data, derive a model of career choice and identity. They characterize two identities relevant for educational choice: career oriented and social oriented typology. For a career oriented person, career and work are important for a meaningful life. Conversely social oriented persons assign more importance to cooperation, social responsibility and social issues, such as other people's well-being. They find that these two underlying factors vary systematically with the investments in level and field of education and they consider this result as evidence that identity pay-offs are an important part of educational decision making. According to their findings, students' educational choices are consistent with their self-images. Students with a career oriented identity, choose according to the financial incentive they believe is associated with their choice. This finding implies that policy makers and higher education institutions should seriously consider identity-related issues to attract high-ability students to certain careers.

Sikora and Pokropek (2012b) find that intergenerational transfers of preferences for science careers vary considerably across countries, but there are certain regularities. In many countries relevant paternal employment enhances sons' interest in science careers regardless of their field. In contrast, maternal employment inspires daughters in fewer countries and the influence tends to be limited to biology, agriculture and health careers.

3 Chilean Higher Education System

There are two types of high schools in Chile: scientific-humanist (regular), and technical-professional (vocational). Most students who intend to continue their studies in a university attend the scientific-humanistic type. In their 11th grade, students choose to follow a certain academic track based on their interests, where a track can be humanities, sciences, or arts. That way, students receive more advanced training in subjects corresponding to their tracks. Therefore, students are already choosing certain areas of study and can prepare for the college admission tests in the last two years of high school.

The higher education system consists of three types of institutions: universities, professional institutes, and technical formation centres. Universities offer licentiate degree programs and award academic degrees. There are two types of universities: 25 traditional (public and private) universities created before the year 1980, and the over 30 non-traditional private universities created after 1980. Traditional universities are coordinated by the Council of Chancellors of Chilean Universities (CRUCH), and are eligible to obtain partial funding from the state.

Chile has a single centralized admission system for its traditional universities, administered by the Department of Educational Evaluation, Measurement and Registration (DEMRE, by its Spanish abbreviation) at the University of Chile, which is under the authority of the CRUCH. Since 2003, the 25 traditional universities of CRUCH have used the group of standardized tests that comprise the University Selection Exam (PSU, by its Spanish abbreviation) -which is similar to the United States' SAT test- and the high-school GPA (NEM, by its Spanish abbreviation) to select students for admission. Starting in 2012, eight non-traditional private universities have joined the PSU admission system, thus, the 33 most selective universities of the country use this single centralized system to select their students.[2]

All entrance exam takers complete mandatory tests in mathematics and language, and they also take optional tests in other subjects (social sciences and/or sciences). Scores are scaled to a distribution with range 150 to 850 and a mean and median of 500. Entrance exam scores, along with high-school GPA, are the primary components of the composite scores used for postsecondary admissions, scholarships, and student loan eligibility. Each university must set the guidelines, requirements and selection factors for admittance to each degree program it offers and choose the weightings it deems appropriate in accordance with the rules established by CRUCH.

The score of an applicant to a degree program is calculated by applying the weightings to his or her results

[2] Starting in 2013, CRUCH decided to include high school in-class rank (besides high-school GPA) as a new selection factor in the university admissions process. This factor has a strong correlation with the high-school GPA.

for each selection factor. After taking the entrance exam and receiving their scores, students choose where to apply and submit their application to the SUA. As in other postsecondary education systems, a choice indicates both an institution and a major; we will refer to a college-major combination as a program. Students submit one application with up to ten ranked program choices. Once students apply, their entrance exam scores and GPAs are used by the universities to assign a score for each program. Once the final application score is calculated, the candidates for each program are placed into a strictly decreasing order based on their scores. Then, the program proceeds to fill their vacancies by starting with the applicant ranked first on the list, following a rigorous order of precedence until they fill all vacant slots. Applicants who are selected for their first choice are eliminated from the lists of their remaining choices. Applicants who are not selected for their first choice are placed on a waiting list and move on to compete for a spot in their second-choice program, and so forth.

Students have an incentive to rank order their choices correctly (they should not list a less-preferred choice over a more-preferred choice), nevertheless they may incorporate overall probability of admission in deciding which options to list (as they are allowed to list a maximum of ten options[3]). While students apply with some knowledge of where they might be admitted, cutoff scores may vary from year to year as demand shocks for various programs ripple through the system.

In our analysis, we only consider the 33 universities (traditional and private) that use this single centralized system to select their students since we want to examine the factors that affect the gender gap on college-major applications, thus we need students' preferences and constraints that are only available through the centralize system. In fact, we use the first preference or most wanted college program to estimate our model.

4 Data

We use data on students' characteristics and the schools they attended, obtained from the Chilean Ministry of Education. Specifically, we use data of a cohort of students graduated from high school in 2014, and applied to enter university in 2015. The data about the characteristics of university candidates, their applications and the final acceptance was provided by DEMRE.

76,680 who students graduated from high school in 2014 applied to a major-university through the centralized admission system, but only 67,426 of them applied as a first option to an area-group of universities in which they both satisfy the entrance requirements and they are not more than 100 points under the cutoff score. Only 53,942 of that set have no missing values in the variables needed for our model. To have a sample size that would allow estimating the models in a reasonable time, we chose a random sample of 20,000 cases. All area-groups have at least 50 cases in the sample.

[3]The most selective schools only consider the first four preferences, that is, if students apply to a major in the fifth preference to a highly selective school, the school does not consider that application.

Students apply to the centralized admission system ranking their major-university choices. To have a reasonable number of possible options, we group majors into disciplines or areas of study: (1) medicine & odontology, (2) health, (3) sciences, (4) civil engineering,[4] (5) technology, (6) business, (7) arts, (8) social sciences & humanities, (9) law, (10) education. Universities were clustered into four groups, where group 1-3 have the CRUCH institutions by selectivity,[5] and group 4 has the private non-traditional ones.[6]

Table 1 shows some descriptive statistics in terms of applications, enrollment and the gender gap for each area of study. Gender gaps are generally higher at the application process than in enrollment, which means that the selection process seems to decrease the gender gap. There are important gender gaps in health, civil engineering, technology and education. The reduction of the gender gap is higher at the female dominated areas. On the other hand, Table 2 depicts applications and enrollment and the gender gap for each university group. It becomes clear that the selection process increases the gender gap for females in group 2.

5 Model

This section presents our model for college-major application, guided by the institutional rules described above. In the model, there is a continuum of students with a set of high school history (h_{ij}), socioeconomic and demographic characteristics (g_{ij}), final score (a_{ij}) and academic interests. There are U universities, each with A areas of study (group of majors). Let $j \in A$ be a major, and $k \in U$ any university. Let (j, k) denote a college program. Each program differs by admission requirements, fields of study and university quality. Thus, students choose among $A \times U$ options.

Students have a certain level of knowledge in mathematics, language, social science and science, which is summarized by the vector of test scores $s_i = (s_{i1}, s_{i2}, ..., s_{iS})$. This knowledge generates a student's final application score:

$$a_{ij} = \sum_{l=1}^{S} \omega_{jl} s_{il} \tag{1}$$

where $\omega_j = [\omega_{j1}, ..., \omega_{jS}]$ is the vector of major-j-specific weights and $\sum_{l=1}^{S} \omega_{ml} = 1$.

Students apply to a college-major combination (program) as their first preference in order to maximize their

[4]This area refers to all the engineering majors than in Chile last 6 years, which are called civil engineering even though the specialties include mathematics, computer science, industrial, mechanics, electrical, civil works, mining, geology, biotechnology, chemistry, etc. These majors are highly selective, whereas the engineering majors that last 4 years are included in the technology area of study.

[5]Group 1 includes the Universidad de Chile and the Pontificia Universidad Católica de Chile; group 2 includes Universidad de Concepción, Universidad de Talca, Pontificia Universidad Católica de Valparaíso, Universidad Austral de Chile, Universidad de Santiago, Universidad Técnica Federico Santa María, and Universidad de la Frontera; group 3 includes the remaining universities belonging to the CRUCH system.

[6]Only the eight universities that use the centralized system in 2015.

expected utility:

$$U_{ijk} = \alpha_j z_{ij} + \beta_k x_{ijk} + \epsilon_{ijk} = V_{ijk} + \epsilon_{ijk} \tag{2}$$

where x_{ijk} is the vector of students characteristics relevant to their college-major choice: gender, father's area of occupation, mother's area of occupation, mixed high school class (between 40% and 60% of females in a class), female high school class (more than 60% of females in the class), geographic location (region of the country where the student currently lives), high school type (public, private-voucher o private), PSU scores s_i=(language, math, sciences, social sciences), high school class ranking[7], high school GPA (biology-chemistry, math-physics, music-arts, humanities), and per capita income).

z_{ij} is the vector of program characteristics relevant to their choice of major: application scores (a_{ij}), cutoff scores[8], average score of programs, vacancies per region, tuition, student aid[9] and the percentage of students from the same high school that the previous year enrolled in this group of universities.) Finally, ϵ_{ijk} is the error term.

Note that a student's relevant characteristic to choose an option (j, k) is the probability of being accepted by that option (p_{ijk}). In the Chilean case, the probability of being accepted by option j depends only on the PSU tests' scores. Before students apply to college, they have access to the following information: (i) their scores at the different tests ($s_i = (s_{i1}, s_{i2}, ..., s_{iS}$), (ii) their average high school GPA, (iii) the vector of major-j-specific weights ($\omega_j = [\omega_{j1}, ..., \omega_{jS}]$) and (iv) the application score of the last student enrolled at each program the year before (cutoff score).

Let $C_{ijk} \in \{0, 1\}$, $C_{ij} \in \{0, 1\}$ and $A_{ijk} \in \{0, 1\}$ be dummy variables. $C_{ijk} = 1$ indicates that the student i chooses as their first preference major j at k university group and $C_{ijk} = 0$ otherwise. $C_{ij} = 1$ indicates that the student i chooses as their first preference major j and $C_{ij} = 0$ otherwise.

The probability of being accepted at program (j, k) can be described by the following equation:

$$p_{ijk} = \theta_k(a_{ij} - \bar{a}_{ij}) + \eta_{ijk} \tag{3}$$

where a_{ij} is the final application score of student i at major j, $\bar{a}_{ij}$ is the application score of the last student admitted the year before, and η_{ijk} is the error term as the cut-off score could change every year.

Although U_{ij} and p_{ij} are not observable, students decisions are observable. Therefore, if U_{ijk*}=Max U_{ijk} student i applies to the program (j, k). In other words, we used the revealed preference principle.

We will assume that the conditional distribution of ϵ_{ijk} given a choice of major j follows a generalized

[7]Standarized score of the student ranking in high school according to their academic performance considering his/her high school cohort. The higher the grades, the higher the ranking and higher scores.

[8]We use the difference between the application score of the student minus the application score of the last student enrolled at the area-university group in 2015 (cutoff).

[9]Using tuition and student aid we compute the out-of-pocket costs for each major in each school.

Gumbel's extreme-value distribution (see equation (4)). This assumption corresponds to a nested logit model.

$$F_{U|A}(\epsilon_{jk}|j) = exp\left[-\left(\sum_{k\in U_j} exp\left(\frac{\epsilon_{jk}}{\tau_j}\right)\right)^{\tau_j}\right] \tag{4}$$

In this type of model, we have that

$$P\left[C_{ijk}=1|C_{ij}=1\right] = \frac{exp(x_{ijk}\beta/\tau_j)}{\sum_{l\in U_j} exp(x_{jk}\beta)/\tau_j)} \tag{5}$$

$$P\left[C_{ij}=1\right] = \frac{exp(\tau_j IV_j)}{\sum_{n\in A} exp(\tau_n IV_n)} \tag{6}$$

where

$$IV_j = ln\sum_{l\in U_j} exp(x_{jk}\beta/\tau_j)$$

$$P[C_{ijk}=1] = P[C_{ij}=1]P[C_{ijk}=1|C_{ij}=1] \tag{7}$$

6 Estimation Results

We estimate the model for the entire sample and also for male and female, independently. Tables 3-5 present the marginal effects for these models. Pseudo R^2 of the model for the whole sample, male sample and female sample are: 0.25, 0.24 and 0.23, respectively. The prediction errors are smaller than 1%. These estimations allow us to analyze how gender affects students' application to different majors or areas of study in higher education.

Our most relevant results show that females are, on average, 14.8% more likely to apply to health majors other than medicine & odontology, 14.7% less likely to apply to civil engineering and 7.9% less likely to apply to technology programs, everything else equal. Mother's field of occupation have higher effects on daughters in female dominated majors and law (2-4%), while father's field of occupation have higher effects on daughters in medicine & odontology, civil engineering and education (3-7%). By contrast, for male students, in almost all areas, having a father related to the area has a higher effect than having a mother with an occupation in the area. Additionally, males and females who graduated from a mostly female high school class, have higher probabilities of applying to health majors (3.5% and 5.8% respectively), while females show a 3% decrease in the probability of applying to engineering and males decrease their likelihood of applying to technology by 1.9%.[10]

[10]Although one might think that this result is due to the fact that schools with a majority of females students per class have certain characteristics that would cause this effect (such as being more conservative), this seems not to be the case. We tested whether this result is maintained considering only schools whose classes have different female proportions, that is, mainly male classes (0-40% females), mixed classes (40-60% females) and mainly female classes (60-100% females). The results we obtained for

We also find that the Math PSU score has the highest effect on the area of study to which student applies. Indeed, an increase in one standard deviation in Math PSU score increases, on average, 20.1% the likelihood of applying to civil engineering for males and 13.6% for females. Also, it decreases the likelihood of applying to health, medicine & odontology, social sciences & humanities and law. Similarly, mathematics-physics GPA are the GPA with higher effects on the study area that the student applies: an increase in one standard deviation in Math-Physics GPA increases 9.1% the likelihood of applying to civil engineering for males and 6.4% for females. Additionally, it decreases 5.1% the likelihood of females applying to health majors.

Also, the increase of biology-chemistry GPA and sciences PSU scores has effects on students' application choices. A one standard deviation increase in biology-chemistry GPA increases the likelihood of applying to health major by 3.4%, on average, for males and 4.1% for females while decreasing the likelihood of applying to civil engineering by 5.2% for males and 3.4% for females. An increase in one standard deviation in sciences PSU score increases by 7.3% the likelihood of applying to medicine & odontology for males and 6.6% for females.

In summary, higher GPA in subjects related to an area and higher PSU scores in the test related to that area are associated with a higher probability of applying to that area. Since females tend to have lower scores than males in the Mathematics and Science PSU tests[11], performance on these tests could partially explain why females tend to apply less to civil engineering and Technology than males.

We also find that students from private high schools are 6.2% more likely to apply to business majors, 3.3% less likely to apply to civil engineering and 3.1% less likely to apply to education.

Furthermore, we study how different variables affect students' application to different university groups. Table 6 shows the average marginal effects for these models. We find that, both female and male students increase their probability of applying to a group when their score is above the cutoff score (the average marginal effect of the score difference is around 4%). However, the effect of this variable is also quadratic and negative, indicating that this probability decreases when a score is much higher than the cutoff.

The selectivity of the university group also affects application; one standard deviation increase of the program's average score increases the probability of applying to the university group by nearly 6%, on average. Additional positive but small effects of applying to a university group include an increase in the percentage of regional vacancies in the program, co-payment options, and high-school colleagues' enrollments in the same university group.

the female sample show that 25.1% of those who studied with a mostly male class chose health majors, 30.7% of those who studied with a similar male/females distribution chose this major, and 30.2% of those in classes with mostly females did as well.

[11] See Appendix 1.

7 Profiles

In the first exercise of this section, we estimate the predicted choice probabilities for each area of study and university group for students with average values in all variables except gender (male or female). This exercise isolates the unobservable preferences in college-major choices that are tied specifically to gender. Figure 1 shows the simulated probability of applying to each area of study and group of universities, for males and females with average values in all the remaining control variables. In Figure 1A, we can observe important differences in the probabilities by gender in health, civil engineering, technology, social sciences & humanities and education. In fact, an average male applicant is much more likely to apply to engineering and technology majors than a female with average characteristics, and an average female applicant is much more likely to apply to non-physician health majors than a male with average characteristics. Otherwise, as Figure 1B shows, the differences between both genders in the probabilities of applying to different groups of universities tend to be small.

In addition, we simulate the probabilities for each area of study and university group for students of different gender, gender composition of the class (mostly female or mixed high school class), and parent's area of occupation (of the same and different gender of the student), and average values for the rest of the variables. This exercise allows us to create 94 different student's profiles. We find that only changing the gender variables, the probability of choosing an area of study can change up to 41%, while the probability choice for a university group can change up to 18%.

We go further and analyze students' profiles related to their achievements. In particular, we consider male and female students with average characteristics, but who have different achievement in a given area, measured by the PSU and GPA score of the subject related to the area. The results can be seen in Figure 2A, which shows that an average applicant, male or female, with higher PSU score and GPA in the subject related to the area, have higher probability of applying to the most selective major in the area of study, that is, to medicine & odontology looking at the science PSU score and GPA in biology or chemistry, to civil engineering looking at the Math PSU and GPA in mathematics and physics, and to law looking at the social science PSU and GPA in the humanities.[12]. Notice that, in all cases, this trend is more pronounced for males than females, which would mean that, given the same GPA and tests results, males have higher probability than females of choosing the most selective major in each area, including humanities.

These results are similar if we analyze the probability of an average applicant (male and female) applying to different groups of universities. Figure 2B shows that higher scores increase application to groups of universities 2, 3 and 4, to a greater extent for females than for males, while the opposite occurs for group 1. Males are more sensitive than females to the difference between their score and the cutoff score of the group of universities,

[12] Appendix 2 displays the average PSU scores by area. It is easy to see that for the case of Chile, the three majors mentioned are the most selective ones.

proxy for the probability of being accepted to the most selective group, and women are more sensitive than men to the less selective groups of universities. In other words, males tend to apply to the most selective universities more than females, everything else equal.

In summary, these gender gaps on college major choices are base on differentiated behaviors of males and females: we find that high achievement students (males and females) tend to increase their probability of application to the most selective areas (medicine & odontology, civil engineering and law), and the most selective universities. However, when we compare average males and females students with the same academic results, this tendency is stronger for males, showing that males are more likely to bet on the most selective majors considering their test results. That is, males apply to competitive programs even where they are marginal candidates while equally qualified women will tend not to apply to those same competitive programs. These findings reveal a crowding out effect caused by prevailing male preferences, which in turn contributes significantly to the gender gap in STEM fields.

8 Counterfactuals

We run three different counterfactual exercises. First, we look at the effects on the probability of applying to each area of study if females students had the same preference parameters as males students, that is, if females consider the same factors and in the same magnitude as males. Second, we look at the effect of applying to each area of study if males students had female students' preference parameters. Third, we look at the effect of applying to each area of study if males had female preference parameters and, simultaneously, females had male preference parameters. Table 7 shows the percentage of female and male applicants by area.

Figure 3A shows the first counterfactual exercise, if females students had male preference parameters, females would apply less to health, social sciences & humanities, education and medicine & odontology while applying more to civil engineering and technology. Over all, considering both male and female students, there would be less applicants for health (13% vs 21% on the original case), education (6% vs 8%), social sciences & humanities (12% vs 13%) and medicine & odontology (6% vs 7%). Table 7, column 1 and 7 shows that: (i) female dominated majors would continue to be dominated by females: 62% vs 76% in health, 62% vs 68% in arts, 60% vs 64% in social sciences & humanities, and 59% vs 69% in education; (ii) civil engineering would continue to be male dominated, but less: 44% vs 25% female; (iii) sciences and technology which are male dominated (49% and 26% female) switch to have equal or more females than males (53% and 50%). We also find that the average of applicant's mathematics and science PSU score would decrease in civil engineering, the average science PSU score would decrease in business, technology and social sciences & humanities and increase in education, and the average history PSU score would increase in medicine & odontology and business but decrease in arts &

music and education.

The second counterfactual shows that if males had female preference parameters, males would apply more to health, education, medicine & odontology, social sciences & humanities and art; less to civil engineering and technology (see figure 3B). In aggregate terms, it would be more applicants for health (28% vs 21%), education (10% vs 8%), social sciences & humanities (14% vs 13%), medicine & odontology (8% vs 7%) and arts & music (4% vs 3%). Table 7, column 1 and 4 show the results of this exercise: (i) female dominated majors would continued be dominated by females, but the women's proportion in most cases is lower: 57% vs 76% in health, 54% vs 68% in arts, 57% vs 69% in education, 52% vs 61% in medicine & odontology, 59% vs 64% in social sciences & humanities, but 54% vs 53% in La; (ii) civil engineering, technology and business would continue to be male dominated, but less with 39% vs 25% of females applying to civil engineering, 48% vs 26% to technology and 47% vs 46% to business. Moreover, we find that civil engineering and technology would have applicants with higher class ranking, mathematics and science PSU scores, and higher GPA. The average of applicants' science PSU score would increase in business.

The third counterfactual shows that if males had female preference parameters while females had male preference parameters. Over all, the percentage of applicants for each area would remain the same (with less than 1% difference). Table 7, column 1 and 10 show the applicants responses to this counterfactual excercise: (i) there are differences in the percentage of female applicants: social sciences & humanities and law would continue being female dominated (55% and 55% vs 64% and 53% of women); (ii) other areas that used to be female dominated would be male dominated: 41% vs 76% of woman in health, 44% vs 61% in medicine & odontology, 47% vs 69% in education, and 48% vs 68% in art; (ii) the areas which used to be male dominated would be female dominated or equitable: 73% vs 26% in technology, 61% vs 25% in civil engineering, 53% vs 49% in sciences and 50% vs 46% in business. Finally, we find that civil engineering and technology would have applicants with higher ranking and GPA, while health would have applicants with lower ranking and GPA. Also, technology would have applicants with lower mathematics and science PSU scores, while business and education would have applicants with higher science PSU score.

These counterfactual exercises allow us to compute the gender gap for each area in each case. We simulate the percentage of male and female applicants using their original preference parameters, using the female preference parameters for all students, using the male preference parameters for all students, and using the other sex's parameters. Columns 2 and 3 of table 7 display the original percentage of applicants by gender using their own parameters and column 4 shows the gender gap for each area as a comparison. Columns 5-6 display the percentage of applicants by gender using female parameters and column 7 shows the gender gap if male and female use female preference parameters. Columns 8-9 display the percentage of applicants by gender using male parameters and column 10 shows the gender gap if male and female use male preference parameters.

Finally, columns 11-12 display the percentage of applicants by gender using other sex's parameters and column 13 shows the gender gap in this case. As a result of these exercises, we find that the gender gap decreases more in female dominated areas if males had female preference parameters or all students have the other sex's parameters rather than females having male parameters. Otherwise, the gender gap decreases usually more in male dominated areas if females had male preferences rather than all students having the other sex's parameters or females having male parameters.

9 Understanding Gender Differences in College Major Choice

In order to understand the importance of the underlying factors that contribute to the difference in college major choices among females and males, we rely on the classic Oaxaca decomposition, that will allow us to to express the difference in the average predicted value of the dependent variable as:

$$\overline{\hat{Y}}_A - \overline{\hat{Y}}_B = (\overline{X}_A - \overline{X}_B)\hat{\beta}_A + \overline{X}_B(\hat{\beta}_A - \hat{\beta}_B) \tag{8}$$

where $\overline{\hat{Y}}_A$ and $\overline{\hat{Y}}_B$ are the average predicted value of choosing an area of study and a group of universities for males and females, respectively. $\overline{X}_A$ and $\overline{X}_B$ correspond to the average values of the independent and observable variables for males and females, respectively. $\hat{\beta}_A$ and $\hat{\beta}_B$ are the estimated coefficients for males and females. Note that the first term on the right-hand side of equation 8 is the gender difference on the mean probability of the choice of each program due to different observable characteristics ($X = [x_{ijk}; z_{ij}]$), while the second term is the difference due to unobservable variables that affect the university program choice.

Table 8 displays the results of this exercise for the 10 majors considered. We use bootstrapping and re-sample 100 times in order to estimate the confidence intervals.[13] It is hard to find statistically significant differences between the contribution of the parameters or the data to the gender gap. However, in the three areas with higher gender gaps: Health, Civil Engineering and Technology, we find statistically significant differences at 95% confidence level. In fact, in these three cases, the gender gaps could de attributed more to the unobservables than to the data, even though it is known that in the data, particularly in the PSU tests scores, there are gender differences that affect the observable differences in application.[14]

10 Conclusions

This paper looks at the gender gaps in applications to college programs in Chile. We estimate the model using a nested logit and simulate counterfactuals. Our results suggest that the gender differences we observe

[13]The estimation of the model takes a considerable amount of time (around 15 hours), we only re-sample 100 times.
[14]See appendix 1 for distribution of PSU score by gender.

in college major application and enrollment are derived from students' preferences, since males and females show different application patterns to different areas of study. In particular, females are 14.8% more likely to apply to health majors, 14.7% less likely to apply to civil engineering and 7.9% less to technology. It's worth pointing out that by preferences we mean the decision behavior patterns we observe males and females make, patterns that can be based on individual motivations but also be the product of social constructions. In this sense, the different application probabilities by area according to student gender, suggest the existence of gender stereotypes that affect college major application.

These stereotypes are also linked with the role model of parents. Our results show that the mother's area of occupation have higher effects on daughters in female dominated majors and law (2-4%), whereas father's area have higher effects on daughters in medicine & odontology, civil engineering and education (3-7%). The father's area of occupation usually has a higher effect than the mother's area for male students. Therefore, while males seem to have a higher tendency to reproduce gender patterns of the previous generation, females seem to be influenced by both, their father's and mother's area of occupation.

Gender differences in application also appear to be a product of high school performance. We find evidence of the importance of previous achievement (related to the probability of being selected) in the application process. Mathematics PSU test score and Math-Physics GPA have the highest effect on area of application. Also, scoring higher than the cutoff score of a program increases the probability of application. This is relevant for increasing the percentage of females in STEM majors, considering that females tend to have lower scores in the Mathematics PSU test and Science PSU test than males.

More importantly, looking at the area of study and university group, we see that males have a higher tendency, compared to females, to choose the most selective program if they have good results on the PSU tests. This also suggests that decisions could be influenced by social stereotypes, as males could feel more social pressure to be successful, choosing the most selective option.[15] On the other hand, it could be the case that females may feel more insecure about their own knowledge, tending to believe that they are less apt for more selective options.

We also find that females and males from mostly-male high school classes are less likely to apply to health majors, and more likely to apply to civil engineering. Consequently, a higher interaction with students of the same gender, increases the probabilities of following the patterns of application of that gender. This finding suggests that class composition tend to differentiate student choices from the areas historically dominated by their gender.

All these results show that the gender gaps on college major choice are related not only to the female choice behavior, but also to the choice of males. In fact, we find that males apply to competitive programs even when

[15] Appendix 3 shows the majors of application of students of outstanding achievement by gender.

they are marginal candidates while equally qualified women will tend not to apply to those same competitive programs. These findings can be understood as a crowding out effect caused by prevailing male preferences being an important contributor to the gender gap in STEM fields. In this sense, our counterfactual analysis illustrates that the gender gap decreases more in female dominated areas if males have female preference parameters rather than females having male parameters, with the exception of civil engineering and technology where the gender gap decreases more if females have male preference parameters rather than males having female parameters. These findings are relevant for policy makers, since we demonstrate that the effort to reduce the gender gap must seek to not only have females encouraged to pursue male dominated programs, but also that males choose female dominated programs, decreasing the missallocation of talent.

Acknowledgements

We would like to thank the Department of Evaluation, Measurement and Educational Registration (DEMRE) and the Ministry of Education for granting us access to the individual PSU scores, application and enrollment data. Additionally, we would to thank the Agency for Quality of Education for providing individual information regarding family background and tests scores of the Standardized tests SIMCE.

Alejandra Mizala acknowledges financial support from FONDECYT Project N° 1140834 and from PIA-CONICYT (Basal Project FB0003).

Paola Bordón acknowledges support from (CONICYT/FONDAP/15130009) and the Millennium Nucleus on Social Development.

This research was also partially supported by the supercomputing infrastructure of the NLHPC (ECM-02) at the Faculty of Physical and Mathematical Sciences, Universidad de Chile.

Finally, we would like to thank the participants at SECHI, PEGNet, ESPE, SCHPP and AEDE conferences for their valuable comments.

References

[1] Akyol, P., Krishna, C. and Lychagin, S.(2017). Gender Differences in College Major Choice: Preferences versus Performance.

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

[3] Bartolj, T. and Polnec, S. (2012). College major choice and ability: Why is general ability not enough. *Economics of Education Review*, 31, 996-1016.

[4] Booth, A. and Nolen, P. (2011). Choosing to compete: how different are girls and boys? *Journal of Economic Behavior & Organization*, 81 (2), 542-555.

[5] Booth, A. and Nolen, P. (2012). Gender differences in risk behavior: does nurture matter? *The Economic Journal*, 122 (558), F56-F78.

[6] Bordón, P. and Fu, C. (2015). College-Major Choice to College-then-Major Choice. *The Review of Economic Studies*, 82, 1247-1288.

[7] Croson, R. and Gneezy, U. (2009). Gender differences in preferences. *Journal of Economic Literature*, 47 (2), 448-474.

[8] Delaney, J. and Devereux, P. (2019). Understanding Gender Differences in STEM: Evidence from College Applications. *forthcoming Economics of Education Review.*

[9] Dickson, L. (2009). Race and gender differences in college major choice. *The annals of the American Academy of Political and Social Science*, 627 (1), 108-124.

[10] Favara, M. (2002). The Cost of Acting "Girly": Gender Stereotypes and Educational Choices. *IZA DP* N 7037, November.

[11] Gneezy, U., Leonard, K.L. and List, A.J. (2009). Gender differences in competition: evidence from a matrilineal and a patriarchal society. *Econometrica*, 77 (5), 1637-1664.

[12] Gneezy, U., Niederle, M. and Rustichini, A. (2003). Performance in competitive environments: gender differences. *Quarterly Journal of Economics*, 118 (3), 1049-1074.

[13] Gneezy, U. and Rustichini, A. (2004). Gender and competition at a young age. *American Economic Review*, 94 (2), 377-381.

[14] Gupta, N.D., Poulsen, A. and Villeval, M.C. (2005). Male and female competitive behavior: experimental evidence. IZA Discussion Paper (1833).

[15] Humlum, M.K., Kleinjans, K.J. and Nielsen, H.S. (2012). An economic analysis of identity and career choice. *Economic Inquiry*, 50 (1), 39-61.

[16] Jurajda, Š. and Münich, D. (2011). Gender Gap in Admission Performance under Competitive Pressure. *American Economic Review Papers and Proceedings*, 101(3), 514-518.

[17] Justman, M. and Méndez, S.J. (2018). Gendered choices of stem subjects for matriculation are not driven by prior differences in mathematical achievement. *Economics of Education Review*, 64, 282-297.

[18] Killgore, W.D.S. and Yurgelun-Todd, D.A. (2004). Sex related developmental differences in the lateralized activation of the prefrontal cortex and the amygdala during the perception of facial affect. *Perceptual and motor skills*, 99 (2), 371-391.

[19] Kimura, D. (2000). Sex and cognition. The MIT Press.

[20] Kurtz-Costes, B., Rowley, S.J. , Harris-Britt, A. and Woods,T.A. (2008). Gender stereotypes about mathematics and science and self-perceptions of ability in late childhood and early adolescence. *Merrill-Palmer Quarterly*, 54 (3), 386-409.

[21] Lenroot, R.K., Gogtay, N., Greenstein, D.K., Wells, E.M., Wallace, G.L., Clasen, L.S., Blumenthal, J.D., Lerch, J. Zijdenbos, A.P., Evans, A.C. (2007). Sexual dimorphism of brain developmental trajectories during childhood and adolescence. *Neuroimage*, 36 (4), 1065-1073.

[22] Niederle, M and Vesterlund, L. (2007). Do women shy away from competition? Do men compete too much? *Quarterly Journal of Economics*, 122 (3), 1067-1101.

[23] Niederle, M. and Vesterlund, L.. (2010). Explaining the gender gap in math test scores: the role of competition. *The Journal of Economic Perspectives*, 24 (2), 129-144.

[24] Niederle, M. and Yestrumskas, A. H. (2008). Gender Differences in Seeking Challenges: The Role of Institutions. NBER WP N 13922.

[25] Noe, C. (2010). Family background and female's choice of male's subjects: evidence from Italy. Child Working Paper N 1.

[26] OECD (2015). The ABC of Gender Equality in Education: Aptitude, Behavior, Confidence.

[27] Örs, E., Palomino, F. and Peyrache, E. (2013). Performance Gender-Gap: Does Competition Matter? *Journal of Labor Economics*, 31(3), 443-499.

[28] Park, H., Behrman, J. and Choi, J. (2018). Do single-sex schools enhance students' STEM (science, technology, engineering, and mathematics) outcomes? Economics of Education Review 62, 35-47.

[29] Saltiel, F. (2019). What's math got to do with it? Multidimensional ability and the gender gap in STEM. Department of Economics, University of Maryland.

[30] Speer, J.D. (2017). The gender gap in college major: Revisiting the role of pre-college factors. *Labour Economics*, 44, 69-88.

[31] Sikora, J. and Pokropek, A. (2012). Gender segregation of adolescent science career plans in 50 countries. Science Education DOI 10.1002/sce.20479. Published online 9 February 2012 in Wiley Online Library (wileyonlinelibrary.com).

[32] Sikora, J. and Pokropek, A. (2012). Intergenerational transfers of preferences for science careers in comparative perspective. *International Journal of Science Education*, 34(16), 2501-2527.

[33] Sutter, M. and Rützler, D. (2010). Gender differences in competition emerge early in life. IZA Discussion Paper N 5015.

[34] Traferri, A. (2011). Gender differences in major choice and college entrance probabilities in Brazil. Working paper IE-PUC, N 403.

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

[36] Wiswall, M. and Zafar, B. (2014). Determinants of college major choice: Identification using an information experiment. *The Review of Economic Studies*, 82(2), 791-824.

[37] Wiswall, M. and Zafar, B. (2017). Preference for the workplace, investment in human capital, and gender. *The Quarterly Journal of Economics*, 133(1), 457-507.

[38] Zafar, B. (2009). College major choice and the gender gap. *Journal of Human Resources*, 48 (3), 545-595.

Table 1: Area Applications and Enrollment 2015

Area	% Applications			% Enrollment		
	Females	Males	Gap	Females	males	Gap
Medicine & Odon.	61%	39%	-22	58%	42%	-16
Health	78%	22%	-55	76%	24%	-51
Sciences	47%	53%	6	46%	54%	8
Civil engineering	24%	76%	52	25%	75%	51
Technology	27%	73%	46	27%	73%	46
Business	49%	51%	2	48%	52%	4
Arts & Music	63%	37%	-26	67%	33%	-34
Social Sc. & Hum.	64%	36%	-28	62%	38%	-24
Law	54%	46%	-8	52%	48%	-4
Education	68%	32%	-36	68%	32%	36

Source: Own calculation.

Table 2: University Applications and Enrollment 2015

University	% Applications			% Enrollment		
	Females	Males	Gap	Females	males	Gap
Group 1	54%	46%	8	49.5%	50.5%	-1
Group 2	48%	52%	-4	43%	57%	-14
Group 3	56%	44%	12	51%	49%	2
Group 4	57%	43%	14	52%	48%	4

Source: Own calculation.

Table 3: Average marginal effects for area (whole sample)

Variable	Medicine & Odon.	Health	Sciences	Civil Eng.	Technology	Business	Arts & Music	Social Sc & Hum.	Law	Education
Female	2.4%	14.8%	-0.2%	-14.7%	-7.9%	-0.6%	1.1%	2.2%	-0.1%	3.4%
Parent's area same sex	1.3%	2.2%	0.3%	1.1%	0.5%	1.8%	-0.2%	1.9%	4.0%	0.03%
Parent's area different sex	4.9%	1.1%	0.2%	2.0%	-0.2%	0.4%	1.8%	0.8%	1.2%	1.0%
Female high school class	0.3%	3.4%	-0.2%	-0.8%	-0.6%	-1.5%	0.1%	-0.3%	-1.3%	1.1%
Mix high school class	0.6%	2.2%	-0.4%	-0.6%	-0.3%	-0.7%	0.3%	-0.5%	-1.0%	0.6%
High school ranking	0.3%	-0.9%	-0.3%	2.8%	-0.5%	-0.5%	0.4%	-0.3%	-0.2%	-0.7%
Language PSU Score	0.8%	0.6%	0.2%	-2.3%	-1.7%	-3.0%	0.4%	3.4%	0.1%	1.7%
Math PSU Score	-4.7%	-12.7%	0.1%	16.5%	0.6%	6.8%	-0.2%	-4.4%	-3.0%	0.4%
Sciences PSU Score	7.2%	3.4%	0.5%	-2.8%	-0.6%	-2.0%	-0.6%	-1.8%	-0.7%	-1.9%
History PSU Score	-0.5%	-2.3%	-0.1%	-0.8%	-0.4%	-0.1%	-0.4%	0.05%	5.4%	-1.0%
Biology-Chemistry GPA	2.5%	4.0%	0.8%	-4.2%	-0.01%	-0.7%	-0.5%	-1.1%	-0.04%	-0.4%
Math-Physics GPA	-1.9%	-5.0%	-0.1%	7.6%	0.5%	2.0%	-0.4%	-2.0%	-0.4%	-0.7%
Arts-Music GPA	-0.1%	0.5%	-0.1%	-0.6%	0.1%	0.1%	0.5%	-0.2%	-0.2%	-0.1%
Humanities GPA	1.4%	1.0%	-0.2%	-2.8%	-1.1%	-1.2%	-0.4%	2.8%	0.9%	-0.1%
Per capita income	0.1%	-0.4%	-0.03%	0.7%	-0.3%	0.1%	0.2%	0.3%	0.2%	-0.9%
Voucher high school	0.4%	-0.5%	0.4%	-0.8%	-0.2%	1.1%	0.04%	-0.3%	0.6%	-0.7%
Private high school	0.8%	-1.9%	0.4%	-3.3%	-1.2%	6.2%	1.2%	0.1%	0.9%	-3.1%

All coefficient are statistically significant at 99% confidence level, except the effect of History PSU score at Social Sc.& Humanities, which is statistically significant at 95% confidence level, and the effect of Biology-Chemistry GPA at Technology, which is statistically significant at 90% confidence level. Estimations have fixed effects by region.

Source: Own calculation.

Table 4: Average marginal effects for area (female sample)

Variable	Medicine & Odon.	Health	Sciences	Civil Eng.	Technology	Business	Arts & Music	Social Sc & Hum.	Law	Education
Parent's area same sex	0.6%	2.9%	-0.2%	-1.9%	0.3%	1.2%	0.8%	2.0%	4.2%	-0.4%
Parent's area different sex	7.1%	-0.7%	1.5%	3.2%	0.1%	-0.1%	1.9%	0.6%	0.6%	3.0%
Female high school class	0.8%	5.8%	-0.1%	-3.0%	0.03%	-0.7%	-1.5%	-2.2%	-0.3%	1.2%
Mix high school class	1.0%	5.6%	0.1%	-2.8%	-0.3%	0.1%	-1.3%	-2.8%	0.03%	0.4%
High school ranking	0.04%	-1.2%	-0.2%	1.9%	-0.2%	0.5%	0.3%	-0.4%	0.4%	-1.1%
Language PSU Score	1.4%	-0.03%	-0.2%	-2.3%	-0.8%	-2.1%	0.3%	3.5%	-1.1%	1.6%
Math PSU Score	-3.6%	-14.1%	0.9%	13.6%	1.3%	7.5%	0.2%	-4.4%	-2.7%	1.0%
Sciences PSU Score	6.6%	3.8%	0.2%	-1.9%	-0.2%	-1.9%	-0.6%	-2.1%	-0.7%	-2.4%
History PSU Score	-0.7%	-3.3%	-0.2%	-0.3%	-0.1%	-0.1%	-0.3%	0.2%	6.1%	-1.3%
Biology-Chemistry GPA	2.0%	4.1%	0.8%	-3.4%	0.7%	-0.9%	-0.9%	-1.4%	0.2%	-0.9%
Math-Physics GPA	-1.5%	-5.1%	0.1%	6.4%	0.2%	2.7%	-0.4%	-1.9%	-0.2%	-0.5%
Arts-Music GPA	-0.1%	0.6%	0.03%	0.1%	-0.4%	0.2%	0.7%	-0.03%	-0.5%	-0.5%
Humanities GPA	1.7%	0.6%	-0.8%	-2.1%	-0.6%	-2.4%	-0.3%	3.3%	0.4%	0.4%
Per capita income	0.3%	-0.2%	-0.2%	0.3%	-0.2%	0.1%	0.3%	0.4%	0.2%	-0.9%
Voucher high school	0.7%	0.6%	0.2%	-1.9%	-0.01%	0.3%	0.04%	0.6%	0.8%	-1.2%
Private high school	0.8%	-0.2%	0.00%	-5.9%	-0.9%	4.2%	2.2%	2.0%	0.9%	-3.0%

All coefficient are statistically significant at 99% confidence level, except the effects of Language PSU score in Health and Private high school in Science, which are not statistically significant.

Estimations have fixed effects by region.

Source: Own calculation.

Table 5: Average marginal effects for area (male sample)

Variable	Medicine & Odon.	Health	Sciences	Civil Eng.	Technology	Business	Arts & Music	Social Sc & Hum.	Law	Education
Parent's area same sex	2.2%	1.8%	0.6%	1.6%	0.6%	2.1%	-0.6%	1.6%	3.8%	0.3%
Parent's area different sex	2.9%	1.7%	-1.0%	-1.4%	-1.3%	1.4%	1.8%	0.9%	2.7%	0.2%
Female high school class	0.6%	3.5%	0.2%	-0.5%	-1.9%	-1.7%	0.4%	-0.2%	-1.0%	0.6%
Mix high school class	0.8%	0.3%	-0.6%	-0.1%	-0.2%	-0.8%	0.5%	0.4%	-0.8%	0.5%
High school ranking	0.3%	-0.02%	-0.4%	3.8%	-0.9%	-1.7%	0.4%	-0.4%	-0.9%	-0.2%
Language PSU Score	0.3%	1.1%	0.6%	-2.0%	-2.8%	-3.9%	0.5%	3.1%	1.4%	1.6%
Math PSU Score	-5.7%	-11.5%	-0.9%	20.1%	-0.02%	6.1%	-0.7%	-4.1%	-3.4%	-0.5%
Sciences PSU Score	7.3%	3.5%	1.0%	-4.2%	-1.0%	-2.1%	-0.5%	-1.4%	-0.6%	-1.2%
History PSU Score	-0.2%	-1.2%	-0.1%	-1.3%	-0.7%	0.02%	-0.4%	-0.3%	4.7%	-0.6%
Biology-Chemistry GPA	2.9%	3.4%	0.8%	-5.2%	-0.5%	-0.4%	-0.1%	-0.6%	-0.2%	0.2%
Math-Physics GPA	-2.3%	-4.8%	-0.3%	9.1%	0.6%	1.3%	-0.3%	-2.1%	-0.5%	-0.9%
Arts-Music GPA	-0.1%	0.1%	-0.2%	-1.0%	0.7%	0.1%	0.4%	-0.2%	0.04%	0.2%
Humanities GPA	1.1%	1.2%	0.3%	-3.5%	-1.7%	0.1%	-0.5%	2.4%	1.3%	-0.6%
Per capita income	0.1%	-0.7%	0.1%	1.3%	-0.4%	0.1%	0.1%	0.1%	0.2%	-1.0%
Voucher high school	0.01%	-1.5%	0.6%	0.1%	-0.4%	2.0%	0.02%	-1.1%	0.4%	-0.1%
Private high school	0.5%	-3.5%	0.7%	-0.6%	-1.4%	8.2%	-0.1%	-1.8%	1.1%	-3.2%

All coefficient are statistically significant at 99% confidence level, except the effect of High school ranking at Health, which is statistically significant at 95% confidence level, and the effects of Math PSU score at Technology and History PSU score in Business, which are not statistically significant. Estimations have fixed effects by region.

Source: Own calculation.

Table 6: Average marginal effects for group

Variable	Whole sample	Female sample	Male sample
Difference of score	4.21%	4.06%	4.34%
Difference of score^2	-1.27%	-1.35%	-1.20%
Program average score	6.00%	6.14%	5.37%
% of regional vacancies of the program	0.03%	0.03%	0.03%
Copayment	0.02%	0.01%	0.03%
% Previous high school class enrolled in the group	0.03%	0.03%	0.03%

All coefficients are statistically significant at 99% confidence level.

Source: Own calculation.

Table 7: Counterfactuals Excercises: Percentage of female and male applicants by area

	% of Applicants											
	True parameters			Female parameters			Male parameters			Other sex's parameters		
Area	Female	Male	Gap	Female	Male	Gap	Female	Male	Gap	Female	Male	Gap
Medicine & Odon.	61%	39%	-22%	52%	48%	-4%	54%	46%	-8%	44%	56%	12%
Health	76%	24%	-52%	57%	43%	-14%	62%	38%	-24%	41%	59%	18%
Sciences	49%	51%	2%	49%	51%	2%	53%	47%	-6%	53%	47%	-6%
Civil engineering	25%	75%	50%	39%	61%	22%	44%	56%	12%	61%	39%	-22%
Technology	26%	74%	48%	48%	52%	4%	50%	50%	0%	73%	27%	-46%
Business	46%	54%	8%	47%	53%	6%	50%	50%	0%	50%	50%	0%
Arts & Music	68%	32%	-36%	54%	46%	-8%	62%	38%	-24%	48%	52%	4%
Social Sc. & Hum.	64%	36%	-28%	59%	41%	-18%	60%	40%	-20%	55%	45%	-10%
Law	53%	47%	-6%	54%	46%	-8%	54%	46%	-8%	55%	45%	-10%
Education	69%	31%	-38%	57%	43%	-14%	59%	41%	-18%	47%	53%	6%

Source: Own calculation.

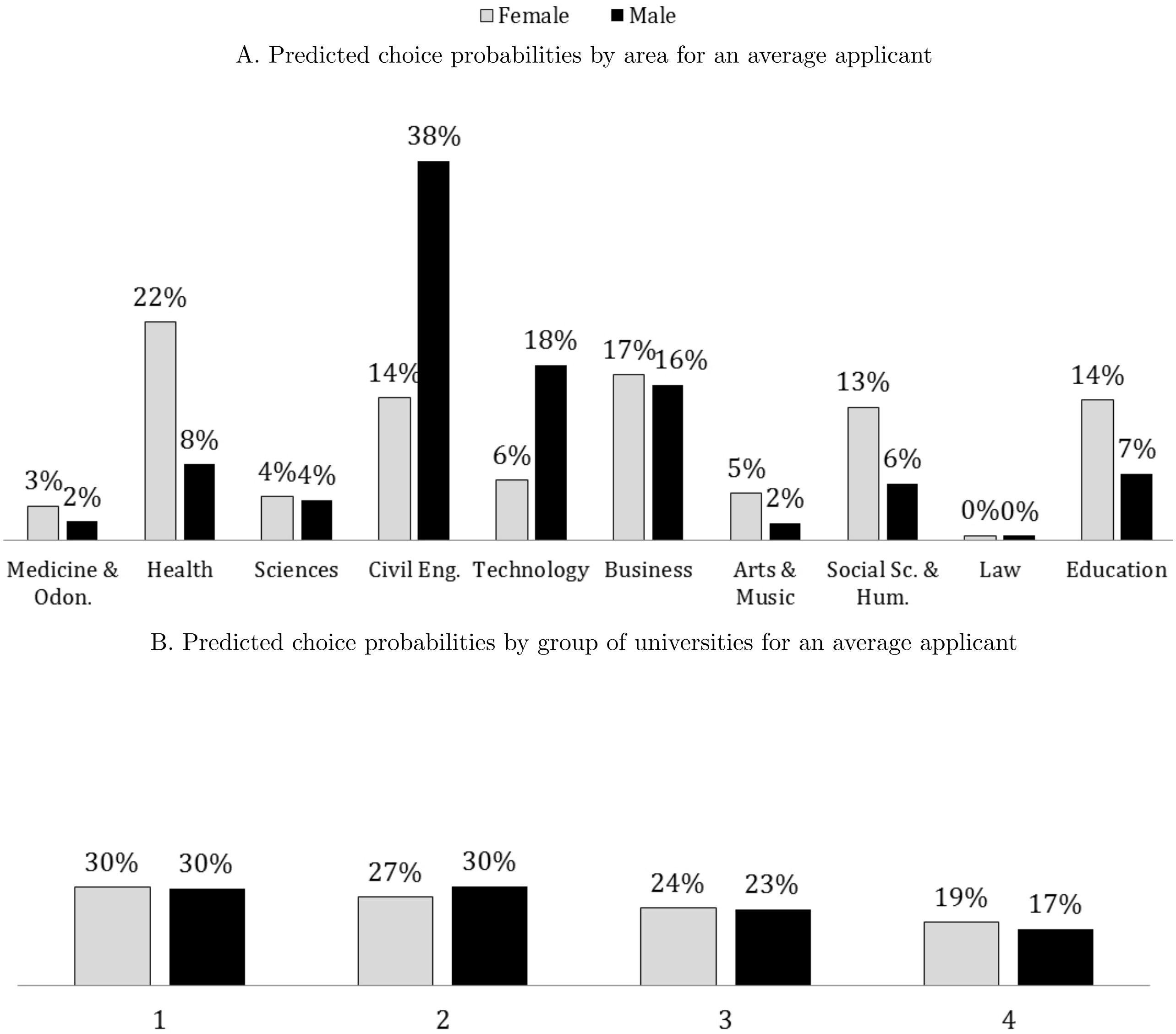
Table 8: Decomposition Analysis to explain the gender gap

	Confidence interval by Area						
	Average probability for		Gender difference	$(\overline{X}_A - \overline{X}_B)\hat{\beta}_A$	$\overline{X}_B(\hat{\beta}_A - \hat{\beta}_B)$	Contribution of the	
	Males ($\overline{\hat{Y}}_A$)	Females ($\overline{\hat{Y}}_B$)	in probability			Data	Parameters
Medicine & Odon.	[6.3%,7.0%]	[8.8%,9.7%]	[-3.2%,-2.1%]	[-1.4%,-0.5%]	[-2.3%,-1.1%]	[17.2%,52.4%]	[47.6%,82.8%]
Health	[9.8%,10.9%]	[29.0%,30.4%]	[-20.2%,-18.4%]	[-5.2%,-3.3%]	[-16.0%,-14.0%]	[17.1%,26.3%]	[73.7%,82.9%]
Sciences	[4.1%,4.7%]	[3.1%,3.7%]	[0.5%,1.4%]	[-0.1%,0.8%]	[0.2%,1.3%]	[-6.3%,72.6%]	[27.4%,106.3%]
Civil engineering	[31.2%,32.6%]	[9.3%,10.2%]	[21.4%,23.0%]	[8.2%,9.9%]	[12.2%,14.1%]	[37.3%,44.5%]	[55.5%,62.7%]
Technology	[12.1%,13.0%]	[4.0%,4.6%]	[7.8%,8.8%]	[-0.3%,1.5%]	[6.9%,8.7%]	[-4.1%,17.6%]	[82.4%,104.1%]
Business	[10.0%,11.1%]	[7.6%,8.3%]	[1.9%,3.1%]	[0.8%,2.0%]	[0.6%,1.9%]	[32.3%,74.7%]	[25.3%,67.7%]
Arts & Music	[1.9%,2.2%]	[3.1%,3.7%]	[-1.6%,-1.0%]	[-1.3%,-0.6%]	[-0.8%,0.0%]	[43.5%,104.3%]	[-4.3%,56.5%]
Social Sc. & Hum	[9.1%,10.0%]	[15.2%,16.1%]	[-6.7%,-5.5%]	[-3.8%,-2.4%]	[-3.8%,-2.2%]	[39.6%,61.9%]	[38.1%,60.4%]
Law	[5.9%,6.7%]	[6.2%,6.8%]	[-0.6%,0.4%]	[-0.9%,0.2%]	[-0.4%,0.9%]	[-1112.2%,1386.1%]	[-1286.1%,1212.2%]
Education	[5.3%,6.1%]	[10.0%,10.7%]	[-5.1%,-4.1%]	[-2.5%,-1.2%]	[-3.6%,-2.1%]	[26.0%,53.9%]	[46.1%,74.0%]

Notes: Gender difference in probability is defined as: $\overline{\hat{Y}}_A - \overline{\hat{Y}}_B$. A negative number means the contribution is in the opposite direction of the gender difference in probability

Source: Own calculation.

Figure 1: Predicted choice probabilities for an average applicant

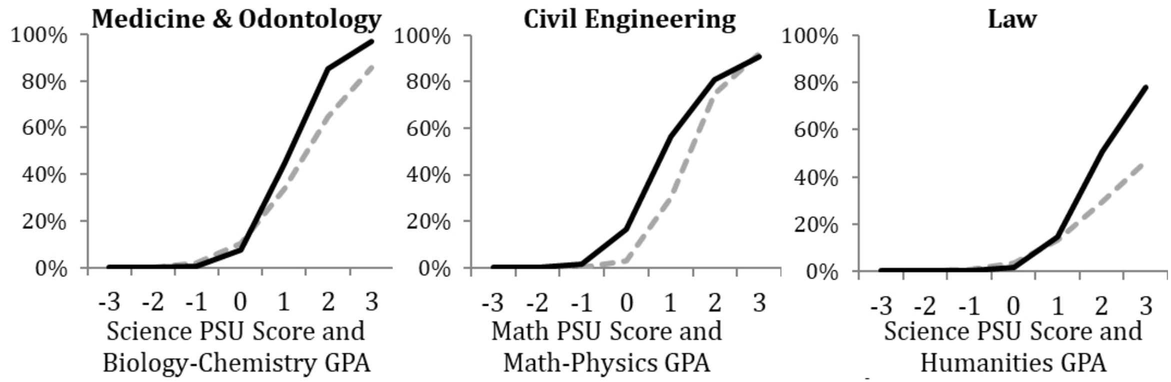


Source: Own calculation.

Figure 2: Predicted choice probabilities for an average applicant by previous achievement

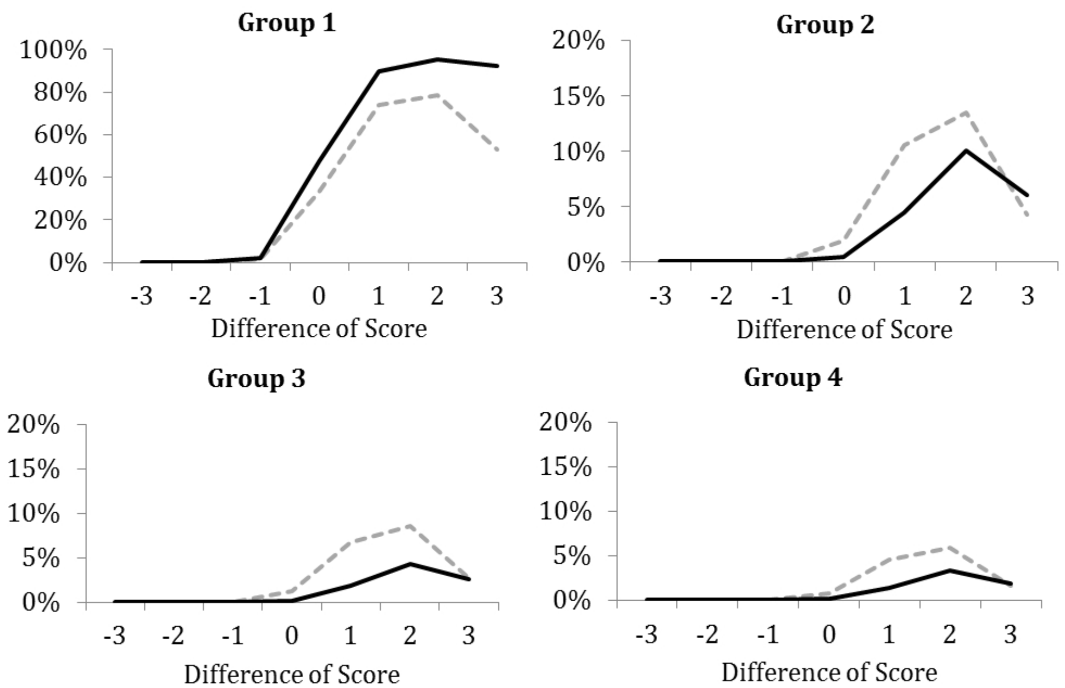
- - - Female —— Male

A. Predicted choice probabilities to most selective areas for an average applicant, with different GPA and PSU Scores



Note: Prediction made using average values for all variables except gender, PSU Score and GPA of the corresponding area.

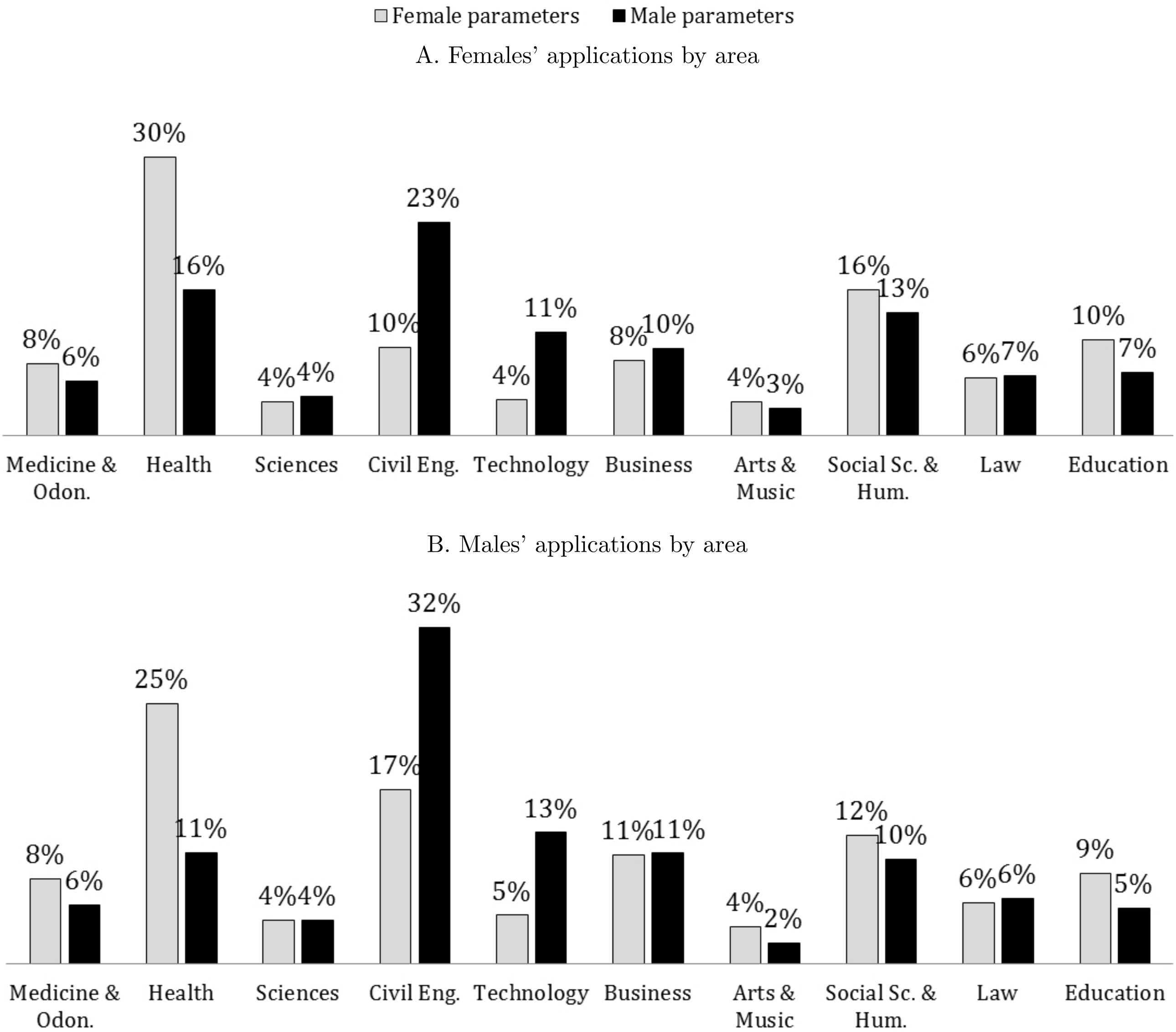
B. Predicted choice probabilities by group of universities for an average applicant, with different Difference of Score



Note: Prediction made using average values for all variables except gender and Difference of Score for corresponding university group.

Source: Own calculation.

Figure 3: Percentages of applicants by area for the counterfactuals exercises



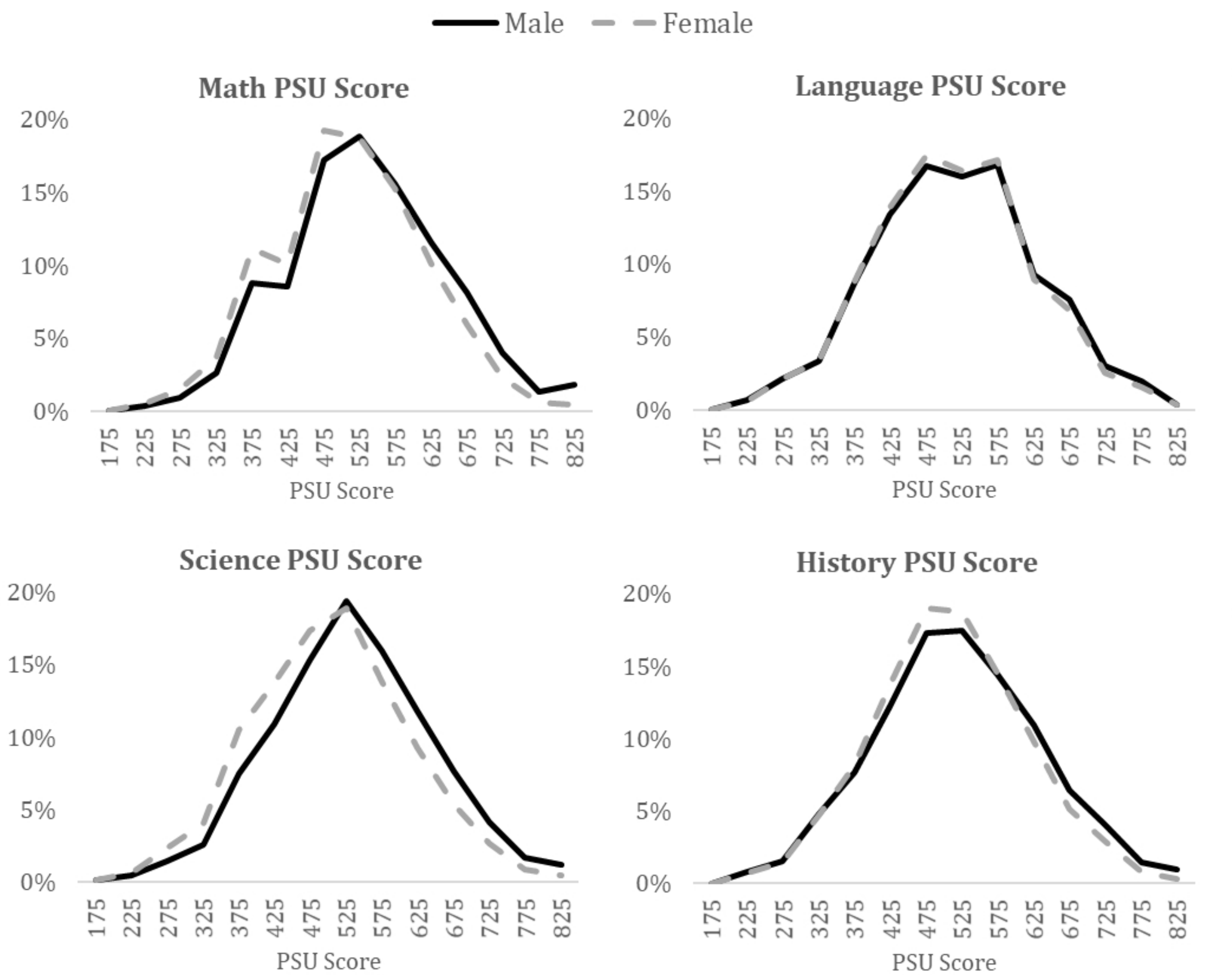
Source: Own calculation.

11 Appendixes

11.1 Appendix 1: Gender differences in PSU Scores

Figure Supplementary 1 shows that female students tend to have lower math and sciences PSU Scores than male students.

Figure Supplementary 1: PSU Scores for the Admission Process of 2015



Source: Own calculation.

11.2 Appendix 2: PSU Scores by area

As you can see in Table Supplementary 1, medicine & odontology is the area with the highest average of PSU Scores of language, math and science. Civil engineering is the second area with higher average of math PSU Scores. Law is the area with the highest history PSU scores and the second area with higher average PSU score of language.

Supplementary Table 1: Average of PSU Scores by area.

Area	Language	Math	History	Science
Medicine & Odon.	667.7	686.3	614.0	691.7
Health	585.0	587.4	542.6	583.5
Sciences	588.3	611.9	562.2	601.9
Civil engineering	581.8	630.1	558.8	583.1
Technology	577.1	614.1	553.8	581.0
Business	576.3	603.8	578.9	546.3
Arts & Music	606.0	573.4	600.1	549.7
Social Sc. & Hum.	615.5	564.7	624.4	536.4
Law	634.4	576.9	658.6	540.2
Education	586.0	560.6	577.7	536.6

Note: PSU Scores have a normal distribution with a mean equal to 500 points and a maximum score of 850 points. Estimations only consider the students who applied through the Unique Admission System (SUA) in 2015.

Source: Own calculation.

11.3 Appendix 3: Majors of application of students with outstanding achievement

From Table Supplementary 2, in particular for the case of the areas related to mathematics (sciences, civil engineering, technology and business), we see that outstanding male students have a higher tendency than females to apply to Civil engineering, the most selective degree of these areas. The same phenomenon is observed for degrees related to sciences (sciences, health and medicine & odontology) and humanities (social sciences & humanities and law), in which outstanding male students apply more than females to law. However, in those two last cases the differences are lower.

Supplementary Table 2: Specific majors of application of students with outstanding achievement

Students with at least 700 points in Math PSU test who applied to the areas of sciences, Civil engineering, technology or business		
	Male	Female
Civil Engineering	41%	24%
Economics or Administration[a]	12%	14%
Physics and/or Astronomy	1%	1%
Pedagogy in Mathematics	1%	1%
Biochemistry	0%	1%
Other	45%	59%
Students with at least 700 points in science PSU test and who applied to the areas of sciences, medicine & odontology or health		
	Male	Female
Medicine	71%	69%
Nursing	2%	8%
Odontology	5%	7%
Biochemistry	2%	2%
Medical Technology	3%	2%
Other	17%	13%
Students with at least 700 points in history and/or language PSU test who applied to the areas of social sciences and humanities or Law		
	Male	Female
Law	21%	18%
Psychology	4%	7%
Journalism	2%	3%
Literature	1%	2%
Sociology	2%	2%
Other	71%	68%

[a] in Chile this mayor is called "Ingenieria Comercial".

Estimations only consider the students who applied through the Unique Admission System (SUA) in 2015.

Source: Own calculation.