I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 14561

Family Ties, Geographic Mobility and the Gender Gap in Academic Aspirations

Lídia Farré
Francesc Ortega

JULY 2021

IZA Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 14561

Family Ties, Geographic Mobility and the Gender Gap in Academic Aspirations

Lídia Farré
University of Barcelona, IAE-CSIC and IZA

Francesc Ortega
CUNY and IZA

JULY 2021

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Family Ties, Geographic Mobility and the Gender Gap in Academic Aspirations*

This paper provides new evidence supporting that gender differences in post-graduate educational choices contribute to the glass ceiling in the labor market. We study the decision to pursue an advanced degree form an internationally renowned institution, which greatly facilitates access to top jobs. Relying on a unique dataset on applications to a highly selective program that provides merit-based graduate fellowships to Spanish students, we show that women apply for the fellowships at lower rates than observationally equivalent male graduates. We also implemented a large-scale survey on current college students and show that female college graduates have stronger family ties than males, which restricts their geographical mobility and has a negative effect on their educational aspirations. Importantly, the previous pattern is reversed in STEM fields: female graduates in STEM participate in the fellowship program at equal or higher rates than comparable males. In fact, we show that female STEM students originate from more educated families, have higher academic ability, and higher educational and earnings aspirations than women in other fields.

JEL Classification: J3, J7

Keywords: gender, post-graduate, fellowships, family ties, geographic mobility

Corresponding author:
Lídia Farré
University of Barcelona
Gran Via de les Corts Catalanes, 585
08007 Barcelona
Spain
E-mail: lidia.farre@gmail.com

* We are grateful to the *La Caixa Foundation* and the participating universities for providing access to their data and for their overall support. We also thank Pol Campos-Mercadé, Rosa Ferrer, Laura Hospido, Pedro Rey, Carlos Sanz and Judit Vall for helpful comments. We are particularly indebted to Amparo Asensi, Dolors Baena, David Casado, Domènec Espriu, Montserrat Masoliver, Pablo Pareja Santi Roca and Josep Ros. We acknowledge the financial support of the Ministry of Science and Innovation (grant PID2019-110397RA-I00), the Government of Catalonia (grant SGR2017-644) and the *La Caixa Foundation*.

1 Introduction

In the last few decades gender gaps in the labor market have narrowed, largely thanks to increases in women's educational attainment. Women account for over 60% of recent college graduates in many countries, but remain under-represented in top positions, both in academic settings and in the private sector (Blau and Kahn (2017)). A vast literature explores the many factors that contribute to this persistent gender gap.

Our paper focuses on a new angle: a college degree or even an advanced degree often do not guarantee access to top jobs. Both in academia and private sector, being hired for a high-paying position at a leading firm or public institution often requires a post-graduate degree from an internationally renown university, particularly for entry-level positions. Despite accounting for the majority of college graduates, women remain in the minority in many prestigious graduate programs. Evidence from the field of Economics & Business illustrates this point. Top MBA programs remain majority male (Wallen et al. (2017)) and women account for only 32% of the entering cohorts in Economics Ph.Ds (Bayer and Rouse (2016) and Boustan and Langan (2019)).[1]

We hypothesize that college-educated women make less ambitious post-graduate educational choices than their male counterparts. This could be a subtle, yet important, factor to help explain the absence of women in top positions in the labor market, impacting their chances to be hired for a top job or to remain in one when having children (as in Cortes et al. (2020)). More specifically, we examine the post-graduate plans and choices of college-educated women through the lens of participation in a highly selective fellowship program. Gaining admission to internationally renown graduate programs is difficult, because of the harsh competition. It is also expensive in terms of tuition and other expenses, and typically requires moving to a different city or country. For these reasons, and because of high social payoffs in terms of innovation and knowledge diffusion, governments and philanthropic institutions in many countries offer fellowships to academically excellent students interested in pursuing graduate studies at the world's leading universities.[2]

We examine college graduates' decision to participate in the *La Caixa Foundation* (LCF) Fellowship Program, largely aimed at funding graduate studies abroad for Spanish

[1] Women are also under-represented in STEM fields that are typically associated with above-average labor market prospects, and are also less likely to pursue professional degrees and doctoral studies (e.g. Bertrand and Hallock (2001), Black et al. (2008) and Hsieh et al. (2019).)

[2] One of the most famous graduate fellowship programs in the world is the *Fulbright U.S. Student Program*, established in 1946, and offering approximately 2,000 grants each year. Alumni of the program occupy leading positions across a wide range of professions.

citizens with excellent academic records. These data provide a unique window into the educational choices of high-achieving male and female college graduates. Specifically, we combine data on the whole set of applicants to the program over a number of years with administrative graduation records for four large universities. These data allow us to estimate the application rates of male and female college graduates with a high level of disaggregation, and to parse out the effects of academic ability and socioeconomic status, in addition to gender. Furthermore, we conducted a large survey of the current students in the universities included in our study to shed light on students' preferences and constraints regarding post-graduate studies. In particular, the survey contains rich information regarding the student's aspirations, geographic mobility and the intended participation in competitive programs, such as the LCF fellowship program.

Our analysis reveals that GPA is the most important determinant of participation in the LCF program, although socio-economic status and field of study also play important roles. We also find that the aggregate participation rates of male and female graduates are practically the same. However, this masks a large gender gap in participation due to offsetting differences in characteristics (largely field of study and average grades). Our most detailed estimates suggest that female college graduates are 22% less likely to apply for the LCF fellowships than male graduates with equal grades, socio-economic status, and in the same field of study and university. This finding suggests that the educational aspirations of college-educated women are lower, on average, than for comparable men. This finding is in line with the results in Fluchtmann et al. (2020) who document that Danish women's application rates to high-paying jobs are significantly lower than for comparable men.

To investigate further the determinants of the gender gaps in participation in the fellowship program, we conducted a large survey among the (approximately 35,000) students currently enrolled in the four universities included in our study. The survey data corroborate the presence a gender gap in the *intention* to apply for the fellowships after graduation. Furthermore, we document that female college students have stronger family ties, measured by hours of care provided to relatives (e.g. younger siblings or grandparents) and having a romantic partner. We are also able to show that these factors reduce females' geographic mobility to a larger extent than for males, generating the gender gap in the intention to conduct graduate studies abroad.

An important qualification to our previous findings is that women with STEM majors exhibit equal (or higher) participation rates in the LCF program than men with the same observable characteristics. The survey data confirm that females in STEM majors are as likely, or more, than their male counterparts to intend to apply for graduate fellowships and to study abroad. We also document that STEM women are positively selected. They originate from more educated families, have higher academic ability, and higher educational and earnings aspirations than women in other fields.

Our results contribute to the literature on the absence of women in high-earnings, high-status positions, often referred to as the glass ceiling (Bertrand et al. (2019)). Several explanations have been proposed to account for the gender disparities at the top of the labor market. Early studies emphasized gender discrimination (Rouse and Goldin (2000)) and differences in skill levels (Goldin et al. (2006)). More recently, researchers have also documented gender differences in preferences for competition (Niederle and Vesterlund (2007), Buser et al. (2014), Hospido et al. (2019)) and in the balance between family and work (Bertrand (2013), Azmat and Ferrer (2017), Bursztyn et al. (2017), Keloharju et al. (2019) and Hospido et al. (2019)). In addition, several studies have also pointed out the role of reviewers in candidate selection processes, whose decisions may be affected by implicit bias, gender stereotypes or other factors (Bagues and Esteve-Volart (2010), Breda and Ly (2015), Hospido and Sanz (2019), Farré and Ortega (2019), Montalban and Sevilla (2020)).

Our findings indicate that explanations based on differences in human capital accumulation remain important: highly talented women make less ambitious educational choices than their male counterparts, and this is partly due to stronger family ties that constrain their geographical mobility. This result is consistent with the recent studies showing that commuting costs penalize women in the labor market (Le Barbanchon et al. (2019) and Petrongolo and Ronchi (2020)). The issue of the geographical mobility of men and women has regained interest in the recent years. Several studies document that marriage, cohabitation and children reduce women's mobility to a larger extent than men's. For instance, Shauman and Xie (1996) find lower geographic mobility for female scientists. They argue that this is related to their higher likelihood of being in dual career marriages and that their mobility falls further, relative to their male partners, when they have children. Similarly, Jürges (2006) documents that marital status (and cohabitation) reduce the geographical mobility of women relative to men, confirming the earlier work of Bielby and Bielby (1992) and others.

Our paper is also related to the literature exploring the reasons for the lower presence of women in STEM disciplines (e.g. Hill et al. (2010)). Using PISA test scores, Guiso et al. (2008) document gender differences in comparative advantage, which tend to be smaller in countries with a greater culture of gender equality. These authors find that, in most countries, girls typically obtain lower math scores than boys, but higher reading scores. Carrell et al. (2010) show that professor gender has a large effect on women's decision to choose science majors. This finding has been recently confirmed by Porter and Serra (2020) for Economics majors with a field experiment. Relatedly, Carlana (2019) shows that teacher gender stereotypes induce girls to underperform in math and self-select into less demanding high schools. Also, Carpio and Guadalupe (2019) document gender differences in the decision to enter the technology sector driven by social norms and show that de-biasing messages can be an effective policy tool to increase female participation. Our results reveal that STEM women participate in the fellowship program at higher rates than men with the same characteristics. That is, while women may encounter high entry barriers to STEM fields, those that clear them appear to be positively selected. Our analysis also demonstrates that, compared to women in other fields, STEM women originate from more educated families, have higher college-entry grades and higher educational and professional aspirations. This evidence is consistent with the recent findings that highlight the role of comparative advantage to account for the allocation of students across fields of study and subsequent labor market success (Altonji et al. (2012) and Kirkeboen et al. (2016)).

Last, our paper also contributes to the literature comparing the academic achievements of boys and girls. It is well established that, from an early age, girls "leave boys behind" in terms of educational attainment (Fortin et al. (2015)). However, the evidence is less clear in regards to the comparison between the most talented males and females. Our data contain individual records for over 160,000 college graduates and allow us to produce highly detailed comparisons of the GPA distributions by gender with high granularity. We find that while women typically have higher mean GPA than men, they tend to be under-represented in the top 5% of the grades distribution (for a given major and university).[3] This difference could be relevant to explain the gender gaps in highly meritocratic contexts, such as seeking admission to graduate studies at a leading institution. Our estimates show that the probability to apply for a LCF fellowship increases exponentially for high-GPA students. However, the difference in the shares of men and women among high-grade earners are too small to explain away the observed gap in participation rates in the program.

The structure of the paper is as follows. Section 2 presents our data sources. Sec-

[3]The under-representation of women among extremely high-achieving students has also been documented for the US among high school students in math (Ellison and Swanson (2010)). It is also well-known that males' aptitude test scores exhibit larger variance than females', and that males outnumber females among high-achievers along several (but not all) dimensions (Hedges and Nowell (1995)).

tion 3 estimates aggregate participation rates in the fellowship program using data for the four universities included in our study. Section 4 extends the analysis further by focusing on individually linked records for a single university. Section 5 examines intended participation in the fellowship program using our survey data. Section 6 discusses the mechanisms that account for the gender gaps in participation. Section 7 explores why STEM women differ from women in other disciplines, and Section 8 concludes.

2 Data Sources

2.1 Applications to LCF Fellowships Program

The *La Caixa Foundation* (LCF) is a private financial institution in Spain that has been providing graduate fellowships since 1982. The LCF fellowship program is the largest program in Spain funding graduate studies abroad, currently awarding 120 fellowships per year (plus around 100 more for graduate studies in Spain).[4] The awards fund both Master's degrees and PhDs in all fields of study and the fellows typically gain admission to the most prestigious institutions worldwide.[5]

Our data contains applications for the period 2014-2018 to three separate FLC sub-programs: graduate studies in North America or Asia, in European countries (other than Spain) and doctoral studies in Spain.[6] The data contains complete information on roughly 9,000 applications, 55% of which from female candidates. The success rate (relative to complete applications) is around 9%.[7]

[4]Similar programs aimed at Spanish citizens are the Foundation Alonso Martin Escudero (60 fellowships), the Foundation Mutua Madrilena (40), Fulbright (25), the Ramon Areces Foundation (22), Rafael del Pino Fellowships (10) and the Foundation Barrie (10).

[5]To date, the LCF has funded close to 5,000 awards, with 70% funding studies abroad. The top destination countries are USA, Spain, UK, Germany and France. Similarly, the top (narrow) fields of study have been: Art & History (14%), Health Sciences (13%), Engineering (13%) and Economics & Business (12%).

[6]The latter sub-program also requires geographic mobility. The host institution cannot be located in the province where the candidate conducted his/her undergraduate studies.

[7]Importantly, women account for only 49% of the successful applicants. For a detailed analysis of the determinants of success in the program, see Farré and Ortega (2019). Another study employing data from the *La Caixa Foundation* fellowship program is Garcia-Montalvo (2014), which showed that the labor market careers of award recipients experience a large and persistent boost (both in academia and private sector).

2.2 College graduates records

We obtained access to the individual (anonymized) graduation records of the 4 largest public universities in Catalonia: the University of Barcelona (UB), the Autonomous University of Barcelona (UAB), the Polytechnic University of Catalonia (UPC) and University Pompeu Fabra (UPF). These universities are located in the Barcelona metropolitan area and together account for 77% of the enrollment in public colleges in Catalonia.[8] Three of these universities offer a large number of majors across all major fields of study, whereas UPC is almost completely specialized in engineering.

Our period of analysis ranges from academic year 2009-2010 to 2018-2019 and the data have wide coverage across all academic disciplines. Among the roughly 162,000 individual observations, about 43.3% of the graduates belong to Social Sciences, 31.2% to STEM disciplines, 13.7% to Health & Life Sciences, and 11.8% to Arts & Humanities.

The graduation records include student-level information on year of graduation, major, gender and GPA. The data show that 55.4% of all graduates are women. Across fields, we observe that they account for a large majority in all fields, except for STEM where women are only 28.8% of the graduates. The female share rises to 65.7% in Social Sciences, 67.5% in Arts & Humanities and 73.2% in Health & Life Sciences.

Graduation GPA is reported on a 0-10 scale (with a minimum of 5 required to pass a class). Across all graduates, the mean GPA is 7.11. However, we observe differences by gender and also field of study. The average GPA for women is 7.23, about 4% higher than for men (6.96). By field of study, the highest mean GPAs are found in Health (7.45) and Arts & Humanities (7.42), followed by Social Sciences (7.16) and STEM (6.79), indicating that grading tends to be harsher in the latter field.[9]

The University of Barcelona (UB) agreed to share with us information on students' family background (e.g. parental education and occupation) and, more importantly, to link their data with the LCF applications dataset at the individual level.[10] The UB is the largest university in our dataset, accounting for almost half of the graduates.[11] It

[8]Public colleges account for 85% of the overall (in-person) tertiary enrollment in Catalonia, which amounts to 173,485 students in academic year 2018-2019.

[9]Mean GPAs are very similar across all universities (ranging between 7.23 and 7.29) except for the engineering school where the mean value is 6.72.

[10]Obtaining permission to match administrative data across different sources has become much more difficult after the application of the General Data Protection Regulation (GDPR) in the European Union. This regulation was adopted in 2016 but implemented from May 25, 2018). To link the two datasets while preserving student anonymity, each party encrypted the students' National Identification Number using the same key. Then we simply merged the two datasets on the basis of the encrypted identifier.

[11]The overall number of UB graduates for academic years 2009-2010 through 2018-2019 is 75,596, or

is also fairly similar to the other universities in terms of the share of females (65% in academic year 2018-2019) and enrollment distribution across fields of study, with the exception of UPC that specializes in engineering. The LCF applications dataset contains 588 complete applications from UB graduates, corresponding to 506 unique individuals, over the period 2014-2018. The data show that 44 of these applicants were awarded a fellowship, that is, the success rate was 8.7%.

2.3 Survey College Students

We conducted a survey of all students at our four participating universities that had completed over half of the 240 credit hours required for graduation (and had registered for at least one class in academic year 2019-2020). The survey was conducted online in January-February 2020 and the response rate was 14%, leading to 4,848 essentially completed questionnaires out of a target population of 34,559 students.[12]

The survey respondents match well the administrative records in terms of gender, field of study and GPA. In particular, 58% of the respondents are women, 39% of respondents are STEM majors, and the *average* GPA is almost the same for men and women in the same field of study.[13] We present a comprehensive set of descriptive statistics for the survey in Section 5.

approximately 7,500 per year (Table C.15). Further details can be found in Appendix Table C.16.

[12]This response rate is quite typical of online surveys conducted by these universities on their own student population. The response rate for our survey also compares favorably to Paredes et al. (2020). If we include questionaries that are only partially complete the response rate increases to 16%.

[13]In the survey, 39% of the respondents are in STEM fields, 31% in Social Sciences, 18% in Health and 13% in Arts & Humanities. By university, the survey female shares range between 66% and 69% for the UB, UAB and UPF and fall to 30% for the engineering school (UPC). The corresponding values in the administrative records (2009-2018) are 60% to 64% and 27%, respectively. Turning now to GPA, the mean value among all survey respondents is 7.2, only slightly above the average of 7.1 found in the administrative records. The average (self-reported) GPA in the survey ranges between 7.23 and 7.40 for the three universities with a broad range of majors but it is significantly lower at the engineering school (6.77). These figures closely match the corresponding numbers in the administrative records. We also note that the students that complied and completed the survey tend to be positively self-selected in terms of grades: 32.6% and 18.8% have grades above the 75th and 90th percentiles, respectively (relative to the grade distribution based on university-field in the administrative records). The administrative records show that women, on average, obtain slightly higher grades than men in all universities (with female-male ratios ranging between 1.02 to 1.03) with the exception of the UPC, which effectively exhibits gender parity (0.996 FM ratio). The situation is similar in the survey data, although the gender gaps are narrower. For the UPC respondents, the female-male ratio is estimated to be 1.003, while the values for the other universities range between 1.001 and 1.014. The narrower GPA gender gaps are consistent with the higher response rates for women and high-GPA students.

3 Aggregate participation rates

Using administrative records on the graduates of 4 large universities for the period 2009-2010, we now compute the participation rates in the FLC program on the basis of these data.[14] We refer to the resulting participation rates as *aggregate* because the calculation does not require linking individuals across the graduation records and the applications dataset, which we can only do for one university. Naturally, gender differences in academic ability, age, or socio-economic status can introduce differences in the fellowship participation rates of men and women that we cannot control fo. In other words, this section can only provide gender gaps in *unconditional* participation rates but we can utilize the graduation records from all four universities in our study.[15]

Besides estimating participation rates, the graduation records also allow us to carry out a highly detailed characterization of the whole distribution of GPA by gender and field of study. This analysis is particularly important for our purposes because there may be gender differences at the upper tail of the GPA distribution, which is the population more likely to consider participation in the LCF fellowship program.

3.1 Definition

To compute the aggregate participation rate (PR) we tally the number of applicants of a given gender g that graduated in year t in field of study f from university u ($Applicants_{g,f,u,t}$) and divide it by the number of graduates at the same level of aggregation ($Graduates_{g,f,u,t}$).[16] That is,

$$PR_{g,f,u,t} = \frac{Applicants_{g,f,u,t}}{Graduates_{g,f,u,t}}. \tag{1}$$

Clearly, we can compute participation rates at a more aggregated level, adding across

[14] Our data begins with academic year 2009-2010 and ends with 2018-2019. For short, we refer to each academic year on the basis of the Fall semester. Hence, following this convention, our data is for the period 2009-2018. For each of the more than 160,000 graduates in our data, we observe the year of graduation, the major, the gender and the GPA. Importantly, the FLC applicants dataset contains the university of origin, the graduation year, field of study, and the gender of each applicant. The field of study identifies the discipline in a broad sense. Thus, each field of study contains several majors.

[15] We are authorized to identify data pertaining to the UB but, at this point, we can only report figures pertaining to the other universities when pooled together.

[16] Clearly, the LCF applicants in year t may have graduated in a previous year. We examine this issue further below. For now, it suffices to point out that the overall number of graduates in a given university and field of study is fairly constant over our sample period. Thus, the normalization we are applying will not be far from the true participation rate.

years, universities or fields of study. More importantly, our dataset only contains applications to the LCF fellowship program for years 2014-2018. As we explain in detail in Appendix A, this creates a censoring problem. We address it by focusing on graduation cohorts 2012-2014, which are largely free of this problem.

3.2 Gender gap in aggregate participation rates

Table 1 collects our estimates for the participation rates in the FLC program. The top panel reports the aggregate participation rates obtained when using all graduation cohorts (2009-2018). The first column shows that the FLC program received 1,530 applications (between 2014 and 2018) by graduates from the 4 universities in our study, which accounts for slightly less than one fifth of all the applications they received over that period. Restricting to the uncensored cohorts (second panel), the number of applications falls to 815. The second panel (Uncensored cohorts) also shows that the number of female applicants was 20% higher than the number of male applicants. Likewise, the number of female graduates was also 20% higher than the corresponding figure for males. As a result, we estimate a participation rate of 1.66% for both genders.

In light of these estimates, it would be tempting to conclude that male and female college graduates have similar educational ambitions, in terms of attending prestigious graduate programs. However, it is important to recognize that there are important gender differences in fields of study (with women severely under-represented in STEM), average grades and over the preferred type of graduate degree (Master's versus Ph.D.).

We begin our exploration of these issues by computing gender-specific participation rates specific to each field of study relying on the graduation records for the four universities included in our study.

3.3 The role of field of study

To get a sense of the role of field of study, we grouped all majors into 4 broad areas: STEM, Health & Life Sciences, Arts & Humanities and Social Sciences. Table 1 reports the participation rates by field. Focusing on the second panel, which pools all universities, we find participation rates above the 1.66% average in all fields except for Social Sciences. The highest participation rate is found in Health & Life Sciences (3.41%), followed by Arts & Humanities (2.20%) and STEM (1.84%), while the lowest value is found in Social Sciences (0.87%).[17] The same pattern is found in each of the 3 universities offering majors across all fields of study.

In terms of participation gender gaps, the data suggest a female advantage in STEM (female-male ratio FM=1.31), gender parity in Arts & Humanities (FM=1.01), and lower female participation in Health & Life Sciences (FM=0.87) and in Social Sciences (FM=0.80).[18]

Because women are highly under-represented in STEM and highly over-represented in Social Sciences, gender differences in field of study will account for an important part of the participation gender gap obtained when pooling together all fields of study.

3.4 Preferences by type of degree and distance

It is also possible that men and women differ on their geographical mobility or desired length of graduate studies, which would affect their need to seek funding for graduate studies abroad. In fact, the FLC fellowship program funds both PhDs and Master's degrees. In addition, about 30% of the fellowships are given out through a sub-program that funds PhD studies within Spain (but in a province that differs from the one where the applicant attended college). This section examines program participation by type of degree sought and whether the desired graduate institution is located abroad.

As shown in Table 2, the overall application rate (pooling fields, genders and sub-programs) is 1.66%, and the majority of applicants (0.87/1.66=52%) are interested in Master's degrees abroad, 17% in PhDs abroad and 31% seek funds for PhDs in Spain. In terms of gender gaps, the figures in the Table reveal that the female-male ratio in applications seeking to fund Master's degrees abroad is 1.04, indicating that women are slightly more interested than men in this type of program. In contrast, the data suggest that the demand for PhD programs abroad is substantially lower for women than for men, with a corresponding female-male participation ratio of 0.81. Hence, conditional on location outside of Spain, women are less interested than men in enrolling in a PhD program.

It is also informative to compare the application rates by gender seeking to fund

[17] Many factors may explain why interest in graduate studies abroad differs across fields. An obvious one is that Social Sciences includes majors in Law and Social work that involve a great deal of country-specific knowledge.

[18] In fact, for all four universities in our study, we observe STEM female-male participation ratios above one (ranging between 1.06 and 2.55), as shown in Table C.12. It is worth noting that the gender-specific participation rates obtained when pooling all universities are weighted sums of the corresponding values for each of the universities. However, this is not the case for the female-male ratios.

PhD programs in Spain. The data reveal that the female participation rate is 6% higher than the male value. Thus, conditional on being interested in PhD studies, women are substantially less keen than men in studying abroad. While this could reflect differences in 'pure' preferences, it may also stem from the presence of constraints, such as family ties, that may lead men and women to make different choices, even if their objective functions are similar. We return to this point when analyzing the survey data.

3.5 The role of GPA

From an early age, women receive higher grades, on average, than men (Fortin et al. (2015), Hsin (2018)). It is also well known that the GPA distribution of male students typically exhibit greater variance than the corresponding distribution for females. In other words, men are over-represented among weak students but tend to be over-represented among the most academically gifted students.[19]

Given that the decision to apply for a highly selective fellowship is mostly relevant for students with the highest grades, it is important to gauge the differences between men and women at the top tail of the GPA distribution. We start from each student's raw GPA (on a 0-10 scale with a passing grade of 5). As shown in the top panel of Table C.18, average grades are lower in STEM (6.79) than in other disciplines (ranging from 6.99 in Social Sciences to 7.47 in Health & Life Sciences). In addition, on average women's GPA is 4% higher than men's, although the gender gap is practically non-existing in all fields except for Social Sciences. Because grades distributions vary across fields, we compute each individual's position in the GPA distribution corresponding to *his or her major and university*. We then pool all students and compare the resulting distributions for men and women.

Figure 1 plots the gender-specific percentile GPA distributions pooling all fields of study across all universities (but defining percentiles separately by university and major). More specifically, we plot the density of students (by gender) by 5 percentage-point brackets across the whole range of GPA percentiles (based on the combined distribution of males and females). The Figure clearly shows that women are greatly under-

[19] In their analysis of PISA test scores for a large set of countries, Guiso et al. (2008) find that women are under-represented at the highest levels of math aptitude. Specifically, on average across countries, there are 0.6 girls for every boy with a math score above the 95th percentile (of the country-level distribution of scores), with a range from 0.4 in Korea to 1.1 in Indonesia. Focusing on the U.S. alone, the authors find that the ratio of U.S. girls to boys who are above the 95th percentile of the scores distribution is 0.59. In contrast, there are 1.83 girls for every boy with a reading score above the 95th percentile (of the country-level distribution of scores).

represented at the bottom of the grades distribution and also slightly under-represented at the top. Figure 2 plots the difference between the two distributions at each percentile bracket. Below the 35th percentile, the density of women is always below the density of men, for as much as 1.5 percentage points (in the first bracket). Between the 35th and 95th percentiles, women are over-represented relative to men. However, women are under-represented at the top bracket (percentiles 95-100) by about 0.2 percentage points (see also Appendix B). Because the LCF fellowships are highly selective, the relative scarcity of women at the very top of the GPA distribution could account for an important chunk of the gender gap in participation.[20]

In conclusion, when we pool the administrative GPA records across all universities we observe that women have higher GPA than men *on average* but are slightly under-represented among students with grades above the 95th percentile. The regression analysis in the next section will quantify how important this factor is in accounting for gender differences in participation in the fellowship program.

4 The determinants of participation

Importantly, for one of the universities (the UB) we were allowed to merge the individual graduation records with the LCF applications dataset. Thus, we are able to use regression analysis to quantify the role of the determinants of the decision to participate in the fellowship program and to estimate the *conditional* participation gender gap net of differences in observables (including age and family background).

4.1 Summary statistics

It is also worth noting that the UB is the largest university in Catalonia in terms of enrollment, accounting for approximately one third of the college population in the whole region. Let us begin by providing some basic descriptive statistics (Table C.16). Women are as likely as men to be between the 90th and 95th percentiles of their major GPA distribution. However, they are 0.23 percentage points less likely to be in the top 5% of their major's grade distribution than men (although we cannot reject the

[20] Section B.2 contains an alternative method for comparing the GPA of men and women along the whole distribution based on standardizing each student's GPA on the basis of the mean and standard deviation of his/her major and university. That method is less demanding in terms of data but makes distributional assumptions. At any rate, the qualitative conclusions coincide with what we just discussed. In a quantitative sense, the alternative method suggests a larger gender gap at the top 5% of the GPA distribution (Table C.18).

hypothesis of equal values). In addition, female graduates at UB are 9.4 percentage points less likely to be in STEM and 8.5 percentage points more likely to be in Health. These observations will resonate in the regression analysis presented later.[21]

It is also interesting to examine the bottom panels of Table 1, which report the aggregate participation rates for the UB and the other institutions. The data show that the aggregate participation rate is lower at the UB (1.07% for both genders combined) than at the other three universities, partly due to its higher specialization in Social Sciences.[22] In addition, the female-male participation ratio (pooling all fields) for the UB is estimated to be 0.90, revealing a substantial gender gap that largely reflects its relatively small STEM share.

4.2 Regression analysis

Next, we explore the factors that determine the decision to apply for an LCF fellowship using the UB-FLC matched dataset. Specifically, we estimate a linear probability model for participation in the fellowship program and analyze the roles played by GPA (and the relative position in the grade distribution of one's major), field of study, gender and socio-economic status. Naturally, the estimation is restricted to the uncensored cohorts (2012-2014) which entails a sample size of slightly over 18,000 graduates.

The bottom of column 1 in Table 3 shows that the mean participation rate for the estimation sample is 1.04%, which is very close to the 1.07% reported in Table 1, but the unconditional gender gap is larger in the estimation sample (FM ratio 0.79) than what we reported in Table 1 (FM ratio 0.90) due to differences in the samples.[23] At any rate,

[21]The Table also shows that female UB graduates are 0.8 years younger than men and are 5.6 percentage-points less likely to have at least one college-educated parent. The latter finding is not surprising given that the average SES of men and women is the same in the population at large, but men are more highly selected among the college-going population (where they account for 1/3 of the student body).

[22]Within Social Sciences, the participation rate is substantially lower at the UB (0.33%) than at the other institutions (where it ranges between 1.31% and 2.17%). This within-field gap may reflect differences in major composition *within* Social Sciences, in addition to differences in selectivity or institutional support to students seeking to pursue graduate studies abroad.

[23]The discrepancy in the female-male participation ratio stems from two reasons. First, we lack information on parental education for 10% of the UB-FLC graduates. Second, some students have multiple majors and this introduces a discrepancy between the year of graduation in their UB records and in their FLC applications. In the estimation sample we chose to rely on the most recent year of graduation according to university records (whereas the aggregate participation rates in Table 1 employ the information in the FLC applications). In both cases, the gender gap increases relative to Table 1. When we use the information on graduation year from the FLC applications for the students with multiple majors, the resulting female-male participation ratio becomes 0.84, substantially closer to the 0.90 reported in Table 1. The remaining discrepancy is due to the individuals that lack information on

the female participation rate in the estimation sample is 0.25 percentage points lower than the corresponding value for males (with a standard error of 0.16).

When we include controls for field of study and individual characteristics (column 2), the estimates show large significant effects of parental education (positive) and age (negative). College-educated parents may be more effective in shaping the academic and professional aspirations of their children. In turn, age negatively affects the probability of participation because it captures when an individual student took longer to graduate (or experienced grade repetition in primary or secondary education).

Column 2 also reveals important differences across fields in the propensity to participate in the program. The highest participation rates are found in STEM, followed by Health, whereas the lowest value is found in Social Sciences.[24] We also learn that controlling for field of study reduces the gender gap in participation to 0.11 percentage points (down from 0.25) because of the female under-representation in STEM and over-representation in Social Sciences, respectively, the fields with the highest and lower application rates.

Column 3 controls for GPA. Not surprisingly, GPA has a positive and significant coefficient in regards to the probability to apply for a LCF fellowship. Moreover, controlling for GPA increases the gender gap to 0.39 percentage points. The increase is due to the fact that women have lower application rates despite having *higher* GPA than men on average (7.3 versus 7.1 in a 0-10 scale). Because the LCF program is highly selective, a disproportionate number of applications are made by the students with the highest GPA in each graduating cohort. This requires a more in-depth analysis of the role of GPA as a determinant of participation in the program. As we saw earlier, women were slightly under-represented in the top 5% of the GPA distribution of the corresponding majors.[25]

Because participation rates in the LCF program increase strongly with GPA, it is important to experiment with econometric specifications that control more flexibly for GPA. Accordingly, column 4 in Table 3 includes dummy variables for all the 5-point percentile brackets above the median, in addition to GPA.[26] Two observations stand out. First, only the coefficients for the brackets above the 85th percentile are signifi-

parental education.

[24]Based on our preferred specification (column 5), the mean participation rate is 2.7 percentage-points higher in STEM than in Social Sciences.

[25]This fact was established in the dataset pooling all universities. However, the same pattern is also found when we restrict to UB graduates (as seen in column 7 in Table C.17).

[26]Our specification is similar to the one used by Dillon and Stmith (2020) in their analysis of ability sorting in college choice and the resulting effect on student outcomes.

cantly different from zero, which shows that controlling for GPA in a linear fashion is a parsimonious parametric restriction across most of the GPA distribution, but is rejected by the data at the very top range of grades. Second, the gender gap falls slightly to 0.27 percentage-points because of the relatively lower share of women with very high grades.

Column 5 presents our preferred specification, which reduces the number of coefficients by allowing discontinuities in the GPA function only above the 85th percentile. This specification confirms the gender gap found in column 4 (estimated at 0.27 percentage points). The estimates also imply that students with grades in the top 5 percent of their major are 4.1 percentage-points more likely to apply to the fellowship.[27] The estimates imply that having two college-educated parents is associated with a 1.5 percentage-point increase in the probability to apply for the fellowships. Naturally, having a single parent with college education has a smaller effect, increasing the probability to participate in the program by 0.5 percentage points. Last, column 6 shows that the estimated gender gap is robust to a completely different way of modeling the relationship between GPA and the propensity to apply to the fellowships, namely, by considering a third degree polynomial.

Table 4 provides additional insights on the determinants of participation. First, the size of the gender gap increases as we restrict to increasingly higher GPA subsamples, both as a difference and in relationship to the mean participation rate (from $0.27/1.04 = 26\%$ for the whole sample to roughly 32% and 39% above the 75th or 90th GPA percentiles, respectively). Additionally, columns 4 and 5 estimate our preferred specification separately for STEM and non-STEM graduates. The estimates reveal no evidence of a gender gap in STEM (after controls and fixed-effects), where the coefficient on the female dummy is actually positive (0.59). Instead, the gender gap arises in the non-STEM fields.[28]

[27] The specification in column 5 implies the following relationship between GPA and probability to participate. Up to the 80th percentile of GPA, the participation probability increases linearly. Discontinuous jumps are allowed at the 80th, 85th, 90th and 95th percentiles, with magnitudes estimated by the corresponding coefficients (whose values are always in reference to the mean participation rate below the 80th percentile). Within each of the brackets 80-85th percentiles through 95-100th percentiles, the marginal effect of GPA is linear and has the same slope as the below the 80th percentile, which is captured by the coefficient of the GPA variable.

[28] Unfortunately, subdividing the non-STEM category by field (Health, Arts & Humanities and Social Sciences) is not informative due to small sample sizes (columns 6-8).

4.3 Conditional gender gap and factor contributions

Table 5 provides a decomposition of the contribution of each of the factors included in the regression model in explaining the gender gap in participation rates. Specifically, we evaluate the estimated (linear) model at the mean values of the covariates for male and female students, separately. The corresponding predicted participation rates are 1.24 (males) and 0.98 (females) for the whole sample (top panel). These values are very similar to the means in the data (1.20 and 0.95, respectively), and amounts to a female-male ratio of 0.79. That is to say, the predicted female participation is 21% lower than for males with the same characteristics.

Next, we experiment by assigning male distribution over fields of study to women. Because the male fields are more prone to participation in the fellowship program, the resulting gender gap shrinks to 9% (female-male ratio 0.91). We turn now to the contributions of age and socio-economic status. When we account for the fact that female graduates are almost one year younger than male graduates, the female-male ratio falls by 6 percentage points to 0.85 (because of the negative effect of age on application rates). In turn, when we assign women the higher SES of men, the female-male ratio increases back to 0.92, cancelling the effect of equalizing age. Last, when we assign the male mean GPA to female students, the female-male ratio now falls to 0.78 because of the lower mean GPA for males. When we assign women the male percentiles for the right-tail of the GPA distribution and the graduation cohorts of males, the predicted participation rate and corresponding female-male participation ratio barely budge. In other words, while it is true that women are slightly under-represented among top 5% grade earners, this makes practically no difference in a quantitative sense. The resulting female-male participation ratio is 0.78 (a 22% gender gap) after controlling for all observable characteristics. Importantly, it is obtained by comparing predicted participation rates at identical values of the covariates (fields of study, age, SES, GPA distribution and graduation cohorts) and therefore provides a credible estimate of the *conditional* gender gap in participation rates. As it turns out, the conditional and unconditional gender gaps in participation are practically identical. However, this masks two offsetting effects: women specialize in fields of study characterized by low participation rates in the fellowship program but have higher GPA, on average, than male students in the same field of study. Importantly, this finding is confirmed when restricting the analysis to the sample of high-GPA students (presented in the bottom panel of the Table).[29]

[29] We stress that this decomposition exercise is simply a helpful device to characterize the contribution

Summing up, our estimates reveal that many factors shape the decision to participate in the fellowship program. Among these, academic ability (measured by grades) and field of study play the key roles in a quantitative sense. The estimates in this section are also helpful in estimating the *conditional* female-male participation ratio, which is found to be around 0.78 (and 0.74 in the high-GPA sample). Thus, we conclude that female participation rates among UB graduates are 22% (26% in the high-GPA sample) lower than the corresponding value for males graduating with the same GPA, age, socio-economic status, and graduating in the same field of study, cohort and university. While we cannot be totally certain, it is likely that this conclusion applies to the other universities as well (with the exception of the engineering school).

5 Survey on interest in graduate studies

So far we have established the existence of large gender gaps in participation in the LCF program even after equalizing fields of study, grades, age, socio-economic status and year of graduation. To explore what accounts for the unexplained gender gap in program participation, we have carried out a large-scale survey of college students in the 4 universities in our study. To provide a bridge between the findings based on administrative records and the survey, we analyze students' *intentions* to apply for competitive fellowships.

5.1 Summary statistics

In January and February 2020, we surveyed all students at the four universities participating in our study that had completed over half of the credit hours required for graduation. We gathered almost 5,000 essentially complete questionnaires in total. The data show that the majority of students (70%) plan to work after completing the current degree. But, at the same time, 90% also consider studying another major (10%), a Master's degree (70%) or a PhD (10%).

Table 6 presents summary statistics, comparing men and women for the whole sample and the subsample with high grades.[30] The Table is largely in agreement with

of each factor to the gender gap in participation. Causal interpretations should not be made, particularly in regards to field of study, given the clear signs of self-selection in unobservables in the choice of field of study. We will return to this point later in the paper.

[30] We consider two thresholds: *GPAHigh1* are students with GPA above 7.8, which corresponds to the average 75th percentile across all fields of study in administrative data. Similarly, we define *GPAHigh2*

the administrative data discussed earlier. Specifically, female students are substantially under-represented in STEM majors (by 36 percentage-points). They are also slightly younger than men and have lower socio-economic status, measured by parental education. In regards to academic achievement, the survey data show that women have slightly higher GPA, on average, than men, but are 4 percentage-points less likely to be in the top 10% by GPA.[31]

Turning now to the main goal of the survey, we examine gender differences in the intention to pursue post-graduate studies. We find that female students are as interested in pursuing a Master's degree as men, but much less interested in doctoral studies (by 3.7 percentage points relative to a 12.2% mean), a pattern consistent with the administrative data in Table 2.[32] Given the lower interest in post-graduate studies among female college students, it is not surprising that women are 5 percentage-points less likely to know about the existence of (selective) graduate fellowship programs and 2.2 percentage-points less likely to apply.

5.2 Intended participation in selective fellowship programs

Next, we examine whether the observed gender gaps in fellowship applications documented in the previous sections are a reflection of gaps in *intentions* prior to graduation. In other words, do we see differences in intentions translate into differences in behavior?

We begin by examining knowledge acquisition regarding graduate fellowships, and whether there are gender gaps along this dimension.[33] Students' knowledge of the fellowships is a function of many factors, such as how widely advertised these programs are. However, acquiring information also depends on individuals' effort in seeking information, which in turn will be a function of one's interest in conducting graduate studies abroad. As reported in columns 1 and 2 of Table 7, about 67% of the students are aware of the existence of competitive fellowship programs (such as the LCF program), and

as an indicator for students with GPA above 8.3, which corresponds to the average 90th percentile across all fields of study.

[31] Almost 20% of our sample reports grades above the 90th percentile because high-achieving students are more likely to respond to the survey.

[32] Appendix Table C.19 provides a more detailed analysis of the determinants of interest in graduate studies, distinguishing between Master's degrees and doctoral studies. The results show that GPA plays a large and significant role and that interest is higher in some fields (STEM and Arts & Humanities) than others. In addition, after controlling for individual characteristics, women are 5 percentage-points less likely to plan to pursue doctoral studies.

[33] The question we ask lists the most well known graduate fellowship programs in Spain: LCF, Fulbright, Ramon Areces and Rafael del Pino.

this rate increases slightly among high GPA students (to 69%). However, we observe a gender gap in knowledge of 5.1 percentage-points (in column 1) that practically doubles for high GPA students.

We turn now to the intention to apply for a selective graduate fellowship. We consider as our dependent variable an indicator of the intention to apply for a competitive fellowship, which takes a value of zero if the student is not aware of the existence of the program or reports no intention to apply and focus our attention to students with high GPA because it is common knowledge that these fellowships are aimed at students with excellent academic credentials. For this population, the mean intended application rate is 16.4%, and women are 9.2 percentage-points less likely to apply than men (column 5). As we control for field of study (column 6), the gap shrinks but remains at 7.1 percentage points. The estimates in this column also reveal that GPA and SES are highly relevant determinants of the intention to apply for a fellowship. Last, columns 7 and 8 show that there is large gender gap in non-STEM fields, but no gap is present in STEM. In the latter, the point estimate for the female dummy variable is actually positive, suggesting that high-GPA women in STEM have slightly higher intended participation than comparable men.

In conclusion, our findings in this section show that around 11% of students intend to apply for the LCF fellowship (or a similar one), although fewer than 2% actually do.[34] The survey data uncover a large gender gap in the intention to apply, except in STEM where women are equally or more likely to intend to apply than men with similar characteristics. Remarkably, this pattern matches our estimates based on administrative data (Section 4).

6 Mechanism: family ties and geographic mobility

What accounts for the lower intention to apply for a graduate FLC fellowship among women relative to men with the same observable characteristics? The administrative data allowed us to quantify the role of several factors, such as GPA, field of study and socio-economic status, but a large gap remains after controlling for individual heterogeneity along these dimensions. We turn to the survey to shed some light on other factors that may be at play.

We consider several mechanisms but focus our attention on examining whether men

[34] The LCF fellowship program is, by far, the largest in size and the most well known to college students in the region of Catalonia.

and women differ in terms of their geographic mobility. Some recent studies provide evidence that commuting disproportionately penalizes women in the labor market and shapes their work decisions (Le Barbanchon et al. (2019), Fluchtmann et al. (2020) and Petrongolo and Ronchi (2020)). Thus, it is possible that distance also constraints the post-graduate educational investments of women. If, hypothetically, women were less willing to relocate to a distant city or to another country, this would reduce the set of graduate institutions they would consider and lead some of them to settle for less prestigious programs. Preliminary evidence along these lines was presented in Table 2. As we discussed earlier, the female-male ratio in applications aimed at doctoral studies abroad was much lower than for doctoral studies in Spain.

Our survey asked students for the geographic location of their preferred graduate program. As shown in Table 8 (column 1), the majority of male and female students (61%) reported preferring a program located in the province where they currently reside.[35] About 11% of male students (and 15% of women) were considering programs in another province within Spain. However, while 28% of men report planning to attend graduate school abroad, the corresponding value for women was 3.5 percentage points lower.

We also asked students about the location of their graduate program of choice in the event they did not face any *economic or family constraints*.[36] As shown in the second column of the table, students' 'unconstrained' choices would be dramatically different. The share of male students that would choose to study abroad would be 32 percentage-points higher. Among women, the increase in the corresponding share is 3 percentage-points higher than for men. In other words, the underlying preference for studying at a foreign institution seems to be the same for male and female students. However, economic or family restrictions constrain women more than men in terms of their geographical mobility.

In order to explore the role of family more deeply, we use the survey data to build several measures of family ties: weekly hours providing care for siblings or elderly relatives, currently being in a relationship or cohabitating (with a romantic partner). In the survey, around 55% of male and female students report providing care for siblings or older relatives (Table 6). However, conditional on providing care, women report spend-

[35] The universities included in our study are located in the Barcelona metropolitan, which offers plenty of options.

[36] The exact wording of the question is the following: "In the absence of family and economic constraints, your preference would be to carry out graduate studies in (i) the current province of residence, (ii) in another province (within Spain), (iii) in another European country, or (iv) outside of Europe.

ing more hours than men in this task. More specifically, the mean number of care hours is 10.7 per week (conditional on caregiving). But women providing care report spending 1.5 hours more, on average, than men. In addition, 40.8% and 7.4% of men report being in a relationship or cohabitating, respectively, while the corresponding rates for women are 9.5 and 1.8 percentage points higher. More formally, Table 9 confirms that the gender gaps in family ties are statistically significant in regression models that control for GPA, SES and field of study. In addition, the table shows that the gender gap in caregiving time is substantially smaller among STEM students relative to other fields.

Next, we conduct a more formal analysis of the factors that determine the decision to pursue graduate studies abroad, including measures of family ties. Specifically, we estimate linear probability models where the dependent variable is an indicator for preferring graduate studies abroad (on the sample of individuals interested in graduate school). The results are reported in Table 10. Column 1 shows that female students interested in graduate studies are 3.5 percentage-points less likely to want to study abroad. However, this gap vanishes once we control for GPA, socio-economic status and field fixed-effects (columns 2).

As noted above, a larger fraction of women report caring for relatives and being in relationships compared to men. Adding these control variables (column 3), we find that hours of care has a negative effect on the intention to study abroad. This also seems to be the case for being in a relationship and cohabiting, particularly for women (although the effects are imprecisely estimated. It is worth noting that the coefficient on the female dummy is now positive, suggesting that should women have the same family ties as men in the same field of study and with the same GPA and socio-economic background, their intention to study abroad might be 2.6 percentage-points higher than men's.

Columns 4 and 5 provide estimates separately for the STEM and non-STEM subsamples. Once again, we find divergent gender gaps between these two groups. In STEM women appear to be more likely than similar men to plan to study abroad, while the opposite is true in non-STEM fields. Last, columns 6 and 7 control for our measures of family ties. This reverses the gender gap in the preference to study abroad for non-STEM students, but leaves the estimate for STEM students unaffected, perhaps due to the smaller gender gap in family ties among STEM majors.

Summing up, women's lower reported interest in conducting graduate studies abroad appears to be related to the stronger bonds to their families, compared to men, which constrain their geographic mobility. In fact, our estimates suggest that, at equal family ties, the proportion of women who would like to study abroad would overtake the proportion of men. Because the LCF fellowships are aimed mainly at graduate studies abroad, family ties probably contribute to explain the abnormally low number of female applicants relative to male students with the same observable characteristics.[37]

7 STEM Women

STEM majors remain heavily dominated by males. In our administrative records (2009-2018), only 28.8% of the STEM majors are women, compared to a female share of 67.5% in non-STEM disciplines, and a similar gender gap is found in our survey data as well. The goal of this section is to explore the differences between female students in STEM majors and those in other majors a long a wide range of indicators, which may shed light on the nature of those differences.

Our findings are collected in Table 11, which contains three panels. The top panel presents data from our survey for graduates of the UB and the bottom panel uses our administrative data for the same university, containing an additional measure of academic ability not included in the survey (college-entry grades). The middle panel presents survey data for the four universities in our study, which is useful to assess if the findings apply to the other universities as well. Three main findings emerge from the table.

1. *STEM women have better socio-economic background than women in non-STEM fields.* As can be seen in the top panel of Table 11, the average female STEM student is 14-16 percentage points more likely to have a father or mother with a college degree than female students in non-STEM fields. The same pattern is also observed in the survey data for the other universities and in the administrative records for the UB, and has also been documented in Mourifie et al. (2020) for Germany and Canada.

2. *STEM women have higher educational and professional aspirations, and looser family ties, than women in non-STEM fields.* The top two panels of the table show that STEM women aspire to have 14-17% higher hourly wages than non-STEM women (column 3 in Table 11). The survey also suggests slightly looser family ties

[37] Other factors may also be at play. The survey data show that female college students have lower earnings aspirations than observationally similar men. In addition, the desired age to have the first child is lower among female students, relative to comparable men. Clearly, both factors may lower the payoff to ambitious educational investments, but we have been unable to establish that link. The results are available upon request.

(at least among UB students), though the evidence is not conclusive. Columns 8-10 turn to educational aspirations. Both at the UB and at the other universities, we find that STEM women are significantly more likely to plan to attend graduate school, particularly in regards to Master's degrees, and much more likely to plan to study abroad than female students in other fields. Last, column 10 suggests that female STEM graduates are equally or more likely to apply to competitive graduate fellowships than non-STEM women.

3. *STEM women have higher academic ability than women in non-STEM fields.* A simplistic comparison of graduation records shows that STEM women have lower GPA at graduation than women in non-STEM fields (column 4). However, this simply reflects the fact that average grades are lower in STEM fields. A more meaningful comparison is based on the university-access score (known as PAU in its Spanish acronym).[38] As shown in Table 11 (columns 11 and 12), STEM women have a slightly higher PAU score than non-STEM women (by 1.3% or 0.1 grade points), widening to 7% (or 0.5 grade points) when we restrict to high-GPA students.

In sum, STEM women differ systematically from other female college students along a variety of dimensions. In particular, their parents are more likely to be college-educated, they have higher earnings aspirations, higher academic ability (measured by college-entry grades) and are more interested in graduate studies abroad than female students in other majors. While parental background and grades were included as controls in our earlier models, earnings aspirations were not and could contribute to explain our earlier finding of higher educational aspirations (measured by participation in the LCF fellowships) for STEM women than for similar men in the same field of study.

8 Conclusions

Strong credentials, such as advanced degrees from renown universities, help gain access top positions in the private and public sectors (including academia). At the same time, it is well established that women remain under-represented in high-status positions in the labor market and also in prestigious graduate programs across many fields of study.

[38] The PAU score is an average of each student's high-school GPA and the grade obtained in a government-administered test. The PAU is the single measure used to allocate high-school graduates across majors and universities on the basis of their stated preferences.

Our paper analyzes the academic aspirations of male and female college graduates through a new lens: participation in a large fellowship program aimed at funding graduate students abroad for candidates with excellent academic qualifications. Obtaining these highly competitive fellowships provides funding for tuition (plus related expenses) and helps gain admission into the world's leading graduate programs. Using administrative records for 4 large universities in Spain, we document gender gaps in participation in the fellowship program. Female participation rates are typically lower than males' in non-STEM fields (where women are the majority), while the opposite seems to be true in STEM (where women are the minority gender).

Using survey data on students currently in college, we show large gender gaps in the *intention* to apply for the fellowships, establishing a bridge between intentions and actual behavior. Furthermore, we demonstrate that female college students have stronger family ties than similar men, reducing their geographic mobility and educational aspirations. As a result, highly talented female students are more likely to engage in more affordable graduate programs of shorter duration (e.g. Master's degrees) at local institutions, compared to similar men. This may help explain the gender gap in achievement in the labor market, particularly among the most academically talented individuals.

References

Altonji, Joseph G, Erica Blom, and Costas Meghir, "Heterogeneity in human capital investments: High school cur-riculum, college major, and careers," *Annu. Rev. Econ.*, 2012, *4* (1), 185–223.

Azmat, Ghazala and Rosa Ferrer, "Gender Gaps in Performance: Evidence from Young Lawyers," *Journal of Political Economy*, 2017, *125* (5), 1306–1355.

_ , Vicente Cuñat, and Emeric Henry, "Gender Promotion Gaps: Career Aspirations and Workplace Discrimination," 2020.

Bagues, Manuel F. and Berta Esteve-Volart, "Can Gender Parity Break the Glass Ceiling? Evidence from a Repeated Randomized Experiment," *Review of Economic Studies*, 2010, *77* (4), 1301–1328.

Barbanchon, Thomas Le, Roland Rathelot, and Alexandra Roulet, "Gender differences in job search: Trading off commute against wage," *Available at SSRN 3467750*, 2019.

Bayer, Amanda and Cecilia Elena Rouse, "Diversity in the Economics Profession: A New Attack on an Old Problem," *Journal of Economic Perspectives*, Fall 2016, *30* (4), 221–242.

Bertrand, Marianne, "Career, family, and the well-being of college-educated wom-en," *American Economic Review*, 2013, *103* (3), 244–50.

_ and Kevin F Hallock, "The gender gap in top corporate jobs," *ILR Review*, 2001, *55* (1), 3–21.

_ , Claudia Goldin, and Lawrence F Katz, "Dynamics of the gender gap for young professionals in the financial and corporate sectors," *American Economic Journal: Applied Economics*, 2010, *2* (3), 228–55.

_ , Sandra E Black, Sissel Jensen, and Adriana Lleras-Muney, "Breaking the Glass Ceiling? The Effect of Board Quotas on Female Labour Market Outcomes in Norway," *Review of Economic Studies*, 2019, *86* (1), 191–239.

Bielby, William T. and Denise D. Bielby, "I Will Follow Him: Family Ties, Gender-Role Beliefs, and Reluctance to Relocate for a Better Job," *American Journal of Sociology*, 1992, *97* (5), 1241–1267.

Black, Dan A, Amelia M Haviland, Seth G Sanders, and Lowell J Taylor, "Gender wage disparities among the highly educated," *Journal of human resources*, 2008, *43* (3), 630–659.

Blau, Francine D. and Lawrence M. Kahn, "The Gender Wage Gap: Extent, Trends, and Explanations," *Journal of Economic Literature*, September 2017, *55* (3), 789–865.

Boustan, Leah and Andrew Langan, "Variation in Women's Success across PhD Programs in Economics," *Journal of Economic Perspectives*, Winter 2019, *33* (1), 23–42.

Breda, Thomas and Son Thierry Ly, "Professors in Core Science Fields Are Not Always Biased against Women: Evidence from France," *American Economic Journal: Applied Economics*, October 2015, *7* (4), 53–75.

Bursztyn, Leonardo, Thomas Fujiwara, and Amanda Pallais, "'Acting Wife': Marriage Market Incentives and Labor Market Investments," *American Economic Review*, November 2017, *107* (11), 3288–3319.

Buser, Thomas, Muriel Niederle, and Hessel Oosterbeek, "Gender, competitiveness, and career choices," *The Quarterly Journal of Economics*, 2014, *129* (3), 1409–1447.

Carlana, Michela, "Implicit Stereotypes: Evidence from Teachers? Gender Bias*," *The Quarterly Journal of Economics*, 03 2019, *134* (3), 1163–1224.

Carpio, Lucia Del and Maria Guadalupe, "More Women in Tech? Evidence from a field experiment addressing social identity," Working Papers, INSEAD 2019.

Carrell, Scott E., Marianne E. Page, and James E. West, "Sex and Science: How Professor Gender Perpetuates the Gender Gap," *The Quarterly Journal of Economics*, 2010, *125* (3), 1101–1144.

Cortes, Patricia, Jessica Pan, and Anna Sjogren, "Reaching for the Top: Gender Differences in the Labor Market and Household Outcomes of Top Talent in Sweden and the United States," Technical Report 2020.

Dillon, Eleanor W. and Jeffrey A. Stmith, "The Consequences of Academic Match between Students and Colleges," *The Journal of Human Resources*, Summer 2020, *55* (3), 767–808.

Ellison, Glenn and Ashley Swanson, "The gender gap in secondary school mathematics at high achievement levels: Evidence from the American Mathematics Competi-tions," *Journal of Economic Perspectives*, 2010, *24* (2), 109–28.

Farré, Lídia and Francesc Ortega, "Selecting Talent: Gender Differences in Participation and Success in Competitive Selection Processes," IZA Discussion Papers 12530, Institute of Labor Economics (IZA) 2019.

Fluchtmann, Jonas, Anita M. Glenny, Nikolaj Harmon, and Jonas Maibom, "The Gender Application Gap: Do men and women apply for the same jobs?," Technical Report 2020.

Fortin, Nicole M., Philip Oreopoulos, and Shelley Phipps, "Leaving Boys Behind: Gender Disparities in High Academic Achievement," *Journal of Human Resources*, 2015, *50* (3), 549–579.

Garcia-Montalvo, Jose, *Impacto de las becas 'La Caixa' de posgrado en el extranjero*, Obra Social 'la Caixa', 2014.

Goldin, Claudia, Lawrence F Katz, and Ilyana Kuziemko, "The homecoming of American college women: The reversal of the college gender gap," *Journal of Economic perspectives*, 2006, *20* (4), 133–156.

Guiso, Luigi, Ferdinando Monte, Paola Sapienza, and Luigi Zingales, "Culture, gender, and math," *SCIENCE-NEW YORK THEN WASHINGTON-*, 2008, *320* (5880), 1164.

Hedges, LV and A Nowell, "Sex differences in mental test scores, variability, and numbers of high-scoring individuals," *Science*, 1995, *269* (5220), 41–45.

Hill, Catherine, Christianne Corbett, and Andresse St. Rose, "Why so few? Women in Science, Technology, Engineering and Mathematics," ISBN 978-1-879922-40-2, AAUW 2010.

Hospido, Laura and Carlos Sanz, "Gender Gaps in the Evaluation of Research: Evidence from Submissions to Economics Conferences," IZA Discussion Papers 12494, Institute of Labor Economics (IZA) July 2019.

_ , Luc Laeven, and Ana Lamo, "The gender promotion gap: evidence from central banking," 2019.

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

Hsin, Amy, "Hegemonic Gender Norms and the Gender Gap in Achievement: The Case of Asian Americans," *Sociological Science*, 2018, *5* (32), 752–774.

Jürges, Hendrik, "Gender ideology, division of housework, and the geographic mobility of families," *Review of Economics of the Household*, 2006, *4* (4), 299–323.

Keloharju, Matti, Samuli Knüpfer, and Joa cim Tåg, "What Prevents Women from Reaching the Top?," 2019.

Kirkeboen, Lars J., Edwin Leuven, and Magne Mogstad, "Field of Study, Earnings, and Self-Selection," *The Quarterly Journal of Economics*, 2016, *131* (3), 1057–1111.

Montalban, Jose and Almudena Sevilla, "The gender gap in student performance: the role of test-taking environment," Technical Report 2020.

Mourifie, Ismael, Marc Henry, and Romuald Meango, "Sharp Bounds and Testability of a Roy Model of STEM Major Choices," *Journal of Political Economy. Forthcoming*, 2020, *0* (0), 000–000.

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

Paredes, Valentina A., M. Daniele Paserman, and Francisco Pino, "Does Economics Make You Sexist?," NBER Working Papers 27070, National Bureau of Economic Research, Inc May 2020.

Petrongolo, Barbara and Maddalena Ronchi, "Gender gaps and the structure of local labor markets," *Labour Economics*, 2020, p. 101819.

Porter, Catherine and Danila Serra, "Gender differences in the choice of major: The importance of female role models," *American Economic Journal: - Applied Economics - Forthcoming*, 2020.

Rouse, Cecilia and Claudia Goldin, "Orchestrating Impartiality: The Impact of 'Blind' Auditions on Female Musicians," *American Economic Review*, September 2000, *90* (4), 715–741.

Shauman, Kimberlee A. and Yu Xie, "Geographic mobility of scientists: Sex differences and family constraints," *Demography*, 1996, *33* (4), 455–468.

Wallen, Aaron S., Michael W. Morris, Beth A. Devine, and Jackson G. Lu, "Understanding the MBA Gender Gap: Women Respond to Gender Norms by Reducing Public Assertiveness but Not Private Effort," *Personality and Social Psychology Bulletin*, 2017, *43* (8), 1150–1170. PMID: 28903718.

Table 1: Aggregate Participation Rates

Gender	Uni	Counts Applications	Counts Grads	PR All Fields	PR STEM	PR Health	PR ArtsHum	PR Soc. Sci.
All cohorts								
Both	All	1,530	139,298	1.10	1.25	1.94	1.54	0.58
Male	All	715	64,081	1.12	1.17	2.21	1.46	0.67
Fem.	All	815	75,217	1.08	1.44	1.85	1.59	0.54
Ratio F/M	All	1.14	1.17	0.97	1.23	0.84	1.09	0.80
Uncensored cohorts								
Both	All	815	49,107	1.66	1.84	3.41	2.20	0.87
Male	All	371	22,365	1.66	1.69	3.78	2.19	1.00
Fem.	All	444	26,742	1.66	2.22	3.28	2.21	0.80
Ratio F/M	All	1.20	1.20	1.00	1.31	0.87	1.01	0.80

Notes: The participation rates is the number of ever applicants (by field-university-gender cell) over the size of the corresponding graduating cohort, or analogous ratios at lower levels of aggregation. The top panel reports data using all cohorts. The bottom panel panel uses uncensored cohorts only (graduation in academic years 2012-2013 through 2014-2015). *All* refers to the four universities (UB, UAB, UPC and UPF) pooled together.

Figure 1: GPA distributions by gender. All universities pooled

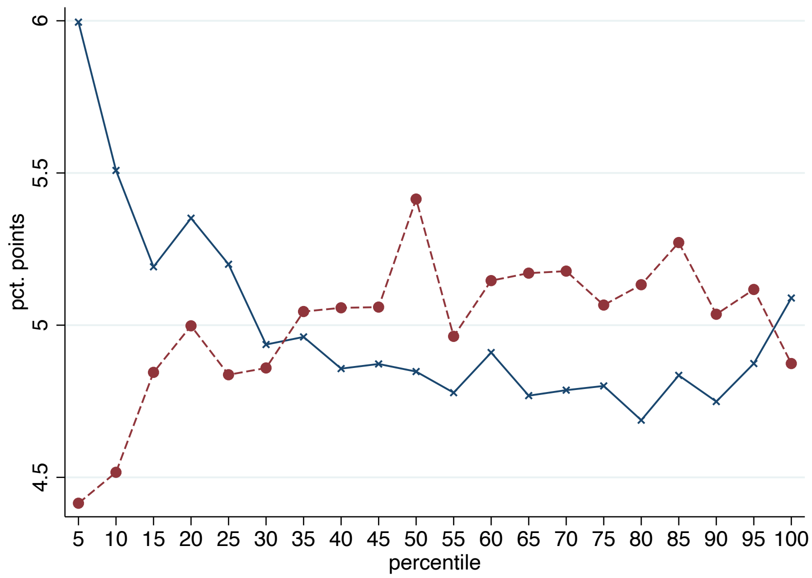
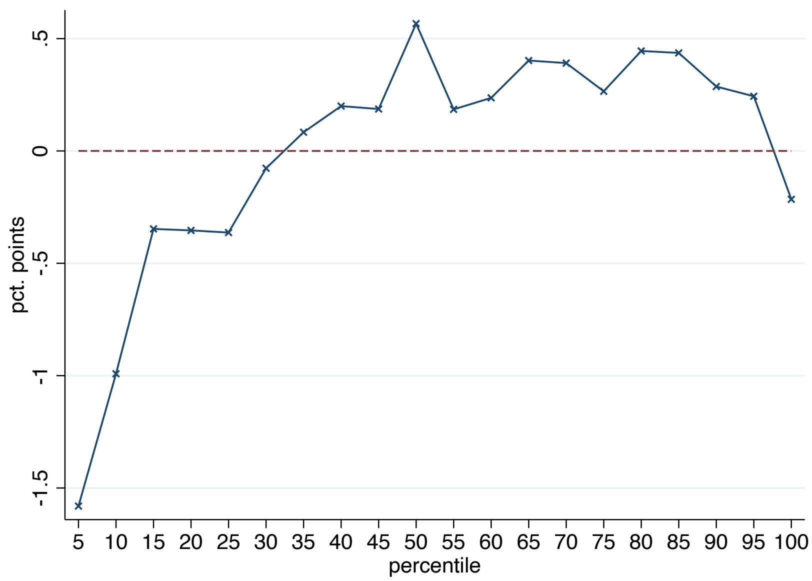


Figure 2: Difference in probability density female - male



Notes: Data for the 4 universities for graduation cohorts 2009-2018 for a total of 161,597 individual records. GPA at graduation administrative records. Each individual data point has been placed in the percentile bracket corresponding to the GPA distribution in its own major and university. The bottom figure plots the difference between the two lines plotted in the top figure (female minus male values).

Table 2: Participation Rates in Fellowship Program by type of degree and location of graduate program

Gender	All	Master Abroad	PhD Abroad	PhD Spain
Both	1.66	0.87	0.28	0.51
Male	1.66	0.85	0.31	0.49
Fem.	1.66	0.89	0.25	0.52
FM ratio	1.00	1.04	0.81	1.06

Notes: Pooled data for the 4 universities (all fields pooled) Uncensored cohorts only (2012-2014). LCF applications for rounds 2014-2018.

Table 3: Participation regressions. Matched UB-FLC Data

	(1)	(2)	(3)	(4)	(5)	(6)
Female	-0.25	-0.11	-0.39**	-0.27*	-0.27*	-0.25
	[0.16]	[0.16]	[0.16]	[0.16]	[0.16]	[0.16]
GPA			1.61***	0.90***	0.85***	120.28***
			[0.11]	[0.18]	[0.14]	[21.41]
GPA^2						-17.41***
						[2.87]
GPA^3						0.84***
						[0.13]
$85-90$ perc.				0.65	0.78**	
				[0.40]	[0.36]	
$90-95$ perc.				0.89**	1.02***	
				[0.43]	[0.38]	
>95 perc.				3.95***	4.11***	
				[0.48]	[0.41]	
Age		-0.07***	-0.08***	-0.08***	-0.08***	-0.08***
		[0.01]	[0.01]	[0.01]	[0.01]	[0.01]
College parent2		1.66***	1.64***	1.54***	1.55***	1.48***
		[0.20]	[0.20]	[0.20]	[0.20]	[0.19]
College parent1		0.44**	0.49**	0.47**	0.46**	0.40**
		[0.19]	[0.19]	[0.19]	[0.19]	[0.19]
STEM		2.72***	2.76***	2.73***	2.74***	2.66***
		[0.29]	[0.29]	[0.29]	[0.29]	[0.29]
Health		1.32***	0.90***	1.06***	1.08***	0.96***
		[0.20]	[0.20]	[0.20]	[0.20]	[0.20]
ArtsHum		1.15***	0.63***	0.87***	0.89***	0.61***
		[0.21]	[0.21]	[0.22]	[0.22]	[0.21]
Observations	18,195	18,195	18,195	18,195	18,195	18,195
Brackets 50p-80p	no	no	no	yes	no	no
Mean dep (%)	1.04	1.04	1.04	1.04	1.04	1.04
Mean Dv Male	1.20	1.20	1.20	1.20	1.20	1.20
Mean Dv Fem	0.95	0.95	0.95	0.95	0.95	0.95
FM ratio uncond.	0.79	0.79	0.79	0.79	0.79	0.79

Notes: The dependent variable is a dichotomous variable taking value of 0 or 100. The latter indicates that the individual applied to the FLC Fellowship program (in any year). The sample contains only uncensored cohorts corresponding to academic years 2012-2013 through 2014-2015 but about 5,000 observations lack socio-demographic observations (because they were transfer students). The GPA percentiles have been computed based on the administrative data for each major. Column 4 includes dummy variables for GPA percentile brackets 50-55 through 75-80, not shown for brevity (and none of those coefficients is statistically significant at 10%). Intercept and dummies for graduation cohort 2013 and 2014 included but not shown. The omitted categories are Soc. Sci. and graduation cohort 2012. The unconditional female-male (FM) ratio is simply the ratio of the mean PR for males to mean PR for females. Heteroskedasticity-robust standard errors. *** p<0.01, ** p<0.05, * p<0.1.

Table 4: Participation regressions. Subsamples

Sample	(1) All	(2) GPA>75p	(3) GPA>90p	(4) STEM	(5) Non-STEM	(6) Health	(7) ArtHum	(8) Soc
Female	-0.27*	-0.85	-1.73*	0.59	-0.23	-0.48	-0.16	-0.08
	[0.16]	[0.52]	[1.03]	[0.93]	[0.15]	[0.51]	[0.46]	[0.12]
GPA	0.85***	2.26***	4.39***	4.27***	0.80***	0.91**	1.23***	0.10
	[0.14]	[0.60]	[1.29]	[1.12]	[0.13]	[0.40]	[0.42]	[0.10]
$85-90$ perc.	0.78**	0.52		5.26**	0.28	1.14	2.36**	-0.14
	[0.36]	[0.65]		[2.57]	[0.34]	[1.08]	[1.13]	[0.26]
$90-95$ perc.	1.02***	0.30		1.55	0.68*	2.39**	1.85	0.09
	[0.38]	[0.70]		[2.59]	[0.36]	[1.09]	[1.14]	[0.28]
>95 perc.	4.11***	2.69***	1.31	11.03***	2.98***	7.30***	5.67***	1.22***
	[0.41]	[0.85]	[1.11]	[2.98]	[0.38]	[1.15]	[1.26]	[0.30]
Age	-0.08***	-0.16***	-0.23***	-0.18	-0.07***	-0.19***	-0.09***	-0.02**
	[0.01]	[0.04]	[0.07]	[0.18]	[0.01]	[0.05]	[0.03]	[0.01]
College parent1	0.46**	1.37**	1.58	-0.16	0.60***	0.46	1.06*	0.27*
	[0.19]	[0.60]	[1.24]	[1.15]	[0.18]	[0.57]	[0.55]	[0.14]
College parent2	1.55***	2.69***	4.40***	3.58***	1.44***	1.97***	1.42**	0.68***
	[0.20]	[0.59]	[1.17]	[1.11]	[0.18]	[0.53]	[0.57]	[0.15]
STEM	2.74***	7.72***	11.21***					
	[0.29]	[0.92]	[1.90]					
Health	1.08***	2.81***	4.18***					
	[0.20]	[0.63]	[1.25]					
ArtsHum	0.89***	2.66***	2.83*					
	[0.22]	[0.75]	[1.53]					
Obs.	18,195	4,557	1,801	1,415	16,780	3,640	2,929	10,211
Mean dep (%)	1.04	2.66	4.39	3.32	0.85	1.87	1.47	0.30
Mean Dv Male	1.20	3.35	5.82	3.30	0.89	2.11	1.61	0.35
Mean Dv Fem	0.95	2.32	3.64	3.31	0.83	1.79	1.39	0.28
Ratio F/M	0.79	0.69	0.63	1.00	0.93	0.85	0.86	0.80

Notes: The dependent variable is a dichotomous variable taking value of 0 or 100. The latter indicates that the individual applied to the La Caixa Fellowship program (in any year). The sample contains only uncensored cohorts corresponding to academic years 2012-2013, 2013-2014 and 2014-2015. The GPA percentiles have been computed based on the administrative data for each major. Intercept and dummies for graduation cohort 2013-2014 and 2015-2015 included but not shown. The omitted categories are Social Sciences and graduation cohort 2012-2013. Heteroskedasticity-robust standard errors. *** p<0.01, ** p<0.05, * p<0.1.

Table 5: Conditional gender gap in fellowship participation. Contributions

	(1) Males	(2) Females	(3) Fem - Male	(4) Fem/Male
Prediction Sample All GPA				
Mean probability	1.20	0.95	-0.25	0.79
Mean predicted probability	**1.24**	0.98	-0.26	0.79
Male Fields		1.12	-0.12	0.91
Male Fields & Age		1.05	-0.19	0.85
Male Fields, Age & SES		1.13	-0.10	0.92
Male Fields, Age , SES & GPA		0.97	-0.27	0.78
Male Fields, Age , SES & GPA & Perc.		0.96	-0.27	0.78
Male Fields, Age , SES & GPA & Perc. & Cohort		**0.97**	**-0.27**	**0.78**
Prediction Sample GPA>75p				
Mean probability	3.35	2.32	-1.03	0.69
Mean predicted probability	**3.33**	2.30	-1.03	0.69
Male Fields		2.92	-0.40	0.88
Male Fields & Age		2.69	-0.63	0.81
Male Fields, Age & SES		2.74	-0.59	0.82
Male Fields, Age , SES & GPA		2.44	-0.89	0.73
Male Fields, Age , SES & GPA & Perc.		2.48	-0.85	0.74
Male Fields, Age , SES & GPA & Perc. & Cohort		**2.48**	**-0.85**	**0.74**

Notes: The predictions are based on the estimates in Table 3 (column 5). We then evaluate the estimated model at the means of the covariates by gender (uncensored cohorts only). The top panel reports predictions for the whole sample (all GPA) and the bottom panel makes predictions only for high-GPA students (but uses the same estimated coefficients as the top panel). Gradually, we assign the mean male values of the covariates to females and examine how the gender gap (female-male predicted participation) evolves. *SES* stands for socio-economic status as is measured as the number of college-educated parents. In terms of GPA, we first equalize mean GPA and then the shares of graduates in the top brackets of the GPA percentile distribution (Perc.). The conditional female-male participation ratio is shown in bold in column 4. It is the ratio of the predicted participation rate for females over males, both of which have been evaluated at the mean values of the male subsample.

Table 6: Survey. Descriptive statistics. All universities

Sample Gender	All Men	All Fem - Male	GPAhigh1 Men	GPAhigh1 Fem - Male
STEM	0.585	-0.361***	0.485	-0.338***
Health	0.092	0.142***	0.135	0.093***
ArtsHum	0.064	0.108***	0.097	0.155***
Soc. Sci	0.236	0.123***	0.204	0.131***
Age	23.815	-0.324***	23.765	-0.235
SES High Edu	0.659	-0.035***	0.731	-0.011
Fam. Income (thousands)	50.428	-0.158	50.843	-2.083
GPA	7.057	0.188***	8.464	-0.121***
GPA high1	0.283	0.008	1.000	0
GPA high2	0.192	-0.041***	0.715	-0.145***
N. children	0.032	-0.007	0.042	0.008
Care indicator	0.556	-0.011	0.511	0.026
Care hours	5.474	0.722***	4.641	1.487***
Relationship	0.408	0.095***	0.361	0.130***
Cohabitates	0.074	0.018**	0.074	0.024
Plans Master	0.686	0.011	0.719	-0.013
Plans PhD	0.122	-0.037***	0.234	-0.099***
Asp. Hourly wage	31.661	-9.992***	25.510	-4.018*
Asp. N. children	1.844	0.028	1.822	0.02
Asp. Age first child	31.085	-1.264***	31.247	-1.251***
Knows Fellowships	0.705	-5.157***	0.752	-9.874***
Intention Apply	0.120	-2.213**	0.214	-8.249***

Notes: The number of observations with non-missing data for gender, field of study and GPA is 4,848. Some of the variables above have missing observations. *GPAHigh1* restricts to students with GPA above 7.8 (in a 0-10 scale), which corresponds to the average 75th percentile across all fields of study in the administrative data when pooling all universities. In this case the sample size falls to 1,093 respondents (i.e. 22.5% of the 4,848 full sample). Similarly, we define *GPAHigh2* as an indicator for students with GPA above 8.3, which corresponds to the average 90th percentile. Caretaking hours are zero for 45% of the sample. The care hours reported include individuals providing zero hours. The *Plans* variables (referring to Master's degree and PhD) are not mutually exclusive. The mean caretaking hours, conditional on a positive value, is 10.7 hours per week. *Asp.* refers to aspirations. Monetary values are in Euros. Heteroskedasticity-robust standard errors. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 7: Survey. Intention to apply to competitive fellowship programs.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. Var.	Know	Know	Apply	Apply	Apply	Apply	Apply	Apply
Female	-5.16***	-9.87***	-2.21**	-8.25***	-9.20***	-7.15**	2.51	-9.95***
	[1.46]	[2.90]	[0.99]	[2.49]	[2.64]	[2.85]	[6.80]	[3.14]
GPA						8.55***	14.84**	5.21
						[3.26]	[6.16]	[3.80]
SES high						5.07**	1.23	6.64**
						[2.44]	[6.21]	[2.59]
STEM						6.52		
						[5.09]		
Life						-7.70**		-7.28**
						[3.02]		[3.04]
ArtsHum						-5.53*		-5.10
						[3.30]		[3.30]
Obs.	4,159	1,012	4,117	999	999	942	213	729
Sample	All	GPA high	All	GPA high	GPA high	GPA high	STEM	NoSTEM
Mean dv.	67.47	69.17	10.68	16.42	16.42	13.69	20.7	14.8

Notes: The dependent variable is dichotomous and takes a value of 100 for individuals that know of the fellowships (columns 1 and 2) or intend to apply to a competitive fellowship program (columns 3-8), and a value of zero for individuals that do not know those programs exist or do not intend to apply to any of them. Except in columns 1 and 3, the sample contains only individuals with GPA above 7.8 because of the highly competitive nature of these fellowship programs. University fixed-effects and intercept included in the models, but not shown in the Table. Column 7 restricts to STEM students and column 8 to non-STEM (with high GPA). Omitted categories are Social Sciences and UB graduate. *High SES* is an indicator taking value one for students with at least one parent with a college degree. Heteroskedasticity-robust standard errors. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 8: Survey. Preferred post-graduate program abroad.

	(1) Actual/Constrained Percent	(2) Unconstrained Percent	(3) Actual - Uncons. Percent
Males			
My province	61.3	32.8	-28.6
Other province	11.0	7.2	-3.8
Abroad	27.7	60.1	32.4
Europe	21.2	35.6	14.4
Outside Europe	6.5	24.5	18.0
sum	100	100	0
obs.	1,109	1,029	
Females			
My province	60.7	31.6	-29.1
Other province	15.1	8.7	-6.4
Abroad	24.2	59.7	35.6
Europe	19.4	37.8	18.4
Outside Europe	4.8	22.0	17.2
sum	100	100	0
obs.	1,554	1,458	

Notes: The sample is only those students who intend to pursue graduate studies. Heteroskedasticity-robust standard errors. *** p<0.01, ** p<0.05, * p<0.1

Table 9: Survey. Family ties

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dep. Var.	Care hours	Relationship	Cohabitates	Care hours	Relationship	Cohabitates	Care hours	Relationship	Cohabitates
Female	0.65**	10.30***	2.13**	0.23	10.25***	2.44	0.92***	10.55***	2.09*
	[0.27]	[1.73]	[0.95]	[0.45]	[2.86]	[1.60]	[0.35]	[2.18]	[1.18]
GPA	-0.57***	-1.35	-1.06*	-0.63***	-2.96*	-1.47*	-0.54**	-0.51	-0.83
	[0.16]	[0.94]	[0.58]	[0.22]	[1.51]	[0.81]	[0.22]	[1.21]	[0.80]
SES	-0.03	-2.82*	-4.44***	-0.63	-6.96**	-7.06***	0.31	-0.54	-3.04***
	[0.27]	[1.61]	[0.91]	[0.44]	[2.73]	[1.59]	[0.34]	[2.00]	[1.10]
STEM	-0.70*	-2.60	-3.04**						
	[0.41]	[2.65]	[1.36]						
Health	-0.57	0.64	-2.57*				-0.69*	0.03	-2.85**
	[0.40]	[2.44]	[1.35]				[0.40]	[2.47]	[1.38]
ArtsHum	0.30	-5.54**	1.50				0.24	-5.68**	1.46
	[0.46]	[2.56]	[1.51]				[0.47]	[2.57]	[1.51]
Obs.	3,737	3,921	3,921	1,318	1,368	1,368	2,419	2,553	2,553
Sample	Al	All	All	STEM	STEM	STEM	No STEM	No STEM	No STEM
mean dv	5.9	46.4	8.4	5.6	44.7	8.2	6.0	47.4	8.6

Notes: University fixed-effects and intercept included in the models, but not shown in the Table. *Care hours* is the number of hours per week taking care of relatives (children, grandparents, etc.). *Relationship* and *Cohabitates* are dichotomous variables taking a value of 100 if the individual reports being in a relationship or cohabitating (and zero otherwise), respectively. Heteroskedasticity-robust standard errors. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 10: Survey. Graduate study abroad.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dep. Var.	Abroad	Abroad	Abroad	Abroad	Abroad	Abroad	Abroad
Female	-3.49**	-0.56	2.63	4.72	-4.08*	4.34	1.30
	[1.73]	[1.86]	[2.70]	[3.13]	[2.29]	[4.88]	[3.36]
GPA		4.88***	4.99***	8.19***	2.65**	8.98***	2.30*
		[1.03]	[1.11]	[1.67]	[1.28]	[1.83]	[1.38]
SES high		12.42***	14.19***	12.92***	12.32***	15.57***	13.47***
		[1.66]	[1.82]	[3.05]	[1.96]	[3.32]	[2.15]
Care hours			-0.20*			-0.26	-0.15
			[0.10]			[0.21]	[0.12]
Relationship			-4.78			-5.03	-3.62
			[3.15]			[4.36]	[4.54]
Fem × Relationship			-6.83*			-3.71	-8.39
			[3.95]			[7.20]	[5.23]
Cohabitate			0.94			-2.23	6.02
			[6.17]			[8.30]	[9.10]
Fem × Cohab			-2.60			-0.48	-7.39
			[7.07]			[12.90]	[9.74]
STEM		22.04***	20.47***				
		[3.08]	[3.22]				
Health		-4.48*	-5.81**		-3.25		-4.57*
		[2.31]	[2.46]		[2.32]		[2.48]
Hum		4.33	2.86		5.18*		3.83
		[2.67]	[2.91]		[2.68]		[2.94]
Observations	2,663	2,632	2,185	964	1,668	780	1,405
Sample	Grad	Grad	Grad	STEM	No-STEM	STEM	No-STEM
Mean dv	25.6	25.6	25.6	31.96	21.88	31.96	21.88

Notes: The sample is only those students who intend to pursue graduate studies. Dependent variables are dichotomous indicators taking values of 0 or 100. University fixed-effects included in all specifications (but not shown). Omitted categories are Social Sciences and UB graduate. *High SES* is an indicator taking value one for students with at least one parent with a college degree. *GPA high* identifies students with GPA of 7.8 or higher. Heteroskedasticity-robust standard errors. *** p<0.01, ** p<0.05, * p<0.1

Table 11: STEM Women. Survey and UB-FLC data

	(1) Coll Fa	(2) Coll Mo	(3) Asp lwh	(4) GPA	(5) Care hours	(6) Relation.	(7) Cohabit.	(8) Master	(9) PhD	(10) Abroad	(11) PAU	(12) PAU 75p	(13) Apply
UB only													
STEM	0.14***	0.16***	0.14**	-0.30***	-0.04	-0.11**	-0.04	0.16***	0.06*	0.29***			0.04
	[0.05]	[0.05]	[0.05]	[0.07]	[0.86]	[0.05]	[0.03]	[0.04]	[0.03]	[0.06]			[0.03]
Obs.	711	724	746	881	674	733	733	869	869	621			755
Mean dv	0.32	0.37	2.71	7.32	6.70	0.51	0.11	0.69	0.10	0.19			0.08
All Uni													
STEM	0.10***	0.10***	0.17***	-0.50***	-0.48	0.00	0.00	0.05**	0.01	0.15***			0.00
	[0.03]	[0.03]	[0.03]	[0.04]	[0.42]	[0.03]	[0.01]	[0.02]	[0.01]	[0.03]			[0.01]
Obs.	2,267	2,295	2,366	2,817	2,182	2,309	2,309	2,765	2,765	1,895			2,403
Mean dv	0.35	0.41	2.74	7.25	6.19	0.50	0.09	0.70	0.09	0.24			0.10
UB-FLC													
STEM	0.07***	0.08***		-0.30***							0.09	0.53***	
	[0.02]	[0.02]		[0.03]							[0.05]	[0.13]	
Obs.	12,126	12,128		15,130							12,228	3,063	
Mean dv	0.29	0.28		7.27							7.13	7.56	

Notes: Only women included in the samples used in this Table. Data from the Survey or UB-LCF matched dataset. Top panels report survey data for UB students only, or for students of all universities included in the survey. The bottom panel reports administrative data for UB graduates (UB-FLC dataset). *CoGrad Father* and *CoGrad Mother* are indicators for a college-educated father and mother, respectively. GPA is the average grade of the student (on a 0-10 scale) and PAU is the score at the university-access test (PAU in its Spanish acronym), also on a 0-10 scale. The last column restricts to students with GPA above the 75th percentile of the major grade distribution. Heteroskedasticity-robust standard errors. *** p<0.01, ** p<0.05, * p<0.1

Appendix

A Uncensored cohorts

A.1 Aggregate participation rates for all universities

Our computation of *aggregate* participation rates in the LCF program proceeds in two steps. First, we tally the number of applicants of a given gender g that graduated in year t in field of study f from university $u(Applicants_{g,f,u,t})$ and normalize it by the number of graduates at the same level of aggregation $(Graduates_{g,f,u,t})$.[39] That is,

$$PR_{g,f,u,t} = \frac{Applicants_{g,f,u,t}}{Graduates_{g,f,u,t}}. \tag{2}$$

Clearly, we can compute participation rates at a more aggregated level, adding across years, universities or fields of study. The results are collected in Table C.12. The top panel reports the aggregate participation rates obtained when using all graduation cohorts (2009-2018). The first column shows that the FLC program received 1,530 applications (between 2014 and 2018) by graduates from the 4 universities in our study, which accounts for slightly less than one fifth of all the applications they received over that period.[40] The Table also shows that the number of female applicants was 14% higher than the number of male applicants. However, as shown in column 2, the number of female graduates was 17% higher than the number of male graduates. Column 3 reports the participation rate in the fellowship program for all fields pooled together. We estimate that 1.10% of the graduates applied for a FLC fellowship.

The second step in the computation of the aggregate participation rates addresses a censoring problem. Our dataset only contains applications to the LCF fellowship program for years 2014-2018. As shown in Table C.15 (for the graduates of the University of Barcelona), the participation rate for graduation cohort 2009-2010 was 0.23%. The rate increased steadily to a peak value of 1.11% for graduation cohort 2013-2014, gradually falling after that until reaching a value of zero for graduation cohort 2018-2019. Thus, the 1.10% participation rate estimated above suffers from a potentially severe censoring

[39] Clearly, the LCF applicants in year t may have graduated in a previous year. We examine this issue further below. For now, it suffices to point out that the overall number of graduates in a given university and field of study is fairly constant over our sample period. Thus, the normalization we are applying will not be far from the true participation rate.

[40] We are only counting complete applications received by the deadline. Many more applications were initiated but were not completed in time or were left incomplete and were not reviewed.

problem. Put simply, the bulk of the applications for the earlier graduation cohorts took place prior to 2014 whereas the applications for the most recent cohorts took place in 2019 or beyond.[41] As we explain in detail in **??**, only graduation cohorts 2012-2014 can be considered uncensored.[42]

A.2 Uncensored cohorts in the UB-LCF dataset

Let us consider first the whole dataset, which contains UB graduation cohorts 2009-2018 and applications data for 2014-2018. It is important to note that a requirement to apply for an LCF fellowship is to have graduated in the year of application or earlier. That is, a student graduating in academic year 2013-2014 (which we refer to as graduation cohort 2013) can first apply to the program in 2014 and any year after that. Additionally, the interest to pursue graduate studies typically fades away a few years after graduating from college. Thus, in our dataset, the participation decisions of many graduation cohorts are severely censored.

Let us now examine our matched data from a longitudinal perspective. It is helpful to begin by focusing on graduation cohort 2013 (whose last academic year was 2013-2014) because for these graduates we are able to follow their application decisions over the full period of applications data (2014-2018). The data show that, among the 7,593 graduates in the 2013 cohort at the UB, only 84 applied to the LCF fellowship program, implying a participation rate of 1.11%.

As shown in Table C.13, 21.4% of the applicants participated in the program in the year of graduation, that is, students graduating in academic year 2013-2014 applied to round 2014 of the fellowship. In fact, they were much more likely to apply one year after graduation (39.3%). Application rates fell sharply two years after graduation (10.7%) and hovered around that level for the next two years. Presumably, applications for this cohort gradually converged toward zero after 2018, but this is not observable within our data. Partly to examine this issue, but also to increase sample size, we widen our analysis to include the two adjacent cohorts 2012 and 2014. These two cohorts suffer from one additional year of censoring (relative to cohort 2013) but they allow us to triple

[41] In particular, individuals graduating in academic year 2018-2019 are not eligible for LCF fellowships until year 2019, which is not part of our applications dataset, which explains the zero participation rate in Table C.15 for this graduation cohort.

[42] The year refers to the beginning of the senior academic year. That is, the uncensored cohorts graduated in academic years 2012-2013, 2013-2014 and 2014-2015. As we discuss in detail in the Appendix, these cohorts may also suffer from a certain degree of censoring, but it is likely to be very small.

the sample size and, in addition, provide a 6-year window on the applications data.[43]

Table C.14 summarizes the years to application for cohorts 2012-2014. First, we note that fewer than 3% of the applications are submitted 5 years after graduation. This low estimate is partly due to the censoring for the cohorts 2012 and 2013, but also suggests the tapering off of applications from the end of our data window. The larger sample size provides a clearer picture of the time profile of applications. The data show that 16% of the applications take place in the year of graduation T, 38% one year later $(T+1)$, 19% in $T+2$, 15% in $T+3$, 10% in $T+4$ and 3% in $T+5$. That is, from second year after graduation onward, the participation rate appears to fall by roughly 5 percentage points per year, suggesting that the degree of data censoring is small for these cohorts. For the sake of simplicity, we refer to graduation cohorts 2012-2014 as *uncensored cohorts*.

Table C.14 also provides a disaggregation by field of study that shows an important distinction: applicants in Arts & Humanities and Social Sciences tend to delay their applications much more than applicants in STEM and Health. In the former two fields the median years to application is 2-3 whereas in the latter two the median is only 1 year.

B Detailed analysis of gender gaps in GPA distributions

B.1 Regression-based estimates

A more formal comparison is presented in Table C.17 where we estimate the gender gap in GPA using linear regression models that control for major, graduation year, university and field of study. The first column confirms the slightly higher GPA for females (0.11 points in a 0-10 scale). Columns 2 and 3 confirm that women are over-represented in the top half and top quarter of the grades distribution of their majors. However, they are under-represented in the top 5% and top 2% (columns 5 and 6) by around 0.25 percentage points. Column 7 in Table C.17 shows that the gender gap for students in the top 5% is essentially the same if we restrict to the UB.[44] Last, column 8 shows that the gap is larger (at 0.4 percentage points) when we exclude the engineering school from the sample.

[43] The number of applications for these three cohorts combined is 248, up from 84 for cohort 2013 alone.

[44] Figure C.4 and Figure C.5 in the Appendix also illustrate this point.

B.2 Alternative method to measure gender gaps along the GPA distribution: standardized GPA

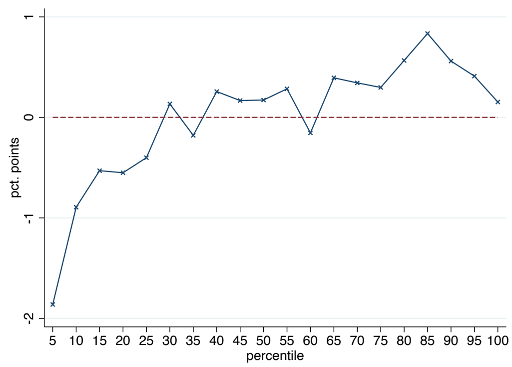
In our main analysis, we characterized the percentiles of the GPA distribution for each university and major. We then pooled individuals on the basis of their position in the corresponding percentiles.

As an alternative ,we now compare the GPA distributions of men and women using a different approach. Pooling observations after standardizing would be an approach if GPA distributions of college graduates were Gaussian. In that case, the resulting pooled distribution would also be a standard normal. However, the data shows that these distributions are not symmetric (and hence, not normally distributed). They tend to bunch slightly over the passing grade (5/10) and exhibit right-skewness. Thus the previous method is preferable as it does not make any distributional assumptions. We find:

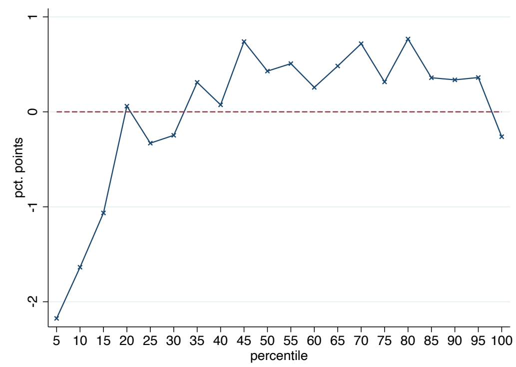
- We start from the raw GPA for each student. As shown in the top panel of Table C.18, average grades are lower in STEM (6.79 on a 0-10 scale) than in other disciplines (ranging from 6.99 in Social Sciences and 7.47 in Health and Life Sciences). In addition, on average women's GPA is 4% higher than men's, although the gender gap is practically non-existing in all fields except for Social Sciences.
- Next, we standardize each student's GPA using the mean and standard deviation of the corresponding major and university, and pool all observations. The resulting data again show that women, on average, have higher GPA than men (by about 0.06 standard deviations). This is the case in all disciplines, except in Arts & Humanities where women have lower grades, on average than men.
- Between the 90th and 95th percentiles, women are slightly over-represented. However, above the 95th percentile they are under-represented. The gap is larger than obtained with our previous method. We now estimate a gap of 0.72 percentage points in this bracket, compared to 0.26 percentage points. However, as discussed above, this method is less reliable for non-normal distributions (like ours).

C Tables and Figures

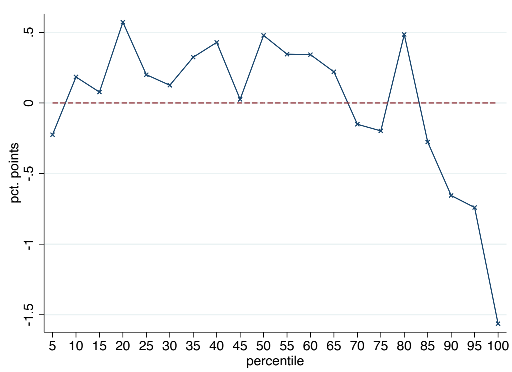
Figure C.3: Difference in GPA probability distribution (in percentage points) by gender and field. All universities pooled.



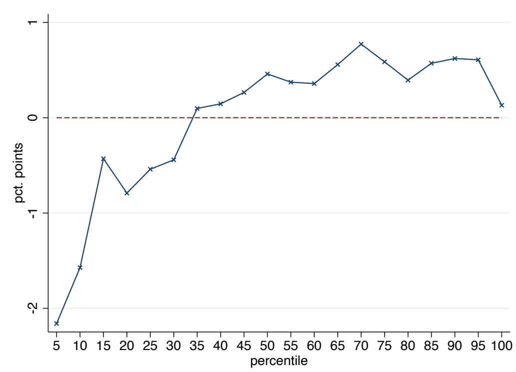
((a)) STEM



((b)) Health



((c)) Arts & Humanities



((d)) Social Sciences

Notes: Data for the 4 universities for graduation cohorts 2009-2018 for a total of 161,597 individual records. GPA at graduation administrative records. Each individual data point has been placed in the percentile bracket corresponding to the GPA distribution in its own major and university. STEM panel plots data for the university specialized in engineering and technology.

Figure C.4: GPA distributions by gender. UB only

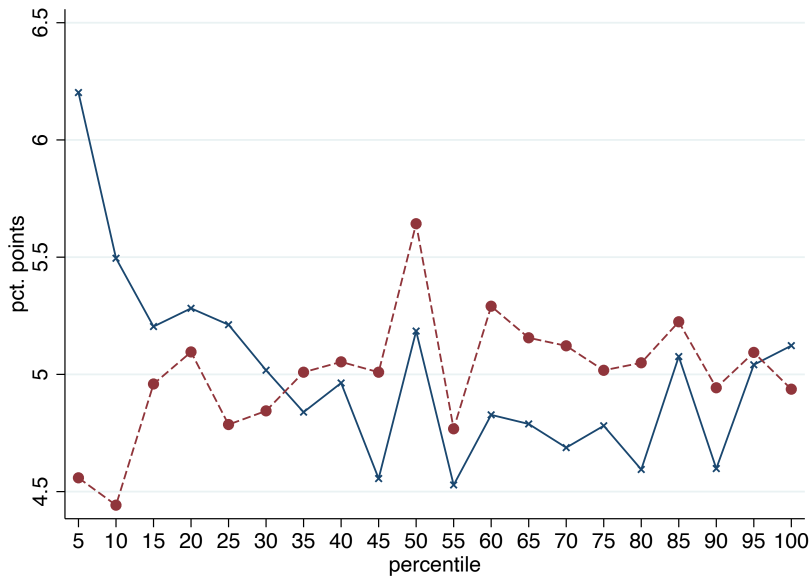
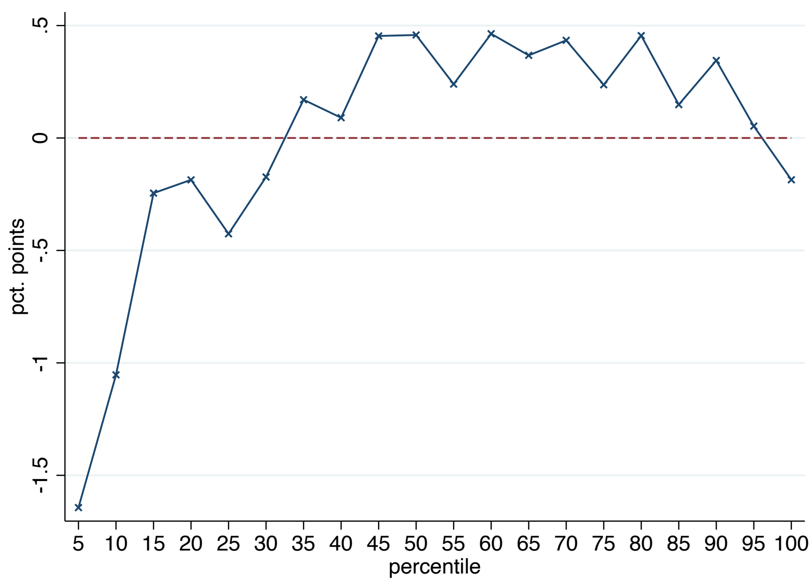


Figure C.5: Difference in probability density female - male



Notes: UB only, graduation cohorts 2009 (academic year 2009-2010) through 2018 (academic year 2018-2019). Based on individual records for 75,478 graduates. GPA at graduation administrative records. Each individual data point has been placed in the percentile bracket corresponding to the GPA distribution in its own major. The data for the UB contains 73 majors.

Table C.12: Aggregate Participation Rates by University

Gender	Uni	Counts Applications	Counts Grads	PR All Fields	PR STEM	PR Health	PR ArtsHum	PR Soc. Sci.
All cohorts								
Both	All	1,530	139,298	1.10	1.25	1.94	1.54	0.58
Male	All	715	64,081	1.12	1.17	2.21	1.46	0.67
Fem.	All	815	75,217	1.08	1.44	1.85	1.59	0.54
Ratio F/M	All	1.14	1.17	0.97	1.23	0.84	1.09	0.80
Uncensored cohorts								
Both	All	815	49,107	1.66	1.84	3.41	2.20	0.87
Male	All	371	22,365	1.66	1.69	3.78	2.19	1.00
Fem.	All	444	26,742	1.66	2.22	3.28	2.21	0.80
Ratio F/M	All	1.20	1.20	1.00	1.31	0.87	1.01	0.80
Both	UB	249	23,217	1.07	2.75	2.30	1.41	0.33
Male	UB	93	8,087	1.15	2.43	2.86	1.53	0.33
Fem.	UB	156	15,130	1.03	3.22	2.11	1.35	0.34
Ratio F/M	UB	1.68	1.87	0.90	1.32	0.74	0.88	1.03
Both	Uni2	206	6,984	2.95	6.12	6.03	2.86	1.31
Male	Uni2	76	2,400	3.17	4.75	5.61	2.91	1.67
Fem.	Uni2	130	4,584	2.84	8.94	6.22	2.84	1.16
Ratio F/M	Uni2	1.71	1.91	0.90	1.88	1.11	0.97	0.70
Both	Uni3	173	5,517	3.14	5.71	10.38	4.79	2.17
Male	Uni3	68	2,132	3.19	4.61	8.57	5.52	2.47
Fem.	Uni3	105	3,385	3.10	11.76	11.05	4.60	1.97
Ratio F/M	Uni3	1.54	1.59	0.97	2.55	1.29	0.83	0.80
Both	Uni4	187	13,389	1.40	1.40			
Male	Uni4	134	9,746	1.37	1.37			
Fem.	Uni4	53	3,643	1.45	1.45			
Ratio F/M	Uni4	0.40	0.37	1.06	1.06			

Notes: The participation rates is the number of ever applicants (by field-university-gender cell) over the size of the corresponding graduating cohort, or analogous ratios at lower levels of aggregation. The middle and bottom panel report data for uncensored cohorts only (graduation in academic years 2012-2013 through 2014-2015). *All* refers to the four universities (UB, UAB, UPC and UPF) pooled together. Except for the UB, the other universities are renamed to preserve confidentiality. The university specialized in engineering has a few majors in Social Sciences but they are very small in terms of enrollment. In fact, our data contain only one applicant to the fellowship program from these majors, which is insufficient to estimate participation rates with any degree of confidence.

Table C.13: Years to application. UB-FLC. Ideal cohort

	All Fields	STEM	Health	ArtsHum	Social Sci
Obs.	84	21	34	19	10
0	21.4	28.6	29.4	10.5	0.0
1	39.3	57.1	50.0	15.8	10.0
2	10.7	9.5	5.9	21.1	10.0
3	13.1	0.0	5.9	15.8	60.0
4	15.5	4.8	8.8	36.8	20.0
Median	1.0	1.0	1.0	3.0	3.0
Mean	1.83	0.95	1.15	2.53	2.90

Table C.14: Years to application. UB-FLC. Uncensored cohorts

	All Fields	STEM	Health	ArtsHum	Social Sci
Obs.	248	65	97	50	36
0	15.7	21.5	17.5	12.0	5.6
1	37.5	43.1	44.3	30.0	19.4
2	19.4	23.1	17.5	18.0	19.4
3	14.9	7.7	11.3	20.0	30.6
4	9.7	3.1	5.2	18.0	22.2
5	2.8	1.5	4.1	2.0	2.8
Median	1.0	1.0	1.0	2.0	3.0
Mean	1.83	1.30	1.60	2.10	2.50

Notes: Matched UB-FLC Fellowship. The UB data contains graduating cohorts 2009-2018. Note that graduating cohort 2013 refers to students that graduated in academic year 2013-2014 (ideal cohort). These students were eligible to apply to the fellowship rounds 2014 onward. The LaCaixa data contains rounds 2014-2018. Uncensored cohorts are the graduating cohorts 2012-2014.

Table C.15: Female shares and participation rates. UB-FLC matched data.

Graduation cohort	Graduates	Share females	PR
2009	7,925	0.69	0.23
2010	8,593	0.68	0.34
2011	7,104	0.64	0.68
2012	8,096	0.65	1.05
2013	7,593	0.65	1.11
2014	7,528	0.65	1.05
2015	7,266	0.67	1.02
2016	7,383	0.65	0.64
2017	7,017	0.64	0.51
2018	7,091	0.66	0
2009-2018	75,596	0.66	0.66
2012-2014	23,217	0.65	1.07

Notes: Matched UB-FLC Fellowship. The UB data contains graduating cohorts 2009-2018. The FLC data contain rounds 2014-2018. Uncensored cohorts are the graduating cohorts 2012-2014.

Table C.16: Summary statistics. UB-FLC. All cohorts

	All	Male	Fem-Male
Female	66.39		
GPA	7.24	7.10	0.21***
90p <GPA< 95p	5.19	5.17	0.03
GPA > 95p	5.21	5.37	-0.23
Age	24.28	24.81	-0.80***
One College parent	20.85	21.80	-1.43***
Two College parent	19.90	22.68	-4.20***
STEM	8.27	14.50	-9.38***
Health	21.31	15.64	8.54***
Social	55.25	54.09	1.75***
ArtsHum	15.00	15.60	-0.90***
Observations	63,701		

Notes: Matched UB-FLC dataset. All graduation cohorts 2009-2018. Only students with information on parental education. GPA is on a 0-10 scale (with passing grade 5.0). All variables (except GPA) in percentage. GPA percentiles 90p and 95p are specific to each student's major, pooling men and women. There are 73 majors at the UB.

Table C.17: Comparison GPA distributions. Admin records all universities.

	(1) GPA	(2) Top50	(3) Top25	(4) Top10	(5) Top5	(6) Top2	(7) Top5	(8) Top5
Fem	0.11***	3.11***	1.33***	-0.03	-0.26**	-0.23***	-0.23	-0.40***
	[0.00]	[0.27]	[0.23]	[0.16]	[0.12]	[0.08]	[0.18]	[0.13]
Obs.	161,207	161,207	161,207	161,207	161,207	161,207	75,478	120,507
Mean dep.var.	7.11	49.9	24.91	9.94	4.97	1.98	4.99	4.99
FE major	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
FE year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
FE uni	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes
FE field	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes
Univ.	All	All	All	All	All	All	UB	No Uni4

Notes: We computed each student's position in the GPA distribution of his/her major and university. We then pooled all observations and constructed indicators for belonging to the top (50, 25, 10, 5 or 2) percent of the pooled grades distribution. Unless otherwise stated, data for all majors and universities are being used (and for all available graduation cohorts). The last column excludes the university specialized in engineering (Uni4).

Table C.18: Comparison GPA distributions 2. Admin records all universities.

GPA	Mean Male	Mean Fem	Mean FM ratio
All Fields	6.96	7.23	1.04
STEM	6.79	6.77	1.00
Health	7.39	7.47	1.01
Hum	7.43	7.41	1.00
Soc	6.99	7.25	1.04
	Mean	Mean	Mean
zGPA	Male	Fem	F-M gap
All Fields	-0.033	0.027	0.06
STEM	-0.013	0.033	0.05
Health	-0.071	0.026	0.10
Hum	0.052	-0.025	-0.08
Soc	-0.076	0.039	0.12
	90-95p	90-95p	90-95p
zGPA	Male	Fem	F-M gap
All Fields	4.72	5.22	0.50
STEM	4.98	4.94	-0.04
Health	4.98	5.01	0.03
Hum	5.41	4.82	-0.59
Soc	4.38	5.30	0.92
	$> 95p$	$> 95p$	$> 95p$
zGPA	Male	Fem	F-M gap
All Fields	5.40	4.68	-0.72
STEM	5.16	4.71	-0.45
Health	5.15	4.95	-0.20
Hum	6.19	4.41	-1.78
Soc	5.21	4.89	-0.32

Notes: Top panel compares raw GPA (in a 0-10 scale), second panel compares GPA standardized ($zGPA$) using mean and standard deviation of the corresponding major-university distribution, third panel compares the 90-95th percentile bracket of the standardized GPA distribution, and bottom panel compares the 95-100th percentile bracket of the standardized GPA distribution.

Table C.19: Survey. Interest in graduate studies.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep.var.	Master	Master	Master	Master	PhD	PhD	PhD	PhD
Female	2.19	1.10	1.49	2.38	-5.32***	-9.55***	-2.18	-7.47***
	[1.46]	[3.03]	[2.24]	[1.93]	[0.99]	[2.78]	[1.43]	[1.35]
GPA	3.49***	-4.07	7.99***	0.60	5.67***	13.43***	7.60***	4.36***
	[0.81]	[3.37]	[1.15]	[1.09]	[0.61]	[3.20]	[1.00]	[0.76]
SES high	-0.04	-0.45	4.95**	-2.58	3.33***	2.00	4.10***	3.06***
	[1.39]	[2.90]	[2.29]	[1.74]	[0.87]	[2.40]	[1.43]	[1.09]
STEM	9.93***	17.26***			6.51***	11.24**		
	[2.25]	[4.80]			[1.61]	[5.02]		
Health	1.42	-3.80		3.27	7.74***	6.71**		8.72***
	[2.11]	[4.19]		[2.13]	[1.45]	[3.24]		[1.46]
Hum	9.44***	13.48***		10.20***	3.00**	4.46		3.68***
	[2.17]	[3.79]		[2.18]	[1.40]	[3.04]		[1.40]
Obs.	4,738	1,068	1,758	2,980	4,738	1,068	1,758	2,980
Sample	All	GPA high	STEM	No-STEM	All	GPA high	STEM	No-STEM
mean dv	69.3	71.2	71.8	67.7	10.1	17.4	9.87	10.27

Notes: Dependent variables are dichotomous indicators taking values of 0 or 100. University fixed-effects and intercept included in the models, but not shown in the Table. Interest in pursuing a Master's degree or PhD studies are not mutually exclusive categories. Omitted categories are Social Sciences and UB graduate. High GPA restricts to individuals with GPA above 7.8. Heteroskedasticity-robust standard errors. *** p<0.01, ** p<0.05, * p<0.1

Table C.20: Survey. Field-specific gender effects

Dep. Var.	(1) Master	(2) PhD	(3) Abroad	(4) Know	(5) Apply
Common coeff.					
All Fields	2.19	-5.32***	0.13	-8.25**	-7.15**
Obs.	4,738	4,738	3,264	954	942
By Field					
STEM	1.49	-2.18	5.33*	-1.39	2.51
No STEM	3.16	-6.74***	-3.46*	-12.05***	-10.14***
Health	-0.47	-9.51***	-1.77	-9.45	-2.44
Arts & Hum	-3.35	-6.09*	4.80	-9.15	-11.66*
Social Sciences	4.72*	-6.61***	-6.65**	-14.02***	-12.82**
FE uni	yes	yes	yes	yes	yes
Controls	yes	yes	yes	yes	yes

Notes: Each coefficient is from a separate regression model. The estimates in the bottom panel are obtained using field-specific sub-samples. All regression models include fixed-effects for university and controls for GPA and SES. In column 3 the dependent variable is an indicator for preferred graduate program abroad. The sample for Columns 4 and 5 restricts to individuals with GPA above 7.8 (out of 10). *High SES* is an indicator taking value one for students with at least one parent with a college degree. Heteroskedasticity-robust standard errors. *** p<0.01, ** p<0.05, * p<0.1