STEREOTYPICAL SELECTION

Martina Zanella*

Abstract

Women are still under-represented and struggling to establish careers in traditionally male-dominated fields. Does minority status in and of itself create a barrier to women's success? Experiments suggest that under-representation exacerbates the detrimental effect of the negative stereotypes that often characterize women's ability in these fields. However, in real-world environments, these results might not hold, as women in male-dominated occupations are often a selected group who chose to challenge the stereotype. In this paper, I answer this question by studying the performance of 14,000 students at an elite university across 16 departments, in a real-world setting that combines a choice with well-defined stereotypes - university major - with exogenous variation in peer identity - quasi-random allocation of students across class groups within the same course. The evidence indicates that those who go against stereotypes (e.g. women in math) do not suffer from being in the minority, but they impose negative externalities on those who select on stereotypes (e.g. men in math). In line with social identity considerations being incorporated into educational choices, the evidence points towards ex-ante "sensitivity" to social norms and preferences to engage with same-gender peers inducing students to select different majors and then reacting to the composition of the environment in a self-fulfilling way.

JEL codes: D91, I24, J15, J16, J24

*Department of Economics, Trinity College Dublin; email: zanellam@tcd.ie. I acknowledge the collaboration with the London School of Economics, that I thank for the data and the possibility to administer the survey. I am grateful for the advice and support from Nava Ashraf, Oriana Bandiera, and Robin Burgess throughout the project. I also thank advice from Manuel Bagues, Tim Besley, Gharad Bryan, Jeremiah Dittmar, Martina Kirchberger, Camille Landais, Alan Manning, Rachael Meager, Guy Michaels, Gaia Narciso, Steve Pischke, Anna Raute, Daniel Reck, Mark Schankerman, Daniel Sturm. I thank Andrea Alati, Luca Citino, Amanda Dahlstrand, Alexia Delfino, Gaia Dossi, Sacha Dray, Jay Euijung Lee, Nicola Fontana, Martina Fazio, Monica Langella, Virginia Minni, Niclas Moneke, Kilian Russ, Viola Salvestrini, Vincenzo Scrutinio, John Spray, Guglielmo Ventura, Anna Vitali, Celine Zipfel and seminar and conference participants for the helpful suggestions. A special acknowledgment goes to Judith Shapiro for priceless guidance in navigating the LSE system, feedback, and support. The use of the data and the survey were approved by the LSE Ethics Committee in October 2017 and April 2021.

I. INTRODUCTION

Despite significant progress has been made in reducing inequality in the labour market over the years, women are still under-represented in traditionally male-dominated occupations, giving rise to an economically consequential gap in labor market outcomes (Blau and Kahn, 2017; Goldin, 2014). This can be traced back to differences in educational choices, that amplify moving up the ladder. Women still make up only a small minority of students enrolling in majors in traditionally male-dominated disciplines, such as STEM (OECD, 2020; Bertrand, 2020; McNally, 2020), and even a smaller share of the people graduating, taking up jobs, and reaching leadership positions in these fields (OECD, 2020; Lundberg and Stearns, 2019; Bertrand, 2018; Beede et al., 2011). Does minority status in and of itself create a barrier to women's success in traditionally male-dominated fields?

Numerous experiments in psychology and economics document that under-representation is detrimental to performance, particularly in "counter-stereotypical" fields - where a negative stereotype exists about the group (Born et al., 2020; Karpowitz and Stoddard, 2020; Bordalo et al., 2019; Chen and Houser, 2019; Spencer et al., 2016; Hoff and Pandey, 2006; Steele and Aronson, 1995). This is the case for women in math and science-intensive fields as a well-known stereotype exists about the performance of women in these areas: "Women are worse in math and science compared to men" (Ellemers, 2018).

However, in real-world environments, minority status is not randomly assigned as in experiments. The women that make it in male-dominated fields are a selected group of women. Women that, at various points in their life, chose to go against societal expectations, and select a counter-stereotypical occupation where women are under-represented. As such, they might not be affected by stereotypes, and as a consequence the identity of their peers, as much as the average woman in the population.

This is what this paper sheds light on: how does minority status, and more generally, peers' identity, affect performance when selection is endogenous? Providing an answer to this question is crucial to design effective policies. Nonetheless, this margin has mostly remained unexplored due to the challenge of finding a setting where choices and minority status could be separated, as under-representation and negative stereotypes about the group often go hand in hand (e.g. women in STEM).

In this paper, I overcome this challenge by studying a real-world setting that combines a choice characterized by well-defined stereotypes with exogenous variation in peer identity. I assembled a rich dataset on the universe of students enrolled in undergraduate programs at an elite UK university across 10 academic years - from 2008 to 2017 - and 16 departments, that combines information on students' applications alongside their academic history and performance in each course, their allocation into courses and tutorials, their background information, as well as some characteristics of their instructors. I then estimate the effect of being in a minority when selection is endogenous and stereotypes play a role by assessing whether students' choice of major - whether it is in line or against gender stereotypes - interacts with the effect of peers' gender on performance once students are in the classroom.

I define stereotypical and counter-stereotypical choices by leveraging the variety of programs offered by the university and the heterogeneity in their gender composition. While 49% of the students enrolled in undergraduate programs at the university are women, women represent the minority of students enrolled in math-intensive disciplines and the majority of students enrolled in humanistic disciplines. This closely reflects the stereotype that "women are worse than men in math and science while better at reading and humanities" (Carlana, 2019; Reuben et al., 2014). Hence, I indicate as “stereotypical” the choice of students who select in departments characterized by a high share of students of their gender, while a “counter-stereotypical” choice is that of a student who selects in a department where their gender is under-represented.[1]

The exogenous variation in peers’ identity comes from three key features of the setting. First, for each course, students are quasi-randomly allocated to *classes* - small groups of maximum 15-20 people - to attend tutorials. Second, as, during their first year, students attend courses common to different departments, first-year courses are big modules characterized by multiple classes. I can hence control for *course fixed effects*, allowing me to account for potential differences across courses. Third, students attend on average 5 courses during their first academic year. Hence, each student can be observed across multiple classes, allowing to control for *student fixed effects*, which take care of student-specific characteristics that affect students’ performance and choices.

The causal effect of class gender composition is then identified by comparing the performance of the *same student* across all the courses that they attend during their first year, where they are allocated to classes with an exogenous composition of classmates, net of *course* and *student* fixed effects.[2] As numerous first-year courses are common across departments and class allocation does not depend on the program of enrolment, students who are in the same class in a given course are not necessarily allocated to the same class in other courses[3]. Hence, men and women who are enrolled in the same program of study are not exposed to the same within-student variation in class composition, allowing to exploit the presence of partially overlapping groups of peers to address the reflection problem that usually characterizes the estimation of peer effects (Bramoullé et al., 2020; De Giorgi et al., 2010). Furthermore, this provides a variation in class composition that goes beyond the program of enrollment, and the common support necessary to compare the effect for same-gender students who made different choices of major.

Besides the clear advantages in terms of identification, the focus on undergraduate students during their first year is also convenient to test the generalizability of experimental findings to real-world environments. In the first year, interactions within the classroom are important as students have yet to form their social networks outside the classroom. The latter has been shown to have an effect on performance on their own (for example, roommates - Corno et al., 2022; Jain and Kapoor, 2015; Stinebrickner and Stinebrickner, 2006; Sacerdote, 2001; or study groups - Oosterbeek and van Ewijk, 2014; Russell, 2017), which might

[1]As a proxy of stereotypical selection, I use the average share of men and women enrolled in undergraduate programs in each department over the period 2008-2017. The relationship between stereotypes and gender segregation of the workforce has been documented in psychology (e.g. Garg et al., 2018), and the share of females as an indicator of friendliness/stereotypes of the sector/field has been widely used in the literature (e.g. Bostwick and Weinberg, 2021; Kugler et al., 2021; Hebert, 2020). The answers to a survey performed on students also confirm that students are aware of the gender composition of departments. The results are however robust to other proxies of stereotypical selection.

[2]To support the validity of the identification strategy, I provide evidence that class allocation does not predict students’ characteristics and that the variation in the share of females in the class is not related to the variation in several predetermined student characteristics, conditional on scheduling constraints. Furthermore, to support the assumption of exogeneity of scheduling constraint, I compare the observed allocation to 1000 simulated unconstrained random allocations, providing evidence that students of different gender do not select systematically different courses among the first-year courses.

[3]The share of same program peers within the classroom is on average 0.49.

confound the effect of the classroom environment. Furthermore, students do not have experience regarding the academic system or their performance in this new system, and this is the first time that they interact with each other. Hence, stereotypes and social interactions can play a strong role in informing priors and expectations regarding students' performance and the performance of their peers.

I present three sets of results. First, I show that students who made different choices of major are affected in a significantly different way by the identity of their peers when they are in the classroom. I find that for students who made a stereotypical choice of major, e.g. men in math, a 10% increase in the share of same-gender classmates *increases* course grades by 0.325 points (2.0% of a standard deviation) on average. Having same-gender students in the classroom lifts performance at every point of the grade distribution, but the effect increases moving toward the top. On the other hand, for students who chose a major stereotypically associated with the opposite gender, e.g. women in math, a 10% increase in the share of same-gender classmates *decreases* final course grades by 0.288 points (1.76% of a standard deviation) on average. For this group of students, an increase in same-gender classmates reduces the probability of being in the mid-top part of the grade distribution but does not affect the probability of excelling in the exam. Lastly, class composition does not affect students who chose a major with a balanced gender ratio. The effect is present for both men and women, albeit it is stronger for men.

These findings significantly diverge from what we would have predicted if we had generalized the evidence gathered through experiments where selection has been shut down by design. Indeed, the students who are enrolled in counter-stereotypical fields, where their gender is negatively stereotyped, do not perform worse when in the minority in the classroom. On the other hand, it is students who are enrolled in stereotypical fields who perform significantly worse when surrounded by opposite-gender peers.

The main challenges to interpreting these results as the interplay between selection and the effect of peers' gender are two. First, the variation in peers' gender might be correlated with differences in other characteristics of peers, programs, teachers, or other potential confounders. Second, choices of major may be important for reasons other than the fact that they confirm or challenge societal expectations.

The second set of results aims at addressing these concerns. I proceed in three steps. First, I provide evidence that alternative mechanisms, such as differences in support across departments, teaching assistants' or peers' characteristics, ability peer effects, potential spillovers, and mechanical effects, are not able to reconcile the estimated patterns, rather is peers' gender that matters. I then continue by showing that gender stereotypes indeed play a crucial role. The results are mirrored when the analysis is replicated by replacing peers' gender with teachers' gender, providing additional evidence that choices of major carry information regarding how sensitive students are to gender stereotypes.[4] Lastly, I demonstrate that the estimated patterns are related to the interplay between stereotypical (counter-stereotypical) choices and the effect of peers' identity in the classroom by carrying out a placebo test and by replicating the analysis along ethnic lines, which additionally confirms that these patterns are not gender-specific.

Taken together, these two sets of results indicate that, when selection into minority status is endogenous, it affects how individuals react to the composition of the environment. What does this imply for policy?

[4]Class teachers have been shown to act as role models, breaking stereotypes regarding gender roles. See for example Breda et al. (2021);Porter and Serra (2020);Olsson and Martiny (2018);Carrell et al. (2010).

To illustrate this, I simulate what would happen to students' performance in the counterfactual scenario of a reallocation policy that enforces a more balanced gender ratio in male-dominated fields. As the only students who benefit from being surrounded by same-gender peers are students who made choices of major in line with stereotypes and belong to the majority group (men), enforcing a more balanced gender ratio in male-dominated fields would reduce differences in performance across groups. However, at the same time, it would lower average performance, as neither men nor women would benefit from a higher share of female classmates.

The third set of results investigates the margins through which selection plays a role. At least three concurring explanations are possible. First, once students took the decision to select a stereotypical (or counter-stereotypical) field, they might be induced to act in a way that confirms their decision to reduce cognitive dissonance (Akerlof and Dickens, 1982; Festinger, 1957). Second, these differences might be due to the particular environment that characterizes male-dominated fields compared to female-dominated fields. Third, as students internalize stereotypes and social identity considerations when making their choices of major (Kugler et al., 2021; Del Carpio and Guadalupe, 2022; Pan, 2015; Oxoby, 2014; Goldin (2014); Bertrand (2011): Card et al., 2008; Akerlof and Kranton, 2000), students who went against stereotypes might be more resilient to belonging to the disadvantaged over-represented group.

I argue that the evidence points towards the third explanation being the most plausible. I do so by building a theoretical framework to rationalize how minority status affects students' performance. I then derive predictions in the absence of selection and show how students internalizing social identity considerations when choosing their major can modify these predictions. Lastly, I bring the model to the data using rich administrative data and a novel survey.

The theoretical framework assumes that students invest in effort to maximize their performance in exams and their social image.[5] In this setting, under-representation affects effort choices through two channels: (i) homophily - students' marginal cost of effort is lower when surrounded by same-gender peers;[6] and (ii) stereotypes - by making identity salient, minority status changes students' beliefs regarding their ability and the ability of their peers, and as a consequence, expected returns to effort.[7]

The model delivers two key testable predictions. First, in the absence of selection, when image concerns prevail, an increase in the share of same-gender classmates increases effort for students who are enrolled in counter-stereotypical departments. Second, the relevance of peers' gender composition is lower if stereotypical associations are weaker and the cost of interacting with the opposite gender is lower. Using a simple Roy model of major choice[8], I show that this might be the case for students who made counter-stereotypical choices. If we assume that selection into majors plays a role because students internalize stereotypes and

[5]This rationalization of effort choices is used also by Bursztyn et al. (2019); Bursztyn et al. (2017); Austen-Smith and Fryer (2005); Akerlof and Kranton (2000).

[6]Under-representation affects effort as it reduces opportunities to benefit from academic and emotional support (Tinto, 1975) due to a preference to engage with same-gender peers (Inzlicht and Good, 2006)

[7]Students do not have perfect information regarding their ability and the ability of their peers. They hold imperfect beliefs, which are affected by stereotypes. I build on Bordalo et al. (2019) I assume that stereotypes-induced distortions are stronger the more group identity is top of mind. By increasing the salience of group identity, under-representation increases the strength of stereotypes-induced distortions, boosting (positive stereotypes - e.g. men in math) or depressing (negative stereotypes- e.g. women in math) beliefs on ability and, as a consequence, affecting expected returns to effort.

[8]The model builds on Del Carpio and Guadalupe (2022)

gender composition when making their choice, students enrolled in different majors are going to be heterogeneous exactly along traits that determine the effect of minority status, such as stereotypical associations and the preferences for same-gender peers.

I test the first prediction by exploiting teaching assistants' evaluations of students' participation in class which, the evidence indicates, mostly reflects image concerns. Contrary to the predictions of the model and the findings of experiments in controlled settings[9], I find no evidence of an effect of class composition on participation for students who are enrolled in counter-stereotypical departments. This indicates that selection against stereotypes moderates the effect of minority status.

The second prediction is tested by exploiting the results of a survey performed on a subsample of students[10] at the university. The evidence is consistent with students internalizing social identity considerations when choosing their major. A Gender-Scientific Implicit Association Test, that elicits implicit associations of men with scientific disciplines and women with humanistic disciplines (Greenwald et al., 1998), shows that the students who uphold the strongest stereotypical associations are those who made a stereotypical choice of major, i.e. men enrolled in math-intensive fields and women enrolled in programs related to humanities or non-math-intensive social sciences. The answers to questions regarding students' propensity to engage with same-gender peers display the same pattern. Those students that display the strongest tendency to engage with same-gender peers are men enrolled in male-majority programs and women enrolled in female-majority programs.

Furthermore, the evidence does not support the two alternative hypotheses. The elicited implicit stereotypical associations and preferences for same-gender peers are not significantly stronger for students that have been exposed more to the environment (as proxied by years of bachelor), as we would expect if the environment affected students' beliefs. Furthermore, not every student making a stereotypical choice of major is affected by peers' gender to the same extent. The students who made stereotypical choices that benefit the most from being surrounded by same-gender classmates are those that come from countries where gender norms are stronger.[11] This heterogeneity in the effect, even for students who made the same decision, seems to suggest that ex-ante "sensitivity" to stereotypes and social norms inducing students to select different majors and then react to the composition of the environment in a self-fulfilling way seems to be a more plausible explanation.

My work contributes to the literature that aims at assessing the effect of stereotypes on behavior, beliefs, and ultimately performance, and their interaction with the composition of the environment (e.g. Coffman et al., 2021; Bonomi et al., 2021; Karpowitz and Stoddard, 2020; Born et al., 2020; Bordalo et al., 2019; Chen and Houser, 2019; Bordalo et al., 2016; Coffman, 2014). This paper shows that selection into minority status plays a role, affecting the extent to which changes in the composition of the environment affect

[9]Coffman et al. (2021); Born et al. (2020); Chen and Houser (2019); Bordalo et al. (2019); Coffman (2014) all find that relaxing minority status increases participation and willingness to contribute to discussions in counter-stereotypical domains by reducing the daunting effect of negative stereotypes.

[10]The survey was administered to all the students enrolled in undergraduate programs in 2021. 500 students (10% of the student body) answered the survey. The response rate is similar to that of other surveys performed on undergraduate students at the university.

[11]Gender norms are proxied by the Global Gender Gap Index (GGI) in the student's country of origin. This categorization is based on findings from Guiso et al. (2008), which provides evidence that the GGI index is significantly correlated with gender gaps in math and reading.

performance by shifting the salience of identity. The effect of peers' identity differs depending on whether individuals made occupational/educational choices that confirm or challenge the stereotypes associated with this identity.

This paper relates also to the significant body of literature that studies the effect of the identity of peers on performance and choices in real-world environments in higher education[12] (Shan, 2020; Zölitz and Feld, 2021; Griffith and Main, 2019; Booth et al., 2018; Huntington-Klein and Rose, 2018; Hill, 2017; Oosterbeek and van Ewijk, 2014; Giorgi et al., 2012; Griffith, 2010). Both the methodology and the analysis of my paper build on this work. However, the majority of these studies focus on a single field or estimate average effects, overlooking the interplay between selection and the effect of peers' identity, and providing only a partial picture. By studying the effect of the composition of the environment on performance across multiple fields within the same setting and exploiting the same source of variation, this paper provides evidence of an additional unexplored margin that might be able to reconcile apparently diverging results: endogenous selection. To the best of my knowledge, the only other paper that assesses the heterogeneity of peer effects based on selection is Pregaldini et al. (2020), which compares students who have self-selected into a STEM specialization with students who have self-selected into a language specialization in a Swiss High School. In line with the results of this paper, they find heterogeneous effects depending on students' choices, which they attribute to differences in willingness to compete. A few other studies assess heterogeneous treatment effects based on the program of enrollment (Hill, 2017) or course type (Zölitz and Feld, 2021 and Oosterbeek and van Ewijk, 2014 distinguish between math and non-math courses), finding, once again, that aggregate results hide significant heterogeneous effects. None of them however explicitly explores selection as a potential reason behind these patterns.

Lastly, the paper contributes to the recent literature that theorizes that individuals incorporate stereotypes (e.g. Kugler et al., 2021; Del Carpio and Guadalupe, 2022; Oxoby, 2014) and the gender composition of the occupation (Pan, 2015; Card et al., 2008) when making occupational choices by showing that students' choices are indeed consistent with the predictions of these models, and by providing evidence of the importance of considering this margin when designing policies to tackle inequality in performance. If we do not acknowledge endogenous selection into minority status, we might design policies that target and nudge minorities in selective environments to level the playing field that end up being ineffective and might even backfire.

The remainder of the paper is organized as follows. The next section explains the institutional setting. Section III describes the data and descriptive statistics. Section IV presents the empirical strategy and the tests in support of the identifying assumptions. Section V illustrates the main results. Section VI presents the theoretical framework and the evidence in support of the different mechanisms underlying the results. Finally, Section VII concludes.

[12]The focus of this paper's contribution is restricted to higher education as the mechanisms at play in this setting might be different from those affecting students down the ladder. Students have different needs, interests, and objectives before they enter this high-stake environment, where their performance will likely determine their lifetime occupation.

II. Institutional Setting

LSE is one of the leading universities in the UK and is characterized by a highly competitive and selective admission process. First, securing an offer of admission from the institution is very challenging. Students apply from every part of the world to attend undergraduate programs at the University, competing for a fixed number of available places. Each year the University admission office receives approximately 26,000 applications for roughly 1,700 places. This fierce competition for places means that roughly 75% of the students who apply are rejected every year. Second, applying to undergraduate programs in the UK is a strategic decision as university applications are not at zero cost. The UK admission system is centralized. Students can only apply to 5 programs across all UK universities. Hence, applying to undergraduate programs with a high rejection rate, such as the LSE, has a high opportunity cost, as it implies not being able to apply to other programs.

The result is a selected, but multicultural and diverse student body: 49% of undergraduate students are women, roughly 36% of students are White, 45% are Asian (47% of which Chinese, and 27% Indian), and 4% are Black.

The university offered a very diverse portfolio of courses during the academic years considered: 44 undergraduate degree programs across 16 academic departments. These programs span different disciplines from Mathematics, Economics, Finance, and Statistics, to Anthropology, Sociology, International Relations, and International History. The variation in gender composition across departments is significant. We can observe majority-male departments, where men represent $>$65% of the students enrolled in undergraduate programs, alongside majority-female departments, characterized by $>$75% of women (Figure I).

The academic system is homogeneous across programs belonging to different disciplines as students take on average 4 courses during each academic year. Each degree program has its list of core courses – either compulsory or offering a very constrained choice – and a list of elective courses. During the first year, which is the focus of this paper, most of the courses are compulsory, and students have very limited possibilities to choose electives[13]. Students might be required (or allowed) to attend courses outside their department or program of enrolment, particularly so during the first year, when students attend several courses outside their department[14].

For each course that students take, they attend lectures and *classes*. Lectures are followed by all the students enrolled in the course and are taught by a lecturer. Classes are taught by teaching assistants and students are divided into small groups of maximum 15-20 students. In total, students attend 10 classes each term, one per week (generally one hour long). The empirical strategy of the paper exploits students' assignment to these small groups and their composition.

Each academic year is composed of three terms. Michaelmas term (September- December) and Lent term (January-April) are 10 weeks long and are teaching terms during which students attend lectures. The third term is called Summer term (May - June). During this term, students primarily prepare and take exams.

[13]Students are generally allowed to choose at most one course

[14]See Table A.3 in Appendix A. for more detailed information

II.A. Assignment of students to classes

Students are allocated to classes before the beginning of the academic year when each one receives a schedule for lectures and classes.

Class allocation can be considered as good as random conditional on scheduling constraints. According to the university's timetable office, only two criteria are used to allocate students to classes. First, students cannot attend more than 4 consecutive hours of teaching. Second, the timetable office considers potential clashes with other courses students attend during the academic year.

Students officially allocated to the same class represent the peers each student will engage with during classes for the whole year. Once students are assigned to a class, they can only change under exceptional circumstances and via an official request that has to be approved. Furthermore, the official allocation of students into classes reflects very closely actual attendance since classes are compulsory and teaching assistants are required to register students' participation in class.[15]

In the analysis, I am going to exploit this conditionally random class allocation and assess whether students' performance is affected by peers' identity by defining treatment based on the characteristics of all the students (undergraduate, general course, intercollegiate, exchange, and postgraduate) officially allocated to the same class at the beginning of the academic year.

III. Data and Descriptive Statistics

The paper combines four sources of data: the university's administrative records, the class register, human resources' records, and additional information gathered through an online survey.

Administrative Records The university's administrative records contain information on academic history and performance in each course for the universe of students enrolled in undergraduate courses at LSE during the academic years 2008/09 - 2017/18. The administrative records also contain some individual background characteristics for students, such as gender, age at enrolment, country of origin, ethnicity, term time accommodation, and the previous schools attended. Lastly, the university collects and stores students' application information. I have the complete list of all the qualifications that students submit when filing their applications for all the students enrolled in undergraduate programs at the university between 2011 and 2017. Furthermore, for the academic years 2007/08 - 2019/202, the database contains information on the total number of students who applied, got accepted, or were rejected to each program, divided by gender, ethnicity, and country of birth.

Class Register The class register contains information on lectures and classes allocation for each student, professor, and teaching assistant. I use this information to (i) identify all the students that attend classes in the same class group and construct a measure of class composition based on students' demographic char-

[15]Whenever students are absent without reason two consecutive times, they receive an "ammunition", and their advisor is contacted. When students fail to attend too many classes, they can be denied access to the final exam. Unofficial changes are possible, but this type of switching is usually limited to one or two sessions. It is difficult to obtain reliable numbers on unofficial switching. From my own experience and consultation with teaching staff, these instances are minimal, given that students have to notify their teaching assistant to be recorded as present when attending a different class

acteristics; (ii) match students' class group allocation with teachers' class group allocation and construct a matched teacher-student database. On top of this information, the class register contains data on attendance and performance in class as the university requires teaching assistants to keep track of students' attendance each week and evaluate their participation and overall class performance (problem set grades and overall class performance assessment). [16]

Human Resources Records The university stores background information (gender and ethnicity) for all the people employed by the institution. I was able to recover background information and match them with the class register for a subsample of teachers who were employed at the university with an official contract over the years considered.

Survey Information from administrative records is complemented with data gathered from an online survey. The survey was administered to all the students enrolled in undergraduate programs at LSE during the academic year 2020/2021. The questionnaire was administered via email. Participants agreed to take part in the survey and signed an informed consent, in which it was explained that the survey was part of a research project aimed at understanding the determinants of academic and labour market performance. The students were informed that if they agreed to participate, they would be asked a few questions regarding demographics, their experience as a student, and their and beliefs and attitudes. The time to complete the survey was around 10 minutes and students were entered in a draft for three amazon vouchers. 498 students completed the survey, representing 10% of the overall population. The survey includes questions on demographic characteristics, educational experience, social networks, explicit attitudes towards gender-specific skills, and an Implicit Association Test eliciting the association male-Scientific, female-Humanistic. More information on the survey can be found in Appendix F..

III.A. Sample Selection

The full sample of students that are enrolled in undergraduate programs at LSE between 2008 and 2018 consists of 14,389 students. I restrict the sample for the analysis to undergraduate students enrolled in their first year that are attending courses for the first time.

The focus on first year courses is motivated by three factors. First, students' choices are strongly restricted in first year courses, allowing to minimize the problem of selection into courses. Second, during the first year students enrolled in different programs have several courses in common. This provides a variation in the composition of the peers in first-year courses that goes beyond the variation in the composition of the students enrolled in each program. Furthermore, it allows to observe significant variation in the composition of classes within each course, given that first-year courses are large and each one is characterized by numerous students and numerous classes. Third, this is the first time that students meet each other. This is the perfect moment to study how behavior and performance are affected by the environment as students have an empty information set regarding each others. Thus, their beliefs and expectations regarding their own performance and the performance of their peers' will mostly rely on priors rather than the signals obtained

[16] These data are stored to prove visa requirements and are used by academic mentors, professors, and teaching assistants to write reference letters for students.

from observing their classmates' behavior and performance in class and exams.

The sample is further restricted to first-year courses that students attend during the first term. A few programs are characterized by half-year courses that are scheduled only for the second term. To limit the problem of endogenous choices of courses in the second term (they could be influenced by the class composition experienced in the first term), the sample is restricted to first-year courses that students attend during the first term, half-unit or full unit courses, excluding half-unit courses that students attend only in the second term (9.2% of observations).

The sample is restricted to class groups with a size between 6 and 28 students (excluding smallest and largest 0.5% of observations) to exclude unreasonable and unlikely class sizes.[17] Furthermore, I exclude from the sample all the students who changed class group during the first term, as in this case I cannot identify the initial exogenous class allocation. This happens 5.09% of times.[18] Lastly, I exclude class groups where I observe more than 50% of students changing group. These occurrences correspond to class group "restructuring" - for instance when a class group gets cancelled and students are re-shuffled to other class groups (2.5% of course-year-class group level observations).

Lastly, since the empirical strategy relies on class group fixed effects and student fixed effects, the sample is restricted to courses with at least two class groups and students for which I can observe, after the sample selection explained above, a test score in at least two courses: 54.603 course-year-class group level observations, corresponding to information on 14.313 students (99% of the original sample), who attend on average 3.8 courses in their first year.[19]

As it can seen from Table II, the sample is made of 512 courses, with an average course size of 138.44 students, split in 9.36 class groups of 13.43 students each.

III.B. Main outcome variables

Overall Course Performance Students are required to get 4 credits during each year of undergraduate programs by attending courses that are worth 0.5 or 1 credit. At the end of each course students get a grade between 0 and 100 that assesses what they learned during the course. This grade can be the result of a final exam, or the combination of final exam and essays or take-home assignments during the year. A student is deemed to have failed a course if his grade is below 40, a third class honor is awarded if the grade is between 40 and 49, a lower second class honor if the grade is between 50 and 59, an upper second class honor if the grade is between 60 and 69 and a first class honor if the grade is 70 or above. Students can also decide to drop-off from the course by not attending the final year exam or not submitting a part of the summative assessments. In this case, a grade of 0 is assigned by the university.

Final grades represent a good measure of performance for two reasons. First, every piece of assignment (exams, essays or take home assignments) is based on absolute grading. Grades are based on the perfor-

[17] The university regulations fix a cap to 17 students per class group. There are frequent exceptions, but class groups with more than 28 students seems very unlikely, and thus are excluded from the sample.

[18] In Appendix A. I present the results of tests that show that the decision to change class group is not driven by the composition of the group.

[19] Excluding an interdisciplinary course that carries no credits and that is excluded from the analysis.

mance of each student in the exam, without any ranking of students or curving of grades.[20]. Second, LSE is characterized by a blind marking system: students do not indicate their names in any type of assessment, but only a candidate number which is secret to examiners and only known to the student.[21] This allows me to disregard potential confounders such as discrimination.

First-year course grades are characterised by substantial variation. This is mainly due to the fact that not all first-year courses count for the final degree classification to the same extent. A student registered on a BA or BSc programme who has completed the first year of the programme and who has passed assessments in courses to the value of at least three credits will be eligible to progress to the second year.[22] The "year-one average" counts for 1/9 of the final classification grade and is calculated by adding together and averaging the best six out of eight grades in first year courses. All first-year one-unit credits will be counted twice, and any half unit credit is counted once to make a total of eight first year grades.

Table I shows some descriptive statistics regarding first-year grades. The average grade is 60.32, with a standard deviation of 16.35 points. Students do not seem to perform in the same way in all courses, as the within-student standard deviation is 8.61.

Class Participation In order to test the model predictions, I analyse data on teachers' evaluation of students' participation in class. Teaching assistants are required to assess students' participation in class at the end of each term. Participation grades are based on the contributions that each student gives in class during the ten weeks of the term. Each teaching assistant is asked: "please mark the student's overall participation in class during the term" and can give a score from 0 to 3, where 0 stands for *"No contribution"*, 1 for *"Occasional contribution"*, 2 for *"Reasonable and alert interest shown"*, and 3 for *"Lively interest and frequent contributions"*.

Teachers' evaluations of students participation in class provide an indication of class dynamics and students interactions. However, since they are grades that teachers give to students, they are not objective measures of students' participation. Thus, results can be considered as indicative of class dynamics conditional on the assumption that teaching assistants do not discriminate or evaluate underrepresented groups differently when they are in a minority.

Even if assessment is in principle compulsory, not all the teaching assistants give a feedback to students. Attrition increases during the academic year: in Michaelmas term, participation is missing for 16.75% of course-year-class group level observations, while 40.68 % are missing in Lent term. For this reason, the analysis of participation grades is restricted to Michaelmas term. In Appendix A., I provide evidence that missing participation information is due to teaching assistants providing no information for everybody in the class and does not depend on individual characteristics or the composition of the class group.

Given the sample restriction, participation grades will shed light on the dynamics that characterize classes in the first half of the academic year. While on the one hand this can represent a limitation since exam grades are the result of students' effort during the whole academic year, on the other hand first-term

[20] For reference: https://www.lse.ac.uk/social-policy/Current-Students/BScProgrammesMarkingframe.pdf

[21] For reference: https://info.lse.ac.uk/Assessment-Toolkit/Marking-and-moderation, https://info.lse.ac.uk/current-students/challenging-results-and-appeals

[22] With the only exception of the Bachelor of Laws in which students progress to the second year if they have passed all four credits

measures are more indicative of how students' behavior is affected by stereotypical distortions and priors, since this is the first time students meet and interact with each other.

The sample for the analysis on class dynamics consists of students for which I can observe a participation grade in at least two class groups: 44.771 course-year-class group level observations, corresponding to information on 13.424 students (93.8% of the sample). Table I shows that the average participation grade received by students is 2.04, with a between-student standard deviation of 0.85, and a within-student standard deviation of 0.58. Figure III shows a histogram of the within-student participation gap. I define the participation gap as the difference between student i's highest and lowest grades across all the seminars they attend during Michaelmas term of the first year. The variation is significant since the difference between the highest and the lowest participation grade for the median student is 1, and 30% of students experience a difference of at least 2 between the highest and the lowest participation grade.

III.C. Additional key variables

Ex-ante Measure of Ability I construct a measure of ability at entry based on the information that students provide during the admission process. The university bases its admission decisions on personal statement, academic achievement, and references. Every program has minimum entry requirements, which are publicly available and clearly stated in the guidelines. They are based on A-level qualifications or equivalents. The A Level is a subject-based qualification conferred as part of the General Certificate of Education, as well as a school leaving qualification offered by the educational bodies in the United Kingdom and the educational authorities of British Crown dependencies to students completing secondary or pre-university education. Students typically study three A levels in different subjects, and the majority of universities set their entry requirements according to this measure. A levels are graded on a scale of A*/A/B/C/D/E. Each A-level grade is worth 30 QCA (Qualifications and Curriculum Authority) points.

Following Campbell et al. (2019), I construct a measure of individual ability at entry based on the best three exam results among the A-level qualification scores students declared when applying to LSE. Some students take courses that are equivalents to A-Levels.[23] In these cases I calculate their A-Level-equivalent scores based on the university conversion tables for foreign students.

Using this criterion, I am able to construct a qualification score for a total of 9.449 students, 90.57% of the students enrolled in undergraduate programs at LSE between academic years 2011/12 and 2017/18.[24] Figure IV displays the resulting qualification score. Due to the high demand for places, the mean qualification score of students enrolled in undergraduate programs at LSE is high: 503.9 QCA points[25] with a standard deviation of 34.26[26].

Information on Previously Attended Schools The university administrative records contain information

[23] Appendix B.3., displays the correlation of qualification scores constructed using different methodologies. The measures are strongly correlated with each others.

[24] I don't have admission information for students who enrolled in LSE before 2011

[25] This corresponds to a score in between a person that got two A-levels with A* and one A-level with A, and a person that has one A-level with A* and two A-levels with A.

[26] This corresponds to one A-level grade, i.e. 30 QCA points.

on the school students attended before enrolling at LSE. I match these data with the UK Government register of schools and colleges. This register contains information on schools and colleges in England, Wales, Scotland, Northern Ireland, or overseas establishments of UK institutions. I recover information on the characteristics of the previous school of enrolment for 63% of the students in my sample.

I use information such as whether the student attended a single-sex vs mixed-sex school, or an independent vs state school in the mechanisms and robustness sections. Mixed- vs single-sex school information is used as a proxy of the extent to which students are comfortable in engaging with other gender students, while whether students attended independent or state schools is used to characterise students' social background.

Global Gender Gap Index (World Economic Forum) The university administrative records contain information on students' country of origin. I match this information with data on the World Economic Forum Overall Global Gender Gap Index for the country between 2006 and 2018.[27] The Global Gender Gap Index measures the level of gender equality for 130 countries around the world. To do this, it ranks countries according to calculated gender gaps between women and men in four key areas: health and survival probability, education attainment, economic participation and opportunities, and political empowerment and representation.

I am able to match the student's country of origin with the Global Gender Gap Index for the country for 92.44% of students in the sample. I use this information as a proxy for the strength of stereotypes regarding gender skills and roles in the student's country of origin.

IV. EMPIRICAL STRATEGY

The main objective of the empirical analysis is to assess what effect selection in line or against stereotypes has on the impact of peers' identity on performance.

The identification of the effect of peers' identity is generally problematic because of two well-known issues that affect the estimation of peer effects: (i) the problem of correlated effects, which stems from correlated unobserved characteristics due, for instance, to endogenous choice of peers; (ii) the reflection problem, i.e. the difficulty of distinguishing between the impact of peers' outcomes (endogenous peer effects) and peers' characteristics (contextual peer effects) due to simultaneity in the behavior of interacting agents (Manski, 1993).

These two concerns are alleviated in this setting thanks to three key features of the environment. First, the quasi-random allocation of students into classes allows to address the problem of endogenous peer choices. Second, since students attend multiple courses during their first academic year, the performance of same student is observed in different courses. Hence, I can control for student fixed effects, alleviating concerns of correlated effects (Bramoullé et al., 2020). Third, numerous first-year courses are common across departments and class allocation does not depend on the program of enrolment. Hence, students who are in the same class in a course are not necessarily allocated to the same class in other courses. This allows to alleviate the reflection problem by exploiting the presence of excluded peers, i.e. the sets of peers of peers

[27] Source: The World Bank data

do not perfectly coincide as classmates of classmates are not necessarily classmates (De Giorgi et al., 2010).

I then leverage this identification strategy to assess how selection into majors interacts with the effect of class composition. To do so, I exploit two additional key characteristics of the setting. First, LSE offers a very diverse portfolio of programs, which allows to define a measure of stereotypical choices. Second, class allocation is orthogonal to the department of enrollment, and students attend several courses outside their department. This ensures substantial common support in class composition and within-student variation in class composition across programs of enrollment.

More detailed explanations on the identification strategy can be found in the following sections.

IV.A. The causal effect of peers' identity

I estimate the causal effect of peers' identity by comparing the performance of the same student i across the courses c that they attend during their first academic year a, where they have been assigned to classes g with exogenous peers' characteristics, net of course and student fixed effects. I estimate the following specification:

$$y_{iacg} = \alpha_{ac} + \alpha_i + \beta \times \mathit{SLM}_{iacg} + \epsilon_{iacg} \tag{1}$$

where the independent variable y_{iacg} is student i's grade in course c in academic year a. SLM_{iacg} is the *share of students like me*, i.e. the share of same gender classmates in class group g of course c: the share of females for women and the share of males for men.[28] β captures the effect of an increase in the share of same gender students in the class[29]. Lastly, α_{ac} and α_i are course $\times$ year and student fixed effects respectively.

Course $\times$ year fixed effects (α_{ac}) are essential to obtain exogenous variation in peers' characteristics. Students are randomly allocated to classes only within courses. Hence, course fixed effects are essential to capture differences in the overall course gender composition, which would lead to a systematically different probability of being exposed to same gender peers for the same student across courses.[30] In addition, course fixed effects control for other observable and unobservable individual traits and course characteristics that determine differences in students' performance across courses. For instance, Statistics and Mathematics courses might be less conducive to interactions and discussions compared to courses as History or International relations. Or else, courses have different types of assessment and different levels of difficulty.

Individual fixed effects α_i allow to address several concerns. First, they capture students' course selection and scheduling constraints. This accounts for the fact that students who are enrolled in the same program and/or make the same constrained choices are more likely to be assigned to the same class, since they share

[28] In Appendix C.2., I provide evidence that defining class composition as the share of same gender peers using the leave-out mean (Angrist, 2014) does not affect the results.

[29] The approach follows the empirical strategy of Brenøe and Zölitz (2020), Anelli and Peri (2017), and Feld and Zölitz (2017).

[30] For example, John is enrolled in the BSc in Economics. John attends a Mathematics course, MA100, and an Economics course, EC100. The Economics course is compulsory also for students enrolled in other undergraduate programs, such as BSc in Accounting, BSc in Government, and BSc in Economic History. Hence, EC100 is attended by a higher number of women compared to MA100. Thus, John will be more likely to have a higher share of female classmates in EC100 compared to MA100 overall. However, once we control for the overall course composition, the identifying variation we are left with is the deviation of class composition from the overall course gender composition. As class allocation is orthogonal to individual characteristics, this is as good as random. Hence, the share of female classmates John is exposed to is exogenous once we control for course fixed effects.

similar schedules for lectures. Second, individual fixed effects control for the fact that students enrolled in different programs have different probability of being exposed to same gender classmates due to the different gender composition of programs (and the combination of courses that are part of each program). Third, individual fixed effects allow to control for observable and unobservable individual characteristics that are student-specific, common across courses, and determinants of program/course selection and average performance, alleviating the problem of correlated effects (Bramoullé et al., 2020).[31] Lastly, including individual fixed effects allows to estimate the effect of class composition by exploiting a *within-student* variation, rather than a *within-course* variation in class composition. While the former compares the performance of the same student across courses, the latter compares the performance of students allocated to different classes within the same course, assuming that classes in different courses are independent from each other. Assuming class independence would be problematic in this setting as we observe the same student in different classes and there might be spillovers across courses. Students' performance in a course might not only be affected by their classmates in the course, but also by their experience in other courses. I discuss the extent to which spillovers affect the estimates in Section V.B., where I show that the estimates controlling for individual fixed effects are more conservative than the estimates without individual fixed effects.

Two final concerns are addressed by the characteristics of the grading system at the university. First, discrimination, or differential treatment of minority students during the examination process, does not represent a concern in this setting, since grading of exams, essays, group projects, etc. is characterized by blind marking.[32] Second, the effect of peers' identity on the performance of men and women separately can be identified only under the assumption of no zero-sum game. This assumption does not seem unreasonable in the LSE setting as marking is based on absolute grading. Hence, grades are based on the performance of each student in the exam, without any ranking of students.[33]

IV.B. The impact of stereotypical and counter-stereotypical choices

The identification strategy explained in the previous session is then leveraged to assess how selection into majors interacts with the effect of class composition. In particular, I consider whether students' choices of undergraduate programs are in line with gender norms and stereotypes - *"stereotypical selection"*.

Following an approach widely used in the literature (see Hebert, 2020; Bostwick and Weinberg, 2021; Kugler et al., 2021 as examples), I infer female stereotypes in the field from its gender composition. Specif-

[31] As an example, John is particularly interested in economics, passionate about history and struggles in math-intensive disciplines. He chooses the BSc in Economics and he attends EC100, MA100, ST102, and chooses EH101 as elective. Given his passions, he performs better in EH101 and EC100, where by chance he is allocated to female-majority classes. However, he struggles in MA100 and ST102, where he has been randomly allocated to male-dominated classes. Without controlling for student fixed effects, we would be mistakenly attributing John's lower performance to class composition rather than person-specific skills and attitudes.

[32] Any piece of assessment is marked by a first marker and a second marker, in most of the cases without seeing the first marker's grades/comments. Where there are any differences in mark, the two markers discuss and agree the final mark. The process is completely blind: students do not indicate their names in the exams, but only a candidate number which is secret to examiners and only known to the student.

[33] Regarding participation in class, grades are based on class interactions over a 10 weeks period. A zero-sum game assumption would require that each week 50 minutes class is fully saturated, i.e. students are racing to answer or ask questions, so that each student's intervention will take from other students' participation time. Although this is possible, from personal experience and consultations with teaching staff, this is quite unlikely, especially in first year courses.

ically, I use as a proxy for stereotypical selection the average share of women and men enrolled in undergraduate programs in each department over the period 2008-2017. A man's choice is considered in line with stereotypes if he is enrolled in a program that belongs to majority male departments, e.g. Mathematics, Statistics, Economics, Finance. On the other hand, it is considered as against stereotypes if he is enrolled in a department such as Anthropology and Sociology, which are majority female departments. In the same way, a woman's choice is considered as more in line with stereotypes the higher is the share of women in her department of enrollment between 2008-2017. A graphical description of the measure of stereotypical selection used can be found in Figure IX.

The average share of women enrolled in undergraduate programs in the department over the period 2008-2017 can be considered a good measure of the extent to which students' choices are in line (against) stereotypes for three reasons. First, this unequal gender distribution barely changed between 2008 and 2017 (Figures A.1), and it closely mirrors the distribution of men and women across subjects in the UK Higher Education system (Figure A.2). Second, it is the result of men and women applying to systematically different programs rather than the university's selection process (Figure A.4). Third, it reflects stereotypes regarding group-specific skills and roles. In line with lab and field experiments in the academic setting documenting a widespread belief that women are worse than men in mathematics and science while being better at reading (Carlana, 2019; Reuben et al., 2014; Ellemers, 2018; Lane, 2012), majority male departments are primarily characterized by math-intensive programs, while majority female departments are characterized by programs related to humanities.[34]

I exploit two definitions of stereotypical selection based on the average share of men and women enrolled in each department over the period 2008-2017: a continuous measure and a categorical measure.

Continuous Measure Exploiting the continuous measure, I estimate the following specification:

$$y_{iacg} = \alpha_{ac} + \alpha_i + \beta_1 \times \mathit{SLM}_{iacg} + \beta_2 \times \mathit{SLM}_{iacg} \times \mathit{STS}_i + \epsilon_{iacg} \tag{2}$$

where SLM_{iacg} is the *share of students like me*, i.e. the share of same gender classmates that student i experiences in class g, course c and academic year a. STS_i stands for *stereotypical selection*, measured as the average share of same gender students in student i's department of enrolment across academic years 2008-2017. β_2 provides an indication of the extent to which being surrounded by same gender classmates has a different effect on performance for students who made choices that are more or less in line with gender stereotypes. A positive coefficient implies that the performance of students who are enrolled in departments that are stereotypically-congruent (women in Humanities and men in Mathematics) is more sensitive to the composition of the class compared to the performance of students who are enrolled in departments that are stereotypically associated with the opposite gender. A negative coefficient, on the other hand, implies that the students whose performance is more affected by being in a minority are those who are enrolled in departments stereotypically associated with the opposite gender (women in Mathematics and men in Humanities). The standard errors are clustered at the class level.

[34]The relationship between stereotypes and race and gender segregation in the workforce has been documented in psychology(e.g. He et al., 2019; Garg et al., 2018).

Categorical Measure The continuous measure of stereotypical selection relies on the assumption that the effect of choices in line or against stereotypes have a linear effect. In order to relax this assumption, I perform a second analysis. I define a categorical measure of stereotypical selection, where I categorize choices using a definition that becomes gradually more stringent. I start by considering students who are enrolled in the top five departments with the highest average share of same gender students as students making stereotypical choices, while students who are enrolled in the five departments with the lowest share of same gender students as students who made a choice not in line with gender stereotypes. I then replicate the same analysis by considering the top and bottom 2 and 3 departments in terms of share of same gender students.[35] I estimate the following specification:

$$y_{iacg} = \alpha_{ac} + \alpha_i + \beta_c \times \mathit{SLM}_{iacg} + \beta_n \times \mathit{SLM}_{iacg} \times \mathit{NS}_i + \beta_s \times \mathit{SLM}_{iacg} \times \mathit{SS}_i + \epsilon_{iacg} \tag{3}$$

where the *share of students like me* (SLM_{iacg}) is the share of same gender classmates that student i experiences in class g, course c and academic year a. *stereotypical selection* (SS_i) and *neutral selection* (NS_i) are a dummy equal to one if the student is enrolled in the top 2,3,and 5 departments with the highest share of same gender students among undergraduates, and a dummy equal to one for students who are not enrolled in neither top 2,3, and 5 nor bottom 2,3, and 5 departments respectively. This way, β_c provides us with an estimate of the effect of a change in the share of same gender classmates for students enrolled in counter-stereotypical departments, while β_n and β_s provide us with a measure of the extent to which the effect of a change in gender composition of classmates is different if a student is enrolled in a "neutral" or stereotypical department. By estimating the effect with a progressively more restrictive measure of stereotypical and counter-stereotypical choices, this measure allows to capture potential non-linearities in the effect. The standard errors are clustered at the class level.

IV.C. Variation in the share of same gender classmates

We might be worried that the random allocation of students into classrooms creates a problem of small variation in classroom composition that could potentially lead to overestimating the effect in an analogous way as a problem of weak instrument in instrumental variable regressions (Angrist, 2014).

This is not a concern in this setting. First, women represent a substantial share of the undergraduate population - around 50%. Second, first-year courses are big - the average course size is 138 students and the average course has 9 classes. Third, as several first-year courses are in common across departments, students can be assigned to classes with students that do not belong to their program of enrolment - the average share of same program peers is equal to 0.49. Lastly, the allocation of students to classes is independent on the program of enrolment and is random conditional on scheduling constraints.[36] Taken together these characteristics of the setting imply that students who are in the same class in a course are not necessarily allocated to the same class in other courses, allowing to exploit a variation in the share of same gender

[35] When I consider top and bottom 5 departments, I am considering the top and bottom tercile of departments given that the total number of departments in 16.

[36] Once students chose their program of enrollment, they attend the courses that are scheduled for their first year, and in each of them, they are assigned to classes considering only scheduling constraints.

classmates that goes beyond the variation within the program of enrollment.

The variation in class gender composition is substantial, as shown in Table III. the average share of same gender classmates is 0.561, with a standard deviation of 0.162. Furthermore, the within student standard deviation in the share of same gender classmates is 0.112. Hence, controlling for course and students fixed effects, the residual variation represents 68.75% of the overall variation.[37] Figure V and Table III also present descriptive evidence on this. Figure V shows the between-students across-courses variation and the within-student across-courses variation in the share of same gender classmates, which is obtained by taking the difference between the highest and lowest share of same gender classmates for each student.

Looking at Figure X and Figure XI, we can also see that both the within-student variation in the share of same gender classmates and the share of same gender classmates for students enrolled in stereotypical and counter-stereotypical departments share significant common support. This allows to assess whether the effect of changes in class composition differs for students who selected into stereotypical or counter-stereotypical majors.

IV.D. Validity of the identification strategy

Specification 1 allows to identify the effect of classmates' gender composition on students' performance β, exploiting the variation from the average share of same gender classmates that students experience in the courses they attend, where they have been assigned to classes g with exogenous peers' characteristics. This identification strategy relies on the following assumptions: (i) students are assigned to classes in a way that is orthogonal to their individual characteristics, conditional on scheduling constraints; (ii) there are no factors that are correlated with class gender composition that differentially affect men and women; (iii) scheduling constraints are exogenous: men and women do not systematically select different courses.

In this section, I discuss the evidence concerning the validity of these assumptions. To support assumption (i), I show that students with different characteristics are not systematically assigned to particular classes. In this direction, I provide evidence that class group allocation does not predict students' individual characteristics, and being a woman does not predict the share of same gender peers in class. Regarding assumption (ii), I perform a series of balance checks to assess whether the variation in the share of same gender classmates is related to the variation in predetermined student characteristics. Lastly, corroborating the validity of assumption (iii), I demonstrate that selection into courses does not systematically differ for men and women. In order to do so, I perform a permutation test where I simulate 1000 unconstrained random allocations and test that the distribution of the variation in the observed share of female students across classes is not statistically different from the distribution obtained with the simulated allocation.

Test 1: Class group allocation does not predict students' characteristics Following Feld and Zölitz (2017) and Braga et al. (2016), I test whether class group allocation predicts students' characteristics. I follow the

[37] Following Olivetti et al. (2020)

specification below:

$$y_{ig} = \sum_{g=1}^{n_g} \alpha_g \times G_{i,g} + \sum_{p=1}^{n_a} \gamma_p \times PC_{i,p} + \epsilon_{ig}, \; \forall a, c \tag{4}$$

where the dependent variables are pre-determined individual characteristics - gender, ethnicity, age at enrollment - of student i enrolled in course c in year t, who has been allocated to class g. The explanatory variables are dummies for each class group g in course c in academic year a. The dummy for class group g is equal to one if student i is assigned to class group g and zero otherwise ($G_{gi} = \mathbb{1}(i\text{'s } group\text{=}g)$). The allocation is constrained by the fact that students that attend more than one course in the same term cannot attend two classes at the same time. To take care of this, I control for a dummy for each course that each student takes during the academic year ($\sum_{p=1}^{n_a} PC_{i,p}$). I run one regression for each combination of course $c \times$ academic year a.[38] The sample for each regression consists of all the students enrolled in course c in the same academic year a that didn't change class group during Michaelmas term.[39] Furthermore, the sample is restricted to all the courses that have at least 2 class groups.

I test that class dummies are jointly significantly different from zero:

$$H_0 : \alpha_g = 0, \; \forall g = \{1; n_c\}$$

i.e. students who take the same combination of courses are allocated to classes independently of their individual characteristics.

According to Murdoch et al. (2008), the p-values of this test should be uniformly distributed with mean 0.5 if class allocation is random. In this case, courses have different sizes, the number of classes is not fixed, and classes belonging to the same course might be of different size. In order to check that also in this case the p-values converge to a uniform distribution with mean 0.5, I Perform a Monte-Carlo simulation where I randomly allocate to class groups the students enrolled in first-year courses, that didn't change class group during the term. I perform this 1000 times, under the assumptions that course size, number of seminars and class size is equal to the observed one. Figure A.6 shows the result of the simulation. Indeed p-values converge to a uniform distribution with mean 0.5. This represents the reference point against which I will check under which conditions the observed allocation of students can be considered as good as random.

The results of the test are consistent with conditional random allocation. Figure VI shows the p-values obtained from the tests of joint significance of the class dummies for the observed sample. For the regressions regarding gender and age at entry, less or equal to 5% of tests display a p-value smaller than 0.05, and less than 10% of tests display a p-value smaller than 0.10, in line with what we would obtain if students were randomly allocated to classes. Regarding ethnicity, slightly more than 5% of tests display a p-value smaller than 0.05, but less than 10% of tests display a p-value smaller than 0.10 when we test whether class composition can predict the probability of being white. Regarding being Asian, slightly less than 5% of tests

[38] This regressions include all first year courses and 15 second year courses, among which 7 are language courses. Students can choose second year courses as elective courses. These correspond to 0.57% of the observations

[39] I exclude students that changed class group since I can't observe their initial allocation. Tests of randomness of the decision to change class can be found in the appendix.

display a p-value smaller than 0.05, and less than 10% of tests display a p-value smaller than 0.10.[40] More details on the randomization tests and the simulation performed can be found in appendix.

Test 2: Being a woman does not predict the share of same gender classmates Following Feld and Zölitz (2017), Brenøe and Zölitz (2020), and Guryan et al. (2009), I test that female students are not systematically assigned to classes where there are more same gender peers, conditional on a given share of same gender peers in the course. The specification used is the following:

$$SGP_{-i,acg} = \alpha_{ca} + \beta \times F_i + \gamma \times SGP_{-i,ca} + \sum_{a=1}^{n_a} \sum_{p=1}^{n_{p,a}} \delta_{p,a} \times PC_{i,ap} + \epsilon_{iacg} \tag{5}$$

For each course c in academic year a, I test if being a woman (F_i) predicts the gender composition of peers in the class g (SGP_{iacg}). I control for the share of same gender peers enrolled in the course ($SGP_{-i,ca}$) to clean for the fact that if there are more women than men enrolled in the course, the probability of being in a class group with students of the same gender is higher for female students than for male students. Lastly, I control for a series of dummies equal to one for all the students that attend the course in the same year, and zero otherwise ($\sum_{a=1}^{n_a} \sum_{p=1}^{n_{p,a}} PC_{i,ap}$), as the allocation of students to classes is random conditional on scheduling constraints. I perform this test for all the courses where there are at least two classes, and standard errors are clustered at the class level.

The results of the above regression confirm that students are not systematically assigned to class group based on their sex. Table IV shows that women do not have a higher probability to be in class with same gender peers than men.

Test 3: Balance checks In this section I provide evidence in support of the assumption that there are no factors that are correlated with class gender composition that differentially affect men and wome. In Table V, I produce an array of "balancing tests" to assess whether the variation in the share of same gender students in the class is related to the variation in an array of predetermined student characteristics. I do so by testing that the share of females in the class is not systematically correlated with ethnicity, previous school characteristics, age at entry, and qualification score at entry.

The results corroborate the assumption that the variation in class gender composition is not systematically related to the variation in other peers' characteristics. As shown in Table V, only one of the estimated correlations appears to be significantly different from zero at 10% level of significance for the sample of analysis. This represents 7% of the tests performed. As expected when running a large number of regressions testing multiple hypotheses, some coefficients are statistically significant. In the absence of a systematic relation between the share of female classmates and other individual characteristics, we would expect 10% of coefficients to be statistically significant at the 10% significance level, 5% at the 5% level,

[40] I aggregated Chinese, Indian and other Asian students in a unique category called Asian in order to have a big enough sample to be able to perform meaningful tests. We need a big enough sample of students that share characteristic t to be able to test if class allocation is as good as random. As a matter of fact, if the number of students is too small, even if students are allocated randomly, class allocation might still predict students' characteristics. To illustrate this, let's consider the extreme case in which there is only one Chinese student enrolled in the course, class assignment will predict students' ethnicity even if the Chinese student is allocated randomly to the class.

and 1% at the 1% level simply as a result of chance. This is consistent with the results found, providing supportive evidence that the estimated effects are due to a change in class group composition along gender lines, rather than other unobserved factors that are correlated with gender.

Test 4: Exogeneity of scheduling constraints The tests presented above provide evidence that students are as good as randomly allocated to class groups, conditional on their course choices. In this section, I provide evidence that the limited course choices, that fist-year students are allowed to make after choosing a major, are not gender specific.

Following Lavy and Schlosser (2011), I perform a permutation test where I simulate 1000 random allocations of students to classes. For each first-year course, I randomly allocate to classes the students enrolled in the course that didn't change class during the first term, under the assumptions that course size, number of seminars, and class size is equal to the observed one. This is an unconstrained random allocation, that does not take into account the constraints deriving from the fact that students cannot be allocated to certain classes due to the clashes with the other courses they attend during the same academic year. As such, this represents a test for the assumption that students of different gender do not select systematically different courses among the first-year optional courses.

For each course c in academic year a, I compute the within-course variation in the share of female students (WCV_{gca}) across classes for both the observed sample allocation and the 1000 simulated random allocations:

$$WCV_{gca} = \textit{Class share of female students}_{gca} - \textit{Course share of female students}_{ca},\ \forall g, c, a, t$$

I then perform a Two-sample Kolmogorov-Smirnov test to test the null hypothesis that the observed distribution of the statistics ($Diff_o$) and the distribution of the statistics for the randomly allocated simulated samples ($Diff_s$) are not significantly different:

$$Ho : Diff_o = Diff_s$$

The results of the test are shown in Figure VII. We fail to reject the null hypothesis that the distribution of the observed statistics and the distribution of the simulated statistics are different at every significance level (1%, 5%, and 10%).

In addition, these results are confirmed also when I perform a Two-sample Kolmogorov-Smirnov test to test the null hypothesis that the observed distribution of the within-course standard deviation in the share of females and the randomly simulated distribution are not significantly different. Figure VIII shows that the observed within-course variation in the proportion of females across classes resembles the variation that would result if the gender composition of each class was randomly generated.

V. MAIN RESULTS

V.A. The impact of selection on the effect of peers' identity on performance

Table VI displays the results of Specification 2 in (Column 1) and Specification 3 in Columns (2)-(4). The effect of an increase in the *Share of students like me* for students who made different choices of major is presented in Table VII. An interesting pattern emerges.

Class gender composition affects students' performance in the course. However, the effect is not the same for every student. Selection plays a significant role. Column (1) shows that the coefficient of the *Share of students like me* is negative and significant, while the coefficient of the interaction *Share of students like me* $\times$ *Stereotypical Selection* is positive and significant in all specifications. This indicates that the students who benefit the most, in terms of the final performance in the exams, when surrounded by same-gender peers are students who chose to enroll in programs in line with stereotypes regarding gender skills and roles (women in humanities and men in math-intensive fields).

These patterns are confirmed when we use the categorical definition of stereotypical selection. Columns (2) - (4) in Table VI show that the impact of a change in the share of same-gender classmates decreases (almost linearly) when we use a less strict definition of stereotypical and counter-stereotypical choices. This indicates that the pattern highlighted in Column (1) is not the result of outliers or students enrolled in particular departments.

What this exactly implies in terms of direction, magnitude, and significance can be inferred from Table VII. Students who made different choices of major are affected in a significantly different way by peers' identity when they are in the classroom. For students who are enrolled in stereotypical departments (men enrolled in math-intensive majors, and women enrolled in humanities and international relations), a 10% increase in the share of same-gender classmates increases course grades by 0.325 points, equivalent to 2.0% of a standard deviation (Column 3). On the other hand, for students enrolled in counter-stereotypical departments (women enrolled in math-intensive majors, and men enrolled in humanities and international relations) a 10% increase in the share of same-gender classmates decreases final course grades by 0.288 points, i.e. 1.76% of a standard deviation (Column 3), albeit the effect is less robust. Lastly, the performance of students who are enrolled in departments with a balanced gender ratio is not affected by class composition.

If we consider that stereotypical choices of major imply selecting a program where the student's gender is majoritarian, these results indicate that students who chose to enroll in departments where being in the majority is the norm (e.g. men in math) suffer from being under-represented once in the class. Students who chose to enroll in departments where being in the minority is the norm (e.g. women in math) perform better when allocated to classes where their gender is under-represented. Those students who chose to enroll in departments with a balanced gender composition are not affected by the composition of the class.

The magnitude of the effect in absolute value is small. However, it has important implications. First, it confirms that even in a setting as competitive and selective as the LSE, classmates' identity affects students' performance. In particular, the magnitude of the effect is in line with what other studies estimated in other higher education settings (e.g. Zölitz and Feld, 2021 in the School of Business and Economics at Maastricht University, and Booth et al., 2018 at the University of Essex). This is surprising considering that the treatment is small – 1 hour of interactions per week, and we would expect highly motivated and selected students to be fully rational. Instead, they are subject to social biases, and their performance is influenced by the composition of the environment.[41]

Second, this result significantly diverges from what we would have predicted if we had generalized the findings of experiments where selection has been shut down by design. According to the literature, stereotypes exercise a detrimental effect on performance when they are made salient in the mind of a person who belongs to a group that is negatively stereotyped (Ellemers, 2018; Spencer et al., 2016; Steele and Aronson, 1995). These are, in this setting, students who are enrolled in counter-stereotypical majors (men in humanities and women in math-intensive fields), who are surrounded by different gender peers in the classroom. Hence, we would expect to see that these students would benefit the most by being surrounded by same-gender classmates. However, the evidence in Table VI points to the opposite direction. The performance of students who made a counter-stereotypical choice of major is not hindered (if anything is boosted) when they are in the minority in the classroom.

What part of the distribution does the effect come from? Table VIII provides evidence of the effect at different points of the grade distribution. Course grades at the University range from zero to one hundred. I define a series of dummies that are equal to one if grades are greater or equal to 40, 50, 60, and 70, respectively. These represent important thresholds for students. A student is deemed to have failed a course if his grade is below 40. The student has achieved a third-class honor if the grade is between 40 and 49, a lower second-class honor if the grade is between 50 and 59, an upper second-class honor if the grade is between 60 and 69, and a first-class honor if the grade is 70 or above. Lastly, I define a dummy equal to one if students drop out from the course and zero otherwise.

Class composition does not affect the probability of dropping out of the course, as displayed in Column (1). This is different from what has been estimated by other papers, that found in different settings that the composition of the environment affects persistence within the course and the major (Shan, 2020; Zölitz and Feld, 2021; Hill, 2017; Griffith, 2010). It is however not surprising in this context. The drop-out rate from first-year courses is low (2.9%). As mentioned in Section II., students can only apply to 5 programs across all UK universities, and the LSE rejection rate is as high as 75% every year. Hence, students who apply at the LSE are likely highly motivated. Furthermore, as only the "year-one average" - calculated by adding together and averaging the best six out of eight grades in a first-year course - counts for the final classification grade, students might decide to take the final exam even if they do not expect to perform well.

Regarding the decomposition of the average estimated effect, Columns (2) - (5) show that having same-gender students in the classroom lifts the performance of students who made stereotypical choices of major at every point of the grade distribution. However, the effect increases when we move toward the top of the distribution. A 10% increase in the share of same-gender students in the classroom increases the probability of getting a distinction in the exam by 0.62 percentage points. Considering that 23% of students on average

[41] One caveat is that first-year courses are not high-stake courses. The "year-one average", calculated by adding together and averaging the best six out of eight grades in a first-year course, counts for only one-ninth of the final classification grade. Hence, the results might be different if performance in second or third-year courses had been studied.

get a distinction, this represents a 2.67% increase in the probability of getting a distinction ($= 0.62/0.232$).[42] On the other hand, a higher share of same-gender students in the classroom reduces the performance of students who made a counter-stereotypical choice of major by preventing students from achieving a first and upper-second-class honor. It does not have an impact on the probability of getting a distinction, and it has a small but not robust effect on the probability of passing the exam.

Gender differences in the effect of stereotypical selection Is the effect of stereotypical selection different for men and women? To answer this question I estimate Specifications 2 and 3 by interacting the share of students like me and the share of students like me $\times$ stereotypical selection with a dummy equal to one if the student is a woman and a dummy equal to one if the student is a man. Table IX displays the results.

I find that men are more sensitive to the composition of the classroom compared to women. The coefficient of the interaction between the share of same-gender classmates and stereotypical selection in Column (1) is positive and significant for men, indicating that those men who chose majors in line with stereotypes benefit more from having classmates of their gender compared to men who made choices against stereotypes. The effect for women is smaller in magnitude and not statistically significant. Moving to Columns (2), (3), and (4), however, we can notice that when we define the measure of stereotypical selection to include only departments that are very male and female-dominated (bottom and top 2 and 3), the coefficients become significant also for women and not statistically significantly different from the effect estimated for men using the same definition of stereotypical selection. This indicates that both men and women are affected by the composition of the environment, but men's performance is more strongly affected compared to women's performance since the effect is significant and strong also for students who didn't enroll in extreme departments in terms of gender composition.

Columns (2)-(4) also provide a quantifiable estimate of the effect of being in a minority on performance for men and women who are enrolled in female and male-dominated fields. Let us consider Column (3), for example. For men who are enrolled in male-dominated fields (math-intensive - stereotypical selection), a 10% increase in the share of men increases course grades by 0.376 points, i.e. 2.3% of a standard deviation in course grades. On the other hand, women who are enrolled in male-dominated fields (math-intensive - counter-stereotypical selection), do not suffer from being surrounded by men. As a matter of fact, a 10% increase in the share of women (equivalent to a 10% increase in the share of men) decreases course grades by 0.234 points, i.e. 1.43% of a standard deviation in course grades. Symmetrically, for women who are enrolled in female-dominated fields (humanities-related - stereotypical selection), a 10% increase in the share of women increases course grades by 0.217 points, i.e. 1.33% of a standard deviation in course grades. On the other hand, men who are enrolled in female-dominated fields (humanities-related - counter-stereotypical selection) do not suffer from being surrounded by women, as a 10% increase in the share of men (equivalent to a 10% increase in the share of women) decreases course grades by 0.435 points, i.e. 2.67% of a standard deviation in course grades.

To the best of my knowledge, no other study before looked at whether the effect of peers' identity differs depending on whether individuals made occupational/educational choices that confirm or challenge the

[42]The same calculation would deliver an increase of 0.4% in the probability of passing the exam, 0.9% increase in the probability of getting at least a lower second pass honor, and 1.48% increase in the probability of getting at least a first-class honor.

societal expectations (stereotypes) associated with this identity. However, these results are not new. Oosterbeek and van Ewijk (2014), after randomly allocating first-year undergraduate students in economics and business at the University of Amsterdam to class groups, find that males in work groups with a high share of females do worse in courses with a high math content (mathematics and statistics). Moreover, Zölitz and Feld (2021) studies undergraduate students in the School of Business and Economics at Maastricht University and finds that an increase in female classmates reduces women's probability, but increases men's probability of choosing a male-dominated major. In line with the results of this paper, they also find that the students whose grades are affected by a change in the gender composition of the classroom are women in non-mathematical courses and men in mathematical courses (albeit men in mathematical courses perform better when surrounded by more women). Lastly, Hill (2017) finds that an increase in female peers is associated with a reduced likelihood of graduating in Business for males and suggestive evidence that females are less likely to complete STEM degrees (conditional on graduating) when exposed to cohorts with greater female shares.

Reallocation policy - counterfactual scenario and implications These findings taken together indicate that the composition of the environment can explain the performance gaps observed across fields. However, the effect is due to the (stereotypical) majority group benefiting from being over-represented, rather than the (stereotypical) minority group suffering from being under-represented. As such, they have important policy implications. To illustrate this, I simulate what would happen to students' performance in the counterfactual scenario of a reallocation policy that enforces a more balanced gender ratio in male-dominated fields. In male-dominated fields[43], the average share of women enrolled in undergraduate programs is 35.6%. The gender gap in performance in first-year courses is negative: male students perform better than female students enrolled in the same course by 2.43 points on average. The estimates indicate that a 10% increase in the share of women in the class, holding the characteristics of students enrolled in these departments constant, reduces the gender gap by 5.9%.[44] Hence, enforcing a more balanced gender ratio equalizes performance across genders. However, this effect is the combination of two negative effects: a negative effect for men (equivalent to 15.5% of the gap), and a negative effect for women enrolled in these fields (equivalent to 9.6% of the gap). As the only students who benefit from being surrounded by same-gender peers are students who made choices of major in line with stereotypes and belong to the majority group (in this case men), enforcing a more balanced gender ratio in male-dominated fields reduces differences in performance across groups. However, at the same time, it lowers average performance, as neither men nor women benefit from a higher share of female classmates.

[43]For this exercise I consider male-dominated departments the 3 departments with the highest share of men. This choice does not undermine results and implications, which are robust to the definition used.

[44]This is obtained by summing the effect for men (0.376/2.43=15.5% of the gap) and women (0.234/2.43=9.6% of the gap). If we re-weight the effect by the share of men and women in the field we obtain a reduction of 4.03% of the gap (0.376/2.43*0.544=8.42% of the gap for men, while 0.234/2.43*0.456=4.39% of the gap for women)

V.B. The importance of choices and stereotypes

The main challenges to interpreting the results presented in the previous section as the interplay between selection and the effect of peers' gender are two. First, the variation in peers' gender might be correlated with differences in other characteristics of peers, programs, teachers, or other potential confounders. Second, choices of major may be important for reasons other than the fact that they carry information regarding stereotypes. In order to do so, I start by providing evidence that the results are not driven by differences in support in the share of same gender classmates and within-student variation in the share of same gender classmates. Second, I show that the results are not driven by differences in teaching assistants or peers' characteristics for students enrolled in different departments. Third, I test whether the results could be explained by spillovers across courses or mechanical effects. Lastly, I provide evidence that the heterogeneity in effect for students enrolled in different departments is related to patterns of selections linked to stereotypes regarding gender skills and norms.

Differences in the strength of numerical minority across departments In Section IV.B., I provided evidence of significant common support in the share of same gender classmates and within student variation in the share of same gender classmates across students enrolled in different departments. Despite the significant common support, the variation in the share of same gender classmates is different for students enrolled in different departments, and is higher for students who are enrolled in stereotypical departments. This could be able to explain the different effects of class composition estimated for students enrolled in different departments if the effect of being in a minority is non linear. Figure XII shows that there is no evidence of non-monotonicity in the effect of a change in the share of same gender classmates. The estimates of a local polynomial regression are plotted for each value of the share of same gender classmates[45]. This provides evidence in support to the fact that the results cannot be explained by potential differences in the effect of being in a minority across departments due to students being exposed to a different variation in the share of same gender classmates or a different overall share of same gender students.

Difference in teachers or other peers' characteristics correlated with gender Students enrolled in different departments might be exposed to different teachers or peers with different characteristics. In particular, as there are more female students in majority-female departments, there might be more female teachers in these departments. This could be problematic if female teachers treat women and men differentially. Furthermore, academic regulations are different across programs: the number of compulsory courses are different and the same is true for courses that students can attend outside the program or department of enrollment. Or else, math-intensive departments might attract students with different characteristics with respect to disciplines that are related to Humanities. Lastly, peers' gender might be related to other peers' characteristics, that might induce us to erroneously attribute the estimated effects to the gender composition of the class when it is instead due to other peers' characteristics (Bramoullé et al., 2020; Angrist, 2014; Manski, 1993).

In order to explore if this affects the results, I estimate Specifications 2 and 3 controlling for teaching

[45]The Figure displays residual course grades and share of same gender classmates obtained by regressing them on course $\times$ year fixed effects and student fixed effects as in Specification 1

assistant fixed effects, share of same ethnicity and same background students (defined as students coming from State funded or Independent schools), and share of same program students. The results are robust to the introduction of these controls.

Ability peer effects or social identity considerations? Another characteristic that might make peers of students enrolled in male or female-dominated departments heterogeneous is ability. Selection and admission processes are different across programs. Thus, students who are enrolled in different departments might be exposed to peers with different ability. A substantial body of literature provides evidence that indeed peers' ability matters for performance (recently for example Feld and Zölitz (2022); Fischer (2017); Feld and Zölitz, 2017; Sacerdote, 2014; for a review of previous studies, Sacerdote, 2011).

Interpreting the estimated effects as the interaction between choices of major and the effect of belonging to the numerical minority becomes problematic if ability is correlated with gender. As relatively more men apply to male-dominated departments than female-dominated departments, the distribution of ability for men and women accepted in male-majority and female-majority undergraduate programs might significantly differ. If this is the case, changes in class composition along gender lines might affect performance by changing the ability of peers. Furthermore, if the identity of the high skill group (men vs women) changes across departments due to selection patterns and admission processes, ability peer effects could also explain why the effect of changes in class composition along gender lines is heterogeneous across departments of enrollment.

In order to address this potential concern, I estimate Specifications 2 and 3 controlling for average peers' ability, defined based on qualification scores at entry.[46], The sample is restricted to students I had admission information for, and for whom I was able to reconstruct a qualification score at entry based on A-level results or equivalent. Results can be found in Column (5) Table X. The results indicate that ability peer effects and the ability of peers changing along gender lines cannot explain the estimated patterns.

Spillovers across courses and mechanical effects The estimated effects could be overestimated in the presence of spillovers across courses. For instance, being in a minority in a course might induce students to perform worse in the course, but also to perform better in other courses where the share of students of their gender in the class is higher. This could rationalize the stronger effect for students enrolled in stereotypical departments considering that the within-student variation in the share of same gender classmates for these students is greater than the within-student variation for students enrolled in counter-stereotypical departments.

In order to test if this is the case, I re-estimate the effects by comparing the performance of students who are enrolled in the same course, but have been as good as randomly assigned to classes with a different share of same gender classmates - "within-course variation". This identification strategy exploits the exogenous allocation of students across classes within the same course conditional on scheduling constraints. Under the assumption that students are assigned to classes in a way that is independent on their assignment in another courses after controlling for scheduling constraints, the effect estimated by exploiting the within-course variation represents a counterfactual to assess the extent to which spillovers across courses induce

[46] As explained in Section III.C.

to overestimate the effect when exploiting the within-student variation. One caveat however has to be kept in mind. Spillovers across courses could also induce us to overestimate the effects when exploiting the within-course variation.[47]

Table XI displays the results of the two specifications. Column (1) displays the result of Specification 2, where I exploit the within-student variation, while Columns (3) and (4) display the results obtained by exploiting the within-course variation. Column (4) exploits one observation per student, which controls for the fact that observations for the same student in different courses are not independent. We can see that the estimates exploiting the within student-variation are smaller in magnitude with respect to the estimated effects obtained by exploiting the within-course variation, confirming that spillover effects are not leading to overestimate the impact of stereotypical selection.

The results in Table XI also confirm that the results are not driven by mechanical effects. Including individual fixed effects is not leading to overestimating the effect due to a mechanical relations between individual observed and unobserved characteristics and class composition (Angrist, 2014). Furthermore, Column (2) displays the result of Specification 2 without controlling for course $\times$ year fixed effects, confirming that the results are not driven by mechanical effects arising because, controlling for course fixed effects forces the average grades in the course to be equal to 0.

Gender stereotypes To support the hypothesis that the results are due to students' choices of major affecting students' sensitivity to peers' gender once in the classroom, I provide three pieces of evidence.

First, I perform a placebo test by estimating specifications 2 and 3 controlling for the interaction between the measure of stereotypical selection and the share of same ethnicity classmates, share of same program classmates, share of same background classmates, and peers' average ability at entry. Table XII shows that the results are robust to the introduction of these controls. Furthermore, none of the coefficients of the interaction between the measure of stereotypical selection and other peers' characteristics are significant, indicating that the coefficient of stereotypical selection is capturing the heterogeneity in the effect of being in a minority caused by students who selected into different fields having different characteristics along gender specific dimensions.

Second, Table XIV shows that the results are robust to other proxies of stereotypical selection related to gender stereotypes: the share of women and men among applicants to undergraduate programs at LSE, the share of men and women enrolled in different subjects in higher education in UK, and the share of men and women among the staff working in each subject in higher education in UK[48].

Lastly, I provide evidence that stereotypes play a role by replicating the analysis exploiting another proxy for the strength of stereotypes for students: the gender of the class teacher. Table XV provides evidence that the results are confirmed when we consider the gender of the class teacher. I replicate the analysis by replacing the share of same gender classmates with an indicator equal to one if the class teacher shares the same gender as the student. The performance of students who are enrolled in counter-stereotypical

[47]For instance, if being in a minority in the Economics course has such a demotivating effect that students perform worse not only in Economics, but also in the other courses they attend, not controlling for student fixed effects would lead to overestimate the effect.

[48]details regarding the measures used can be found in Table XIII.

departments is not affected by the gender of the class teacher, contrary to the performance of students who made a stereotypical choice of major. Since class teachers act as role models, breaking stereotypes regarding gender roles and skills (see for example Breda et al., 2021; Porter and Serra, 2020; Olsson and Martiny, 2018; Carrell et al., 2010), this is an additional piece of evidence that points in the direction of students who are enrolled in counter-stereotypical departments being less affected by gender stereotypes.

Beyond gender-specific dynamics Are these patterns specific to gender or do they concern being in a minority more in general? In order to provide an answer to this question, I replicate the analysis along ethnic lines. The institution is characterized by an exceptionally diverse environment: only 37% of students are White, while 47% of the population of undergraduates are are Asian (primarily Chinese and South Asian). Furthermore, also ethnic groups are distributed unequally across departments. The pattern along ethnic lines is exactly the opposite with respect to the distribution of traits that characterizes departments along gender lines. The fields where ethnic minorities (in this case Asian students) represent the majority of enrolled students are Math and Science-intensive fields, which are also the fields where women are under-represented. As for gender, this reflects stereotypes regarding group-specific skills and roles. As a matter of fact, lab and field experiments in the academic setting document a widespread belief that Asians are better at math and science with respect to white students (see for example the review by Spencer et al., 2016).

I estimate the following specification:

$$y_{iacg} = \alpha_{ac} + \alpha_i + \sum_{e=1}^{3} E_{i,e} \times [\beta_{1,e} \times SLM_{iacg} + \beta_{2,e} \times SLM_{iacg} \times STS_i] + \epsilon_{iacg} \quad (6)$$

where students are divided in three groups $E_{i,e} = \mathbb{1}(\text{i's ethnic group} = e)$, Asian, White, and Other students, which is a residual category that includes other ethnicity students and students who provided no information on their ethnicity[49]. The *share of students like me* (STS_{iacg}) is the share of same ethnicity classmates that student i experiences in class g, course c and academic year a, and *stereotypical selection* (STS_i) is the average share of same ethnicity students in student i's department of enrolment across academic years 2008-2017. Figure XX provides a graphical description of stereotypical selection along ethnic lines. The average share of same ethnicity students in the department of enrolment for the residual group is calculated as one minus the share of Asian and White students enrolled in each department[50]. $\beta_{2,e}$ provides an indication of the extent to which being surrounded by same ethnicity classmates has a different effect on performance for students who made choices that are more in line with stereotypes regarding ethnic specific skills and norms with respect to students who made a choice against stereotypes, for each ethnic group e. The standard errors are clustered at the class level.

This exercise is based on the assumption that students are not systematically allocated to different classes

[49]This categorization is motivated by the necessity of having large enough groups to perform the analysis. Table A.15 in Appendix E. provides evidence that pulling students coming from different Asian countries together is not such a strong assumption given that the performance of Asian students coming from different countries is affected by an increase in the share of classmates coming from other Asian countries in the same qualitative way as an increase in the share of Asian classmates who were born in the same subset of Asian countries (dividing students in Chinese, India, Other Asian).

[50]The results are robust to restricting the sample to only Asian and White students (Table A.1 for evidence)

because they belong to particular ethnic groups. I provide evidence for the validity of this assumption by replicating for ethnicity the same tests performed for gender. Evidence and a detailed explanation on the results can be found in Appendix E..

Table XVI displays the results of this exercise. Stereotypical selection moderates the effect of being in a minority also along ethnic lines. Both White and Asian students benefit from being surrounded by more classmates of their ethnicity when they made a stereotypical selection of major. This is robust to controlling for the share of same gender classmates (Column 2). These results confirm that the estimated patterns concern the interaction between being in a minority and selection into fields, and are not only related to gender-specific dynamics.

VI. MECHANISMS

How does selection interplay with the effect of peers' identity on performance? I provide an answer to this question by building a theoretical framework. I embed the key channels indicated by the experimental literature as the main reasons behind the effect of changes in peers' identity on performance in a theoretical framework to rationalize how class composition affects students' performance in exams. I derive predictions regarding the effect of relaxing minority status for LSE students in the absence of selection. I then enrich the framework to show how considering selection can modify these predictions and explain what we observe in the data.

VI.A. Theoretical Framework

Students are risk neutral and choose educational investments (e_i) to maximize
$U(e_i, a_i^b, a_{-i}^b, g, s_g, \beta_i^T, \beta_i^S)$

$$\max_{e_i} \quad \overbrace{\beta_i^P[f(a_i^b, e_i)]}^{\textbf{Course Performance}} + \overbrace{\beta_i^S g(e_i, a_i^b - a_{-i}^b, g)}^{\textbf{Image concerns}} - c(e_i, \delta_i s_g) \tag{7}$$

Students' choice of effort depends on two factors (following Bursztyn et al., 2019 and Ashraf et al., 2014: (i) learning motives -$f(.)$, students benefit from investing more effort since they want to maximize their course performance, and (ii) image concerns -$g(.)$, since students care about appearing smart in front of their peers. The latter will be more important (higher β_i^S) when effort is visible to the peers. In the case of undergraduate students at LSE, an example of public investments in education is participation in class. For instance, students might refrain from participating to avoid appearing not smart enough in front of their peers, or they might participate only to appear as smart[51]. The classical assumptions of returns to schooling are assumed: $f_a > 0$, $f_{aa} < 0$, $f_e > 0$,$f_{ee} < 0$, thus course performance is increasing and concave in ability and effort. In the same way, I assume that the utility that students derive from social image is increasing and concave in effort: $g_e > 0, g_{ee} < 0$. Peers' ability enters students' utility function because students care about being of higher ability than their peers rather than having higher grades per se (Ashraf et al.,

[51]Evidence regarding the importance of image concerns for participation is discussed in the section VI.B.

2014): $g(a_i^b - a_{-i}^b > 0) > 0$, $g(a_i^b - a_{-i}^b < 0) < 0$, $g_{a_i^b - a_{-i}^b} \geq 0$[52]. The benefit of effort in terms of image is higher the higher is students' relative ability: $g_{e,a_i^b - a_{-i}^b} \geq 0$. This implies that students of higher perceived relative ability get higher utility from image if they make higher investments in effort, and that conditional on investing in effort students with higher perceived relative ability get higher utility. This is a key assumption of the model and evidence for its validity can be found in Section VI.B.. Lastly, following Spence (1974), the cost function is convex and increasing in effort: $c_e > 0$ and $c_{ee} > 0$.

According to the evidence of lab and field experiments in psychology, economics and sociology, we can expect peers' identity to affect students' performance via two main channels. First, minority status can be expected to impact students' performance is by affecting expectations regarding returns and confidence about personal achievement through its effect on beliefs about self and others' ability. Previous work in psychology, sociology and economics suggests that differences in performance can be partially explained by biased beliefs regarding own-self and others' ability caused by stereotypes (e.g. Bordalo et al., 2019; Spencer et al., 2016; Coffman, 2014). For instance, Bordalo et al. (2019) shows that people tend to overestimate their own ability and the ability of other people of their gender in categories that are judged to be stereotypically congruent with their group, and underestimate their own ability and the ability of other people of their gender in categories that are judged to be stereotypically in-congruent with their group. The effect is stronger when individuals are paired with people of the other gender (Bordalo et al., 2019) or when individuals are in groups where their gender is over-represented (Karpowitz and Stoddard, 2020; Born et al., 2020; Chen and Houser, 2019). As a matter of fact, stereotyping is exacerbated by minority status, which reminds people about their social identity and their belonging to the group that makes them distinct (Hoff and Pandey, 2006; Inzlicht et al., 2006).

I model beliefs on ability building on the framework of Bordalo et al. (2019). Students hold imperfect information regarding their ability and the ability of their peers. Beliefs about their and their peers' ability in the subject rely on stereotypes and are affected by stereotypical distortions.

$$a_i^b = A_g + \mu_i + \theta_i \sigma(s_g)(A_g - A_{-g}) \tag{8}$$

$$a_i^b - a_{-i}^b = (A_g - A_{-g})(1 - s_g)(1 + 2\theta_i \sigma(s_g)) \tag{9}$$

A_g is the average ability of group g, and μ_i is individual specific ability, such that $E_i(\mu_i) = 0$ and $E_i(A_g + \mu_i) = A_g$. Stereotypes driven distortions are defined as $\theta_i \sigma(s_g)(A_g - A_{-g})$. The distortion in beliefs that derive from stereotypes contains a kernel of truth since stereotypes exaggerate true differences in ability between groups. If a student belongs to the "better" performing group, stereotypical distortions will lead him/her to overestimate his/her own absolute ability. On the contrary, if a student belongs to the "worse" performing group, stereotypical distortions will lead him/her to underestimate his/her own absolute ability. The size of the bias depends on how much stereotypes are top of mind ($\sigma(s_g)$) for the student. If stereotypes are top of mind, differences in ability between the two groups are exaggerated. Changes in class composition change the extent to which stereotypes are salient in the students' minds. In particular, stereotyping is exacerbated by minority status, $\sigma_{s_g} \leq 0$, i.e. the extent to which stereotypes are top of mind

[52]This is a framework where social comparison enters additively in the utility function (Ashraf et al., 2014; Kandel and Lazear, 1992)

for a member of group g is decreasing in the size of his/her group s_g.[53].

The second channel through which class composition affects students' choices of effort is through the cost of effort. Experiments in psychology show that people prefer to affiliate with those who share their attitudes and beliefs or demographic traits. We have positive affective responses for those who are similar to us, and we also expect increased comfort and trust when interacting with them (Inzlicht and Good, 2006). So, minority status affects performance by reducing opportunities for social interactions and mutual academic assistance, which increase the marginal cost of effort, $C_{s_g} \leq 0$ and $C_{es_g} \leq 0$.[54]

I assume that individuals are heterogeneous along two key dimensions: the strength of stereotypical associations θ_i and their cost of interacting with students of the opposite gender δ_i. The former determines the strength of stereotypical distortions, while the latter determines the extent to which students benefit from being surrounded by students similar to them in terms of marginal cost of effort. These are the dimensions through which stereotypical selection intervenes to moderate the effect of being in a minority on performance.

The effect of changes in class composition on effort in absence of selection Given the assumptions on concavity of g and f and convexity of the cost functions in effort, there exist at least an interior solution such that

$$e_i^* : \beta_i^P f_e + \beta_i^S g_e - C_e = 0 \tag{10}$$

Given the equilibrium first order conditions described above, the effect of a change in class composition is described by the following expression:

$$\frac{\partial e_i^*}{\partial s_g} = -\frac{\beta_i^P \overbrace{f_{ea}}^{?} \frac{\partial a_i^b}{\partial s_g} + \beta_i^S \overbrace{g_{e,a-a_{-i}}}^{\geq 0} \frac{\partial a_i^b - a_{-i}^b}{\partial s_g} - \delta_i \overbrace{c_{es_g}}^{\leq 0}}{\underbrace{\beta_i^P f_{ee} + \beta_i^S g_{ee} - c_{ee}}_{\leq 0}} \tag{11}$$

The effect of a change in class composition on effort depends on three components, a learning component, an image component and a cost of effort component. In order to illustrate how these channels interact in shaping the overall effect, I am going to explain what happens when relaxing numerical minority generates a positive shock on beliefs regarding absolute and relative ability.

(i) Learning component - the effect depends on the sign of f_{ae}, i.e. whether ability and effort are complement or substitute in learning. If ability and effort are complement, students will be induced to invest more in effort, as the positive shock on absolute ability increases the marginal returns of an additional hour of effort on performance in the course. On the other hand, if ability and effort are substitute, students are induced to reduce the amount of effort they invest in the course as they can obtain the same grade with a

[53]More details on the definition and derivations of beliefs on ability can be found in Appendix D.

[54]This assumption implies that students feel more comfortable participating in class when they are surrounded by students that are "similar" to them. This can be due to the fact that we expect increased comfort, trust and positive affective responses from those who are similar to us. But this can also have to do with potential language barriers. Regarding private forms of effort, such as hours spent studying, this might happen because class composition influences students networks formation, and as a consequence their academic support.

lower level of effort.

(ii) Image component - the prediction is straightforward: students invest more in effort since they receive a confidence boost. When students believe they are relatively better than their peers, the benefits of investing in effort deriving from image are higher, as they get higher utility for each level of effort.

(iii) Cost of effort component - relaxing numerical minority will have an effect on the cost of effort, and in particular, it will induce student to invest more when the share of same gender people increases. This is due to the fact that the model assumes that the marginal cost of effort is lower when students expect and/or receive more emotional and academic support since they are surrounded, and can engage, with a bigger pool of students at a lower cost.

Thus, the three channels combined imply that, if ability and effort are complement, a positive shock on beliefs regarding relative and absolute ability induces students to invest more, since image and learning components reinforce each others. On the contrary, if effort and ability are substitute in learning, the effect depends on the weight that learning and image concerns have in the utility function.

Whether students receive a positive or negative shock on beliefs regarding relative ability when relaxing minority status depends on the nature of stereotypes in the field. I assume that men perform better on average ($A_M - A_W > 0$) in stereotypically male departments, while women perform better on average ($A_W - A_M > 0$) in stereotypically female departments, so that stereotypical distortions carry indeed a kernel of truth. This is consistent with what we observe at the LSE and with the definition of stereotypical choices used in the paper.[55]

$$\frac{\partial a_i^b}{\partial s_g} = \overbrace{\theta_i \sigma_{s_g}}^{\leq 0}(A_g - A_{-g}) \tag{12}$$

$$\frac{\partial E(a_i^b - a_{-i}^b)}{\partial s_g} = (A_g - A_{-g})\overbrace{[-1 - 2\theta_i\sigma + 2\theta_i(1 - s_g)\sigma_{s_g}]}^{\leq 0} \tag{13}$$

If students are enrolled in departments in line with stereotypes ($A_g - A_{-g} > 0$), an increase in the share of students of their gender depresses their beliefs regarding absolute and relative ability. An increase in the share of same gender classmates reduces the extent to which stereotypes are top of mind for the student, thus weakening the stereotypical distortions that were leading him/her to overestimate his/her own absolute ability. At the same time, an increase in the share of same gender classmates increases the average perceived ability of peers, depressing the student's perceived relative ability. This for instance would be the case for men (women) in math-intensive (humanities) departments when the share male (female) students in class increases.

On the other hand, if students are enrolled in counter-stereotypical departments ($A_g - A_{-g} < 0$), an increase in the share of same gender classmates improves their beliefs regarding absolute and relative ability.

[55]Figure A.7 displays the raw gender gap in course performance for students enrolled in different departments calculated as the difference in course performance between men and women enrolled in the same course in the same year. Women enrolled in female-majority departments perform better than men, while the gender gap is negative for students enrolled in male-majority departments. The sample contains all the courses attended by students enrolled in undergraduate programs in the department during the bachelor.

As a matter of fact, an increase in the share of same gender classmates, by reducing the extent to which stereotypes are top of mind,reduces stereotype threat, which was leading them to underestimate their own absolute ability. In the same way, an increase in the share of same gender classmates decreases the average perceived ability of the peers, improving the student's perceived relative ability. This is the case for women (men) in math and science-intensive (humanities) fields when the share of women (men) in class increases.

$$\begin{aligned} &\text{if } (A_g - A_{-g}) > 0 \rightarrow \frac{\partial e_i^*}{\partial s_g} = -\frac{\beta_i^P \overbrace{f_{ea}}^{?} \overbrace{\frac{\partial a_i^b}{\partial s_g}}^{\leq 0} + \beta_i^S \overbrace{g_{e,a-a_{-i}} \frac{\partial E(a_i^b - a_{-i}^b)}{\partial s_g}}^{\leq 0} - \delta_i \overbrace{c_{es_g}}^{\leq 0}}{\underbrace{\beta_i^P f_{ee}^e + \beta_i^S g_{ee} - c_{ee}^e}_{\leq 0}} \\ &\text{if } (A_g - A_{-g}) < 0 \rightarrow \frac{\partial e_i^*}{\partial s_g} = -\frac{\beta_i^P \overbrace{f_{ea}}^{?} \overbrace{\frac{\partial a_i^b}{\partial s_g}}^{\geq 0} + \beta_i^S \overbrace{g_{e,a-a_{-i}} \frac{\partial E(a_i^b - a_{-i}^b)}{\partial s_g}}^{\geq 0} - \delta_i \overbrace{c_{es_g}^e}^{\leq 0}}{\underbrace{\beta_i^P f_{ee}^e + \beta_i^S g_{ee} - c_{ee}^e}_{\leq 0}} \end{aligned} \tag{14}$$

The explained mechanisms lead to the following predictions:

Prediction 1 (Counter-stereotypical departments) *Relaxing minority status increases effort for students that are are enrolled in counter-stereotypical departments ($A_g - A_{-g} < 0$) whenever effort and ability are complement, or effort and ability are substitutes but the effect on image and cost of effort prevails on the effect on learning.*

Prediction 2 (Stereotypical departments) *For students enrolled in fields that are stereotypically congruent with the group identity ($A_G - A_{-G} > 0$), relaxing minority status generates an increase in effort for students if the negative effect on image is small and ability and effort are substitute. On the other hand, relaxing minority status generates a reduction in effort if the positive effect on the cost of effort is small and ability and effort are complement.*

It can be seen from equations (14) that the effect of relaxing minority status on effort choices is stronger the stronger is the effect that relaxing minority status has on beliefs regarding relative and absolute ability and the stronger is the effect on the marginal cost of effort. Equations (12) and (13) show that the effect of relaxing numerical minority on beliefs on ability is stronger the stronger are stereotypical associations for the students (θ_i). On the other hand, the strength of the effect of relaxing minority status on the marginal cost of effort depends on δ_i, the individual cost of interacting with students of the opposite gender. Thus, the model delivers a third prediction:

Prediction 3 (Comparative statics) *The effect of changes in class composition on effort choices are stronger the stronger are stereotypical associations for students (θ_i) and the higher is the individual cost of interacting with students of the opposite gender (δ_i).*

The effect of choices of major Recent literature on selection into occupations and majors provide evidence that negative stereotypes regarding group-specific skills affect beliefs about expected returns, driving individuals out of counter-stereotypical occupations (Kugler et al., 2021; Del Carpio and Guadalupe, 2022; Oxoby, 2014). Furthermore, individuals incorporate their preferences toward the fraction of women and men in the occupation when making choices regarding the type of field to specialize in (Pan, 2015; Card et al., 2008). As a consequence, students who decided to enroll into a field dominated by the opposite gender, in spite of the negative stereotypes regarding their group (women in math-intensive fields and men in humanities), might have weaker stereotypical associations and might face lower costs of interacting with students of the opposite gender, being, as a consequence, less affected by minority status.

This can be easily illustrated using a simple Roy model of major choice. Adapting Del Caprio and Guadalupe (2021)'s theoretical framework, let's assume that student i of gender g has to choose between two options, a stereotypical and a counter-stereotypical major, which provide the following returns:

$$\text{Stereotypical major (S) - Returns: } R^S = W^S A_i^S \tag{15}$$

$$\text{Counter-stereotypical major (C) - Returns: } R^C = \frac{W^C A_i^C}{(B\Theta_i)(\Delta_i S_{-g})} \tag{16}$$

The returns depend on the future streams of earnings of the jobs the major offers access to, which depend on the returns to skill (e.g. wage per unit of skill) in each sector (W^C and W^S). Stereotypes and norms prescribe what people are expected to do based on individuals' observable identity, hence choosing a counter-stereotypical major implies a cost of deviating from norms and stereotypes, $B\Theta_i$. This is characterized by two components: an identity cost B, which is higher the more counter-stereotypical is the major, and an individual specific component Θ_i, which depends on how sensitive individual i is to norms and stereotypes. A second component enters the returns for choosing the counter-stereotypical major, $\Delta_i S_{-g}$. In line with the rationalization of students' choices of effort, I assume that students display preferences to engage with same gender peers (homophily). As a consequence, they sustain a cost when they choose a counter-sterotypical major, as the share of same gender peers is lower. This cost is higher the higher is the share of opposite gender peers in the major S_{-g} and the higher is the strength of i's preferences Δ_i.

If we log-linearize and assume log-normality for $a_i^S, a_i^C, \delta_i, \theta_i$, we obtain the following log-linearized returns:

$$\text{Stereotypical major (S) - Returns: } ln(R^S) = w^S + a_i^S \tag{17}$$

$$\text{Counter-stereotypical major (C) - Returns: } ln(R^C) = w^C + a_i^C - \beta - \theta_i - \delta_i - s_{-g} \tag{18}$$

A student chooses the counter-stereotypical major if and only if the returns from it are at least as great as the returns from choosing the stereotypical major. Hence, the probability of choosing the counter-stereotypical major is characterized by the following equation:

$$
\begin{aligned}
Pr(\text{Choosing C}) &= Pr(ln(R^C) > ln(R^S)) \\
&= Pr(w^C + a_i^C - \beta - \theta_i - \delta_i - s_{-g} > w^S + a_i^S) \\
&= Pr(\underbrace{a_i^C - a_i^S - \theta_i - \delta_i}_{h} > w^S - w^C + \beta + s_{-g}) \\
&= Pr\left(\frac{h_i}{\sigma_h} > \frac{w^S - w^C + \beta + s_{-g}}{\sigma_h}\right) \\
&= 1 - \Phi\left(\frac{w^S - w^C + \beta + s_{-g}}{\sigma_h}\right)
\end{aligned}
$$

Having characterised the probability of choosing a counter-stereotypical major, we can derive predictions regarding the conditions under which this probability is higher.

First, keeping everything else constant, the probability of choosing the counter-stereotypical major decreases in the identity cost (β) and the share of opposite gender students (s_{-g}). This is intuitive. As we are assuming that students' display preferences consistent with homophily, and stereotypes prescribe what people are expected to do based on individuals' observable identity, the higher is the identity cost and the share of opposite gender students in the major, the lower will be the perceived returns from enrolling into a counter-stereotypical major.

$$
Pr(\text{Choosing C}) = 1 - \Phi\left(\frac{w^S - w^C + \beta + s_{-g}}{\sigma_h}\right) \Rightarrow \begin{cases} \frac{\partial Pr(\text{Choosing C})}{\partial \beta} < 0 \\ \frac{\partial Pr(\text{Choosing C})}{\partial s_{-g}} < 0 \end{cases}
$$

Second, assuming that students internalize the gender composition of the field and stereotypes when making their choice of major implies that students applying to stereotypical and counter-stereotypical majors are heterogeneous in terms of their sensitivity to stereotypes (θ_i) and their preferences for same gender peers (δ_i). Below I illustrate the case of sensitivity to stereotypes, but an analogous discussion can be made regarding the students' preferences for same gender peers.

The average sensitivity to stereotypes (θ_i) of students who apply to counter stereotypical majors can be characterized as follows:

$$
E(\theta_i|\text{Choosing C}) = E(\theta_i|h > \overbrace{\overbrace{w^S - w^C}^{d} + \beta + s_{-g}}^{z}) = \sigma_\theta \ \rho_{\theta h} \ \lambda(z)
$$

where $\lambda(.) = \phi(.)/(1 - \Phi)$ is the inverse Mills ratio.

This depends on how big the identity cost for the student is, i.e. how strong stereotypes and norms are in the environment the student lives in. The direction of selection depends on the correlation between d, net returns to skills, and θ in the population.

$$
\frac{\partial E(\theta_i|\text{Choosing C})}{\partial \beta} = \frac{\sigma_\theta \sigma_d}{\sigma_h} \left[\rho_{\theta d} - \frac{\sigma_\theta}{\sigma_d} - \rho_{\theta \alpha}\frac{\sigma_\alpha}{\sigma_d}\right] \lambda(z)' \frac{\partial z}{\partial \beta}
$$

Prediction 4 (Stereotypical selection) *Students applying to stereotypical and counter-stereotypical majors* *are heterogeneous in terms of their sensitivity to stereotypes (θ_i) and their preferences for same gender peers (δ_i).*

At least other two concurring explanations are possible. First, once students took the decision to select into a stereotypical (or counter-stereotypical) field, they might be induced to act in a way that confirms their decision to reduce cognitive dissonance (Akerlof and Dickens, 1982; Festinger, 1957). Second, these differences might be due to the particular environment that characterizes male-dominated fields compared to female-dominated fields. In the last part of the next subsection, I discuss the extent to which the evidence is consistent with these two alternative explanations.

VI.B. From Theory to Empirics

In order to test the model predictions, I exploit rich administrative data and a novel survey.

Test of predictions in absence of selection The model delivers a testable prediction regarding the effect of a change in class composition in absence of selection. When social image represents the primary concern and effort is public, the model predicts that relaxing minority status induces students who are enrolled in counter-stereotypical departments to participate more. In this section, I test this prediction by employing data on teachers' evaluation of students' participation in class, a requirement for teaching assistants during the term.

Teachers' evaluations of students' participation in class are ideal to test this prediction for two reasons. First, participation in class can be considered as a form of public effort, since it is observable to peers. Thus, we can expect choices of participation to be motivated primarily by image concerns rather than learning motives. This assumption is supported empirically. Figure XIII displays the relationship between exam grades and participation in class (Panel A), and participation and individual and peers' ex-ante ability (Panels B and C), controlling for student and class group fixed effects. Participation is beneficial for students' learning, since we can see that it is positively correlated with exam grades. However, the students who participate the most are high ability students, and participation appears to be negatively correlated with the ability of the peers in the class group. This is in line with participation in class being primarily motivated by image concerns and in particular with the idea that students care about being better than their peers. This is confirmed also by the evidence gathered through the survey. Figure XIV displays that students are significantly more comfortable answering questions in class rather than asking questions. Furthermore, they admitted that the fear of looking "smart enough" has refrained them from participating at least once. On the contrary they disagree with the statement that they refrained from participating for the fear of appearing "too know-it-all".

Second, teachers' evaluations of students' participation in class represent a good counterfactual to test the generalizability of experimental findings in real-world settings. Lab and field experiments provide clear evidence that, by reducing the daunting effect of negative stereotypes, relaxing numerical minority increases participation and willingness to contribute to discussions for individuals in counter-stereotypical domains (Coffman et al., 2021; Karpowitz and Stoddard, 2020; Chen and Houser, 2019; Bordalo et al., 2019; Coffman, 2014). Hence, I can exploit teachers' evaluations of students' participation in class as a counterfactual to estimate the effect of a change in group composition on participation and willingness to contribute to discussions in a real-life setting where selection against stereotypes plays a role[56].

The evidence indicates that these channels do not play a role in real world environments where selection is plays a role, contrary to the experimental findings. Figure XV shows the results of Specification 3 on participation in class defining stereotypical selection based on the bottom and top 5 departments categorization.[57] There is no evidence of an effect of changes in class composition on participation for students who are enrolled in counter-stereotypical fields. This is despite , and despite changes in class composition affecting

We can exclude that the absence of effect is due to participation being unaffected by class composition, as students who made choices of major in line with stereotypes display significant effects. However, can this zero effect be consistent with the daunting effect of negative stereotypes being so strong that not even a change in class group composition can induce students to participate in class? This does not seem to be the case, as students choosing a counter-stereotypical major appear to be on average more vocal compared to students of the same gender who chose a major in line with stereotypes (Table A.2).

The evidence on the effect of Selection The last prediction of the model, Prediction 4, indicates that if social identity considerations are internalized in choices of major, students who are enrolled in different departments are heterogeneous along traits that determine the effect of peers' identity on performance. In this section, I provide evidence that students are indeed heterogeneous along these traits. In particular, the evidence points towards students enrolled in fields dominated by the opposite gender having weaker stereotypical associations and facing lower costs of interacting with students of the opposite gender, making them more resilient to the effect of changes in peers' identity.

Evidence on the strength of stereotypical associations is provided in Figure XVI, which displays the results of an implicit association test (Greenwald et al. 1998) to elicit students' associations between female-humanistic and male-scientific[58]. A score of 0 indicates no association between male-Scientific and female-Humanistic. A positive score indicates that the student unconsciously associates women with humanities and men with science and math. A negative score indicates the opposite association. Figure XVI shows that students on average associate men with math and science and women with humanities, but this implicit association is significantly stronger for students who enrolled in departments in line with stereotypes.

We might be concerned that these patterns are specific to the sample of students that took part to the survey. Although the respondents might be a selected sample of students not representative of the overall

[56]These data have two caveats. First of all, they are not objective measures of students' participation and performance, but they are grades that teachers give to students. Thus, results can be considered as indicative of class dynamics conditional on the assumption that teaching assistants do not discriminate or evaluate underrepresented groups differently when they are in a minority. Secondly, they only concern Michaelmas term (due to attrition problems, I consider only Michaelmas term grades, not Lent term evaluations). As a consequence, they are able to shed light on the dynamics that characterize classes in the first half of the academic year. This on the one hand represents a limitation, since exam grades are the result of students' effort during the whole academic year. On the other hand, we can argue that first term measures are more indicative of how students' behavior is affected by stereotypical distortions, since this is the first time students meet and interact with each other.

[57]This is the definition used when discussing the mechanisms as it is the least restrictive and most conservative definition of stereotypical and counter-stereotypical choices, and the one that allows me to have a big enough sample size to deter significant effects.

[58]Details on the survey can be found in Appendix F.

population of undergraduate students at the university, this selection bias would be problematic for the interpretation of the results only if it is systematically correlated with choices of major and implicit attitudes. In particular, for the selection bias to lead to an overestimation of the effects, there should be a positive systematic correlation between the probability of participating in the survey and implicit stereotypical associations among students who made a stereotypical choice of major, and a negative correlation among students who made a counter-stereotypical choice of major. This seems unlikely. Moreover, the very same pattern can be observed when analysing the data from a gender-science implicit association test carried out by Harvard University Project Implicit[59], which covers a significantly wider population (Figure A.11 in Appendix F.).[60]

Regarding the cost of interacting with opposite-gender students, I rely on the evidence from the auxiliary survey provided in Figure XVII. Students who are enrolled in counter-stereotypical fields nominate significantly more people of the opposite gender when asked about their friends, people they study with, and people they ask questions on the material to. While this might be partially explained by a different availability of students of the opposite gender to interact with, the difference between students enrolled in stereotypical and counter-stereotypical departments is significantly bigger when it concerns students to ask questions on the material to. This points toward students enrolled in counter-stereotypical departments having a lower cost of engaging with opposite-gender peers, allowing them to benefit from academic support from their peers independent of their gender.

Differences in ability, and in particular, minorities being positively selected, could be an alternative potential explanation for why their performance is less affected by the composition of the environment. Differences in ability do not seem to be able to explain the results. Students who are enrolled in stereotypical and counter-stereotypical departments have equal ability at entrance. Even though math-intensive departments are the most competitive and selective, with the higher overall rejection rates, there is no evidence of a gender gap in qualifications at entry (Figure XIX).

Alternative mechanisms We might be worried that once students decided to select a stereotypical (or counter-stereotypical) field, they might be induced to act in a way that confirms their decision to reduce cognitive dissonance (Akerlof and Dickens, 1982; Festinger, 1957), or that being enrolled in a stereotypical (counter-stereotypical) major might affect students' implicit stereotypical associations and preferences for same-group interactions. Two pieces of evidence suggest that these two mechanisms are not able to reconcile the estimated results.

First, the elicited implicit stereotypical associations and preferences for same-gender peers are not significantly stronger for students that have been exposed more to the environment. Figure A.10 shows that the difference in score between students enrolled in stereotypical and counter-stereotypical majors is significant also when the sample is restricted to first-year students (even if the results are less precise due to the limited sample size), and the score for first-year students is not significantly different to the score for surveyed students enrolled in further years of their undergraduate programs.

[59] Information can be found at this website https://implicit.harvard.edu/implicit/takeatest.html.

[60] Whether we look at respondents who are current students (including 63.648 observations) or all the respondents that declare their university major (421.641 respondents), we can observe that men uphold more stereotypical beliefs compared to women when they attended a science-math related major, while less stereotypical beliefs compared to women when they attended a humanities-related major. This confirms that these patterns are not specific to LSE students.

Second, not every student making a stereotypical choice of major is affected by peers' gender to the same extent. In Table XVII, I exploit the Global Gender Gap Index (GGI) of the country of origin of the student as a proxy for the strengths of gender roles and norms for the student.[61] I estimate Specification 1 interacting the share of same-gender classmates with the GGI of each student's country of origin. I split students into three categories based on the terciles of the GGI of students' countries of origin. A higher tercile implies a higher GGI, indicating that the student comes from a country with more equal gender norms. I estimate this specification for the sample of students who are enrolled in counter-stereotypical fields (Column 1) and for the sample of students who are enrolled in stereotypical departments (Column 2)[62]. Focusing on students who are enrolled in stereotypical departments, we can see that those who benefit the most from being surrounded by same-gender classmates are students who come from countries with low GGI, i.e. students who come from countries with more unequal gender roles. On the other hand, students who come from countries where gender roles are more equal do not benefit as much. The opposite is true for students who made counter-stereotypical choices, indicating that the students coming from countries with more unequal gender roles (low GGI) are those who benefit the most from being in a minority if they went against stereotypes when choosing their major. The effects however are small and not significant. This indicates that confirming the decision taken might play a role, but cannot be the only reason for the estimated patterns to appear. Ex-ante "sensitivity" to stereotypes and social norms inducing students to select different majors and then react to the composition of the environment in a self-fulfilling way seems to be a more plausible explanation.

VII. Discussion and Implications for Policy

This paper provides evidence that, as individuals internalize social identity considerations and the composition of the environment when making their choices, peers' identity affects individuals belonging to the same group (for example, gender or ethnicity) in a different way depending on the environment where we observe them. Ex-ante "sensitivity" to social norms and preferences to engage with same-gender peers induce students to select different majors and then react to the composition of the environment in a self-fulfilling way.

Considering selection into minority status is particularly important when the environment is very selective and the presence of a person from an under-represented group is the result of a series of strategic choices. This is very likely to be the case at the LSE. Application to undergraduate programs is centralized in the UK, and students can only apply to 5 programs (across all universities in the UK) in the same year. Furthermore, applications at LSE are very competitive: 75% of students who apply are rejected every year. Thus students who apply at LSE are very motivated and very confident.

The LSE setting is well-suited to answer the paper's research question as it provides an ideal environment

[61] Guiso et al. (2008) provides evidence that the Country's GGI index is significantly correlated with gender gaps in math and reading. In particular, countries with lower GGI have bigger math gender gaps in favor of men, while countries with high GGI have bigger reading gender gaps in favor of women.

[62] I am using here the top and bottom 5 departments definition of stereotypical selection. This is the most conservative definition of stereotypical and counter-stereotypical, which also allows me to have a big enough sample size to estimate significant effects.

to shed light on the effect of selection based on stereotypes on minority status. Furthermore, it allows to disregard discrimination, thanks to the blind grading system. However, the paper's findings cannot be easily generalized to other educational environments since LSE and its students are not comparable to the average university. Given the competitive nature of the environment and the strategic decision-making process that applying to undergraduate programs in the UK entails, these findings might be informative for selective and competitive working environments. For instance, the paper's findings might inform policies addressing the under-representation of women in decision-making bodies or leading positions. In this regard, the results on participation in class become particularly informative, as this outcome can be indicative of what happens in work environments where social interactions between colleagues occur.

These results have important policy implications. In recent years, numerous initiatives, such as reallocation policies, mentoring programs, networking events, role model provision, and many others, have brought minorities together to improve diversity, reduce the threat of stereotypes, and equalize *performance* at every level of the career ladder. The evidence gathered by this paper suggests that these initiatives might not be enough to level the playing field and might even backfire, especially when they target minorities in selective environments once education and occupational choices have been made. When belonging to the minority group is the result of a series of counter-stereotypical choices and strong selection, those who uphold the strongest stereotypical associations and preferences for same-gender interactions might be those who belong to the majority group (those who select on stereotypes). In these environments, initiatives targeting minorities might not be the most effective in reducing imbalances and might even risk contributing to perpetuating stereotypes. The counterfactual simulation presented in Section V.A. shows that policies such as sex segregation or reallocation to enforce a more equal gender ratio could deliver unintended effects in the short-run - reducing differences in performance across groups, decreasing average performance at the same time. Moreover, by making differences between groups salient and reducing the possibility for the majority group to update their beliefs through experiential learning, initiatives nudging minorities, such as networking events or mentoring programs, might reinforce stereotypical distortions in the mind of the majority group.

The discussion so far abstracted from considering the implications that these results have for the effectiveness of measures such as affirmative action policies in equalizing *participation*. Affirmative action policies affect the composition of the pool of individuals from under-represented groups in counter-stereotypical fields by introducing students that might have otherwise selected more stereotypical majors in environments where their group is severely under-represented (e.g. women in economics). This could potentially negatively affect the performance of the minority group in the field to which these students belong in the short term. Furthermore, it could affect the short-term performance of the majority group in the field via its effect on the gender composition of the environment. However, at the same time, by affecting performance and the composition of the environment in the short term, it could consequently impact selection in subsequent periods. As students derive information regarding their returns to ability in a particular discipline from the existing gender composition of majors (and potentially from the average performance of different groups in each field via stereotypical distortions), the introduction of quotas would affect students' selection in the following years, potentially generating a trade-off between short and long-term effects. This represents another important avenue for research.

REFERENCES

Akerlof, G. and W. T. Dickens (1982). The economic consequences of cognitive dissonance. *American Economic Review 72*(3), 307–19.

Akerlof, G. A. and R. E. Kranton (2000, 08). Economics and Identity*. *The Quarterly Journal of Economics 115*(3), 715–753.

Anelli, M. and G. Peri (2017). The effects of high school peers' gender on college major, college performance and income. *The Economic Journal*.

Angrist, J. D. (2014). The perils of peer effects. *Labour Economics 30*, 98–108. Special Section articles on "What determined the dynamics of labour economics research in the past 25 years? edited by Joop Hartog and and European Association of Labour Economists 25th Annual Conference, Turin, Italy, 19-21 September 2013 Edited by Michele Pellizzari.

Ashraf, N., O. Bandiera, and S. S. Lee (2014). Awards unbundled: Evidence from a natural field experiment. *Journal of Economic Behavior Organization 100*(C), 44–63.

Austen-Smith, D. and J. Fryer, Roland G. (2005, 05). An Economic Analysis of "Acting White"*. *The Quarterly Journal of Economics 120*(2), 551–583.

Bandiera, O., V. Larcinese, and I. Rasul (2010). Heterogeneous class size effects: New evidence from a panel of university students*. *The Economic Journal 120*(549), 1365–1398.

Beede, D. N., T. Julian, D. Langdon, G. McKittrick, B. Khan, and M. Doms (2011). Women in stem: A gender gap to innovation. *Sexuality & Gender Studies eJournal*.

Bertrand, M. (2011). Chapter 17 - new perspectives on gender. Volume 4 of *Handbook of Labor Economics*, pp. 1543–1590. Elsevier.

Bertrand, M. (2018). Coase lecture – the glass ceiling. *Economica 85*(338), 205–231.

Bertrand, M. (2020, May). Gender in the twenty-first century. *AEA Papers and Proceedings 110*, 1–24.

Blau, F. D. and L. M. Kahn (2017). The gender wage gap: Extent, trends, and explanations. *Journal of Economic Literature 55*(3), 789–865.

Bonomi, G., N. Gennaioli, and G. Tabellini (2021, 09). Identity, Beliefs, and Political Conflict*. *The Quarterly Journal of Economics 136*(4), 2371–2411.

Booth, A. L., L. Cardona-Sosa, and P. Nolen (2018). Do single-sex classes affect academic achievement? an experiment in a coeducational university. *Journal of Public Economics 168*, 109–126.

Bordalo, P., K. Coffman, N. Gennaioli, and A. Shleifer (2016). Stereotypes. *Quarterly Journal of Economics 131*(4), 1753–1794.

Bordalo, P., K. Coffman, N. Gennaioli, and A. Shleifer (2019). Beliefs about gender. *American Economic Review 109*(3), 739–73.

Born, A., E. Ranehill, and A. Sandberg (2020). Gender and willingness to lead: Does the gender composition of teams matter? *The Review of Economics and Statistics 0*(ja), 1–46.

Bostwick, V. and B. Weinberg (2021). Nevertheless she persisted? gender peer effects in doctoral stem programs. *Journal of Labor Economics 0*(ja), null.

Braga, M., M. Paccagnella, and M. Pellizzari (2016). The Impact of College Teaching on Students' Academic and Labor Market Outcomes. *Journal of Labor Economics 34*(3), 781–822.

Bramoullé, Y., H. Djebbari, and B. Fortin (2020). Peer effects in networks: A survey. *Annual Review of Economics 12*(1), 603–629.

Breda, T., J. Grenet, M. Monnet, and C. van Effenterre (2021, March). Do Female Role Models Reduce the Gender Gap in Science? Evidence from French High Schools. PSE Working Papers halshs-01713068, HAL.

Brenøe, A. A. and U. Zölitz (2020). Exposure to more female peers widens the gender gap in stem participation. *Journal of Labor Economics 38*(4), 1009–1054.

Bursztyn, L., G. Egorov, and R. Jensen (2019, 05). Cool to be Smart or Smart to be Cool? Understanding Peer Pressure in Education. *The Review of Economic Studies 86*(4), 1487–1526.

Bursztyn, L., T. Fujiwara, and A. Pallais (2017, November). 'acting wife': Marriage market incentives and labor market investments. *American Economic Review 107*(11), 3288–3319.

Campbell, S., L. Macmillan, R. Murphy, and G. Wyness (2019, August). Inequalities in student to course match: evidence from linked administrative data. CEP Discussion Papers dp1647, Centre for Economic Performance, LSE.

Card, D., A. Mas, and J. Rothstein (2008, 02). Tipping and the Dynamics of Segregation*. *The Quarterly Journal of Economics 123*(1), 177–218.

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

Carrell, S. E., M. E. Page, and J. E. West (2010, 08). Sex and Science: How Professor Gender Perpetuates the Gender Gap*. *The Quarterly Journal of Economics 125*(3), 1101–1144.

Chen, J. and D. Houser (2019). When are women willing to lead? the effect of team gender composition and gendered tasks. *The Leadership Quarterly 30*(6), 101340.

Coffman, K., C. B. Flikkema, and O. Shurchkov (2021). Gender stereotypes in deliberation and team decisions. *Games and Economic Behavior 129*, 329–349.

Coffman, K. B. (2014). Evidence on Self-Stereotyping and the Contribution of Ideas. *The Quarterly Journal of Economics 129*(4), 1625–1660.

Corno, L., E. La Ferrara, and J. Burns (2022, December). Interaction, stereotypes, and performance: Evidence from south africa. *American Economic Review 112*(12), 3848–75.

De Giorgi, G., M. Pellizzari, and S. Redaelli (2010, April). Identification of social interactions through partially overlapping peer groups. *American Economic Journal: Applied Economics 2*(2), 241–75.

Del Carpio, L. and M. Guadalupe (2022). More women in tech? evidence from a field experiment addressing social identity. *Management Science 68*(5), 3196–3218.

Delfino, A. (2021). Breaking Gender Barriers: Experimental Evidence on Men in Pink-Collar Jobs. IZA Discussion Paper 14083.

Ellemers, N. (2018). Gender stereotypes. *Annual Review of Psychology 69*(1), 275–298. PMID: 28961059.

Feld, J. and U. Zölitz (2017). Understanding peer effects: On the nature, estimation, and channels of peer effects. *Journal of Labor Economics 35*(2), 387–428.

Feld, J. and U. Zölitz (2022). The effect of higher-achieving peers on major choices and labor market outcomes. *Journal of Economic Behavior Organization 196*, 200–219.

Festinger, L. (1957, June). *A Theory of Cognitive Dissonance*. Stanford University Press.

Fischer, S. (2017). The downside of good peers: How classroom composition differentially affects men's and women's stem persistence. *Labour Economics 46*, 211–226.

Garg, N., L. Schiebinger, D. Jurafsky, and J. Zou (2018). Word embeddings quantify 100 years of gender and ethnic stereotypes. *Proceedings of the National Academy of Sciences 115*(16), E3635–E3644.

Giorgi, G. D., W. G. Woolston, and M. Pellizzari (2012). Class size and class heterogeneity. *Journal of the European Economic Association 10*(4), 795–830.

Goldin, C. (2014, April). A grand gender convergence: Its last chapter. *American Economic Review 104*(4), 1091–1119.

Greenwald, A. G., D. E. McGhee, and J. L. K. Schwartz (1998). Measuring individual differences in implicit cognition: the implicit association test. *Journal of personality and social psychology 74 6*, 1464–80.

Griffith, A. L. (2010). Persistence of women and minorities in stem field majors: Is it the school that matters? *Economics of Education Review 29*(6), 911–922.

Griffith, A. L. and J. B. Main (2019). First impressions in the classroom: How do class characteristics affect student grades and majors? *Economics of Education Review 69*, 125–137.

Guiso, L., F. Monte, P. Sapienza, and L. Zingales (2008). Culture, gender, and math. *Science 320*(5880), 1164–1165.

Guryan, J., K. Kroft, and M. J. Notowidigdo (2009, October). Peer effects in the workplace: Evidence from random groupings in professional golf tournaments. *American Economic Journal: Applied Economics 1*(4), 34–68.

He, J. C., S. K. Kang, K. Tse, and S. M. Toh (2019). Stereotypes at work: Occupational stereotypes predict race and gender segregation in the workforce. *Journal of Vocational Behavior 115*, 103318.

Hebert, C. (2020). Gender stereotypes and entrepreneur financing. 10th miami behavioral finance conference, available at ssrn: https://ssrn.com/abstract=3318245 or http://dx.doi.org/10.2139/ssrn.3318245.

Hill, A. J. (2017). The positive influence of female college students on their male peers. *Labour Economics 44*, 151–160.

Hoff, K. and P. Pandey (2006, May). Discrimination, social identity, and durable inequalities. *American Economic Review 96*(2), 206–211.

Huntington-Klein, N. and E. Rose (2018). Gender peer effects in a predominantly male environment: Evidence from west point. *AEA Papers and Proceedings 108*, 392–395.

Inzlicht, M., J. Aronson, C. Good, and L. McKay (2006, May). A particular resiliency to threatening environments. *Journal of Experimental Social Psychology 42*(3), 323–336.

Inzlicht, M. and C. Good (2006). How environments can threaten academic performance, self-knowledge, and sense of belonging. *In S. Levin and C. van Laar (Eds.), The Claremont symposium on Applied Social Psychology. Stigma and group inequality: Social psychological perspectives*, 129–150.

Jain, T. and M. Kapoor (2015). The impact of study groups and roommates on academic performance. *The Review of Economics and Statistics 97*(1), 44–54.

Kandel, E. and E. P. Lazear (1992). Peer pressure and partnerships. *Journal of Political Economy 100*(4), 801–817.

Karpowitz, Chris, J. P. and O. Stoddard (2020). Strength in numbers: A field experiment on gender, influence and group dynamics. IZA Discussion Paper No. 13741.

Kugler, A. D., C. H. Tinsley, and O. Ukhaneva (2021). Choice of majors: are women really different from men? *Economics of Education Review 81*, 102079.

Lane, K. A., G. J. X. . D.-L. E. (2012). Implicit science stereotypes mediate the relationship between gender and academic participation. *Sex Roles: A Journal of Research 66(3-4)*, 220–234.

Lavy, V. and A. Schlosser (2011, April). Mechanisms and impacts of gender peer effects at school. *American Economic Journal: Applied Economics 3*(2), 1–33.

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

Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *The Review of Economic Studies 60*(3), 531–542.

McNally, S. (2020). Gender differences in tertiary education: What explains stem participation. CEP Discussion Paper No. 1721.

Murdoch, D. J., Y.-L. Tsai, and J. Adcock (2008). P-values are random variables. *The American Statistician 62*(3), 242–245.

OECD (2020). *Education at a Glance 2020*.

Olivetti, C., E. Patacchini, and Y. Zenou (2020). Mothers, Peers, and Gender-Role Identity. *Journal of the European Economic Association 18*(1), 266–301.

Olsson, M. and S. E. Martiny (2018). Does exposure to counterstereotypical role models influence girls' and women's gender stereotypes and career choices? a review of social psychological research. *Frontiers in Psychology 9*, 2264.

Oosterbeek, H. and R. van Ewijk (2014). Gender peer effects in university: Evidence from a randomized experiment. *Economics of Education Review 38*, 51–63.

Oxoby, R. J. (2014). Social inference and occupational choice: Type-based beliefs in a bayesian model of class formation. *Journal of Behavioral and Experimental Economics 51*, 30–37.

Pan, J. (2015). Gender segregation in occupations: The role of tipping and social interactions. *Journal of Labor Economics 33*(2), 365–408.

Porter, C. and D. Serra (2020, July). Gender differences in the choice of major: The importance of female role models. *American Economic Journal: Applied Economics 12*(3), 226–54.

Pregaldini, D., U. Backes-Gellner, and G. Eisenkopf (2020). Girls' preferences for stem and the effects of classroom gender composition: New evidence from a natural experiment. *Journal of Economic Behavior Organization 178*, 102–123.

Reuben, E., P. Sapienza, and L. Zingales (2014). How stereotypes impair women's careers in science. *Proceedings of the National Academy of Sciences 111*(12), 4403–4408.

Russell, L. (2017). Can learning communities boost success of women and minorities in stem? evidence from the massachusetts institute of technology. *Economics of Education Review 61*, 98–111.

Sacerdote, B. (2001, 05). Peer Effects with Random Assignment: Results for Dartmouth Roommates*. *The Quarterly Journal of Economics 116*(2), 681–704.

Sacerdote, B. (2011). Chapter 4 - peer effects in education: How might they work, how big are they and how much do we know thus far? Volume 3 of *Handbook of the Economics of Education*, pp. 249–277. Elsevier.

Sacerdote, B. (2014). Experimental and quasi-experimental analysis of peer effects: Two steps forward? *Annual Review of Economics 6*(1), 253–272.

Shan, X. (2020). Does minority status drive women out of male-dominated fields? Working paper.

Spencer, S. J., C. Logel, and P. G. Davies (2016). Stereotype threat. *Annual Review of Psychology 67*(1), 415–437. PMID: 26361054.

Steele, C. and J. Aronson (1995, November). Stereotype threat and the intellectual test performance of african americans. *Journal of Personality and Social Psychology 69*(5), 797–811. Copyright: Copyright 2018 Elsevier B.V., All rights reserved.

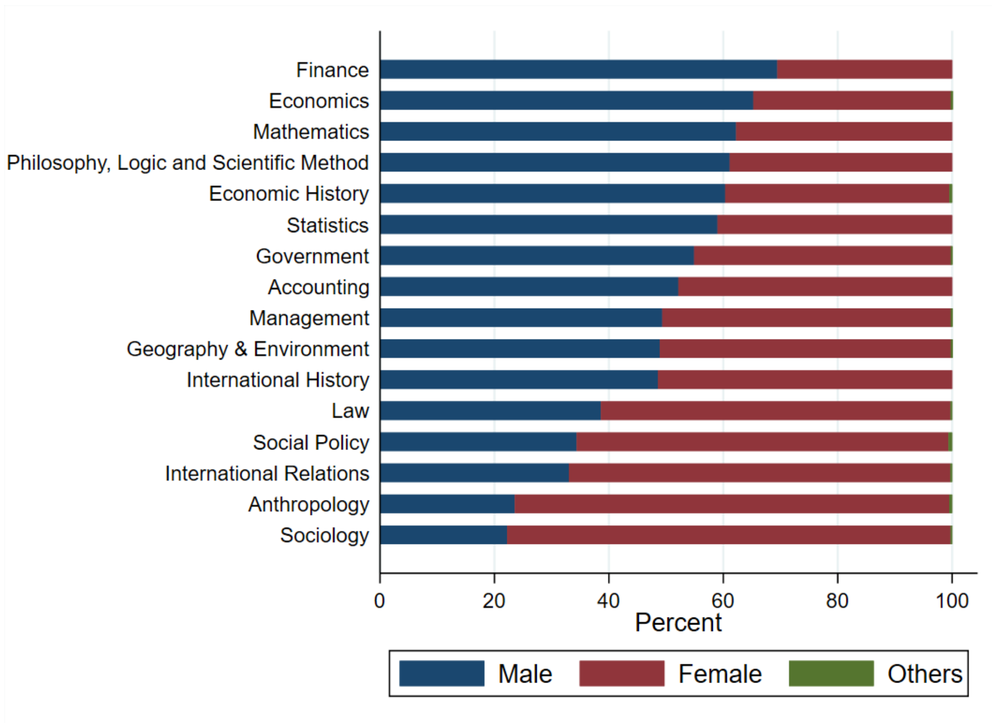
Stinebrickner, R. and T. R. Stinebrickner (2006). What can be learned about peer effects using college roommates? evidence from new survey data and students from disadvantaged backgrounds. *Journal of Public Economics 90*(8), 1435–1454.

Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. *Review of Educational Research 45*(1), 89–125.

Zölitz, U. and J. Feld (2021). The effect of peer gender on major choice in business school. *Management Science 67*(11), 6963–6979.

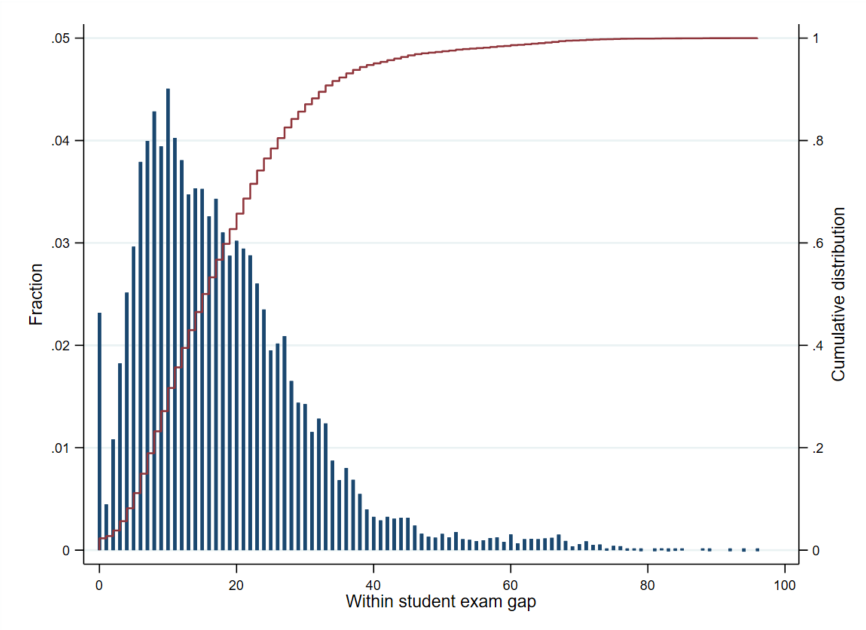
VIII. FIGURES

Figure I
Departments Gender Composition



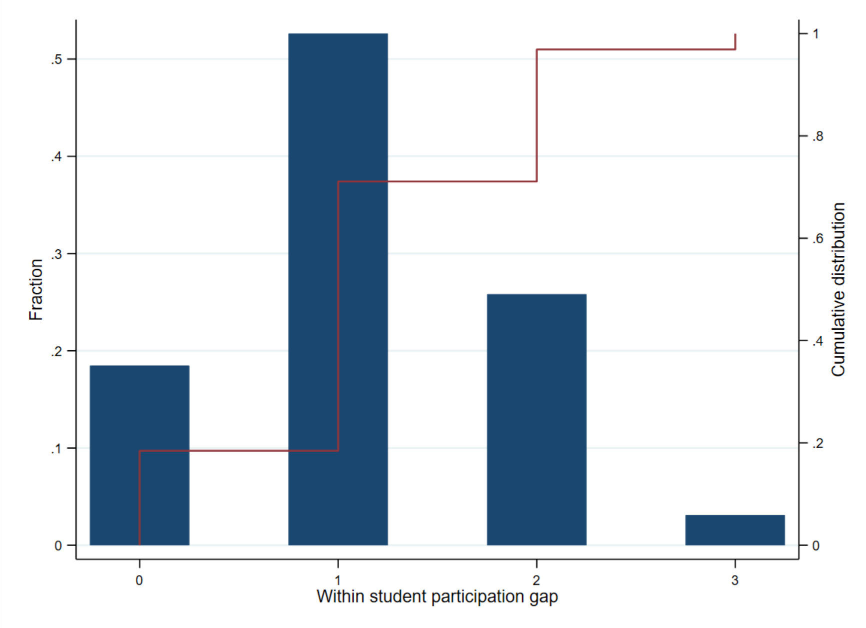
Notes: This figure illustrates the gender composition of each department. It is constructed based on the number of first year undergraduate students that are enrolled in bachelor programs offered by the department between 2008/09 and 2017/18. The figure shows the average across academic years 2008/09 to 2017/18.

Figure II
Within-student Course Performance Gap



Notes: Following Bandiera et al. (2010). The figure displays the within student exam gap. This is defined as the difference between student i's highest and lowest exam mark across all the courses attended during Michaelmas term of the first year. The sample considered in the graph is the sample in analysis: 54.603 course-year-class group level observations, corresponding to information on 14.313 students.

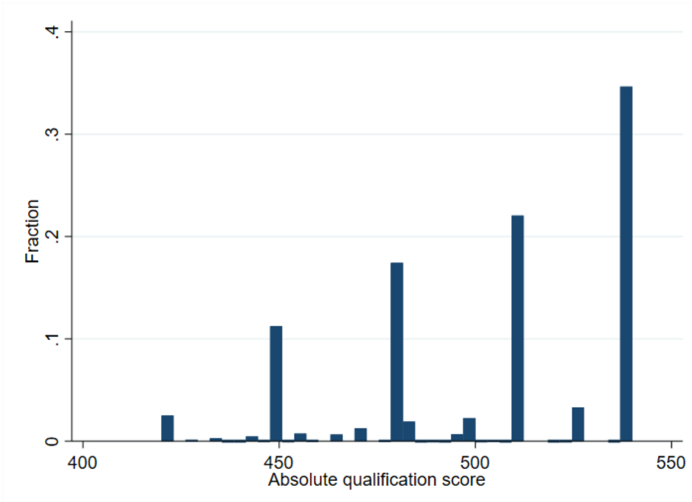
Figure III
Within-student Class Participation Gap



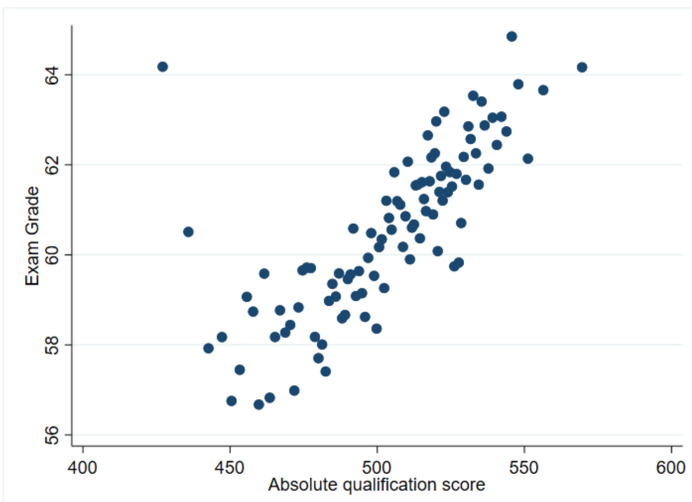
Notes: Following Bandiera et al. (2010). The figure displays the within student participation gap. This is defined as the difference between student i's highest and lowest participation grade across all the classes attended during Michaelmas term of the first year. The sample consist of the sample of analysis restricted to all the students for which I can observe a participation grade in at least two class groups: 44.771 course-year-class group level observations, corresponding to information on 13.485 students.

Figure IV
Qualification Scores at Entry

Panel A



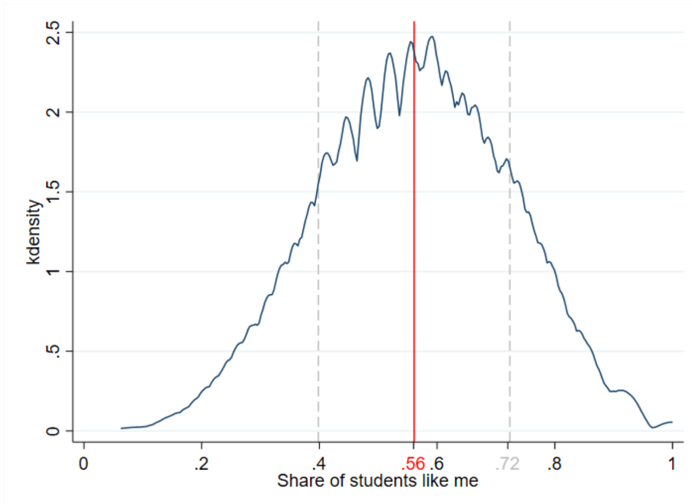
Panel B



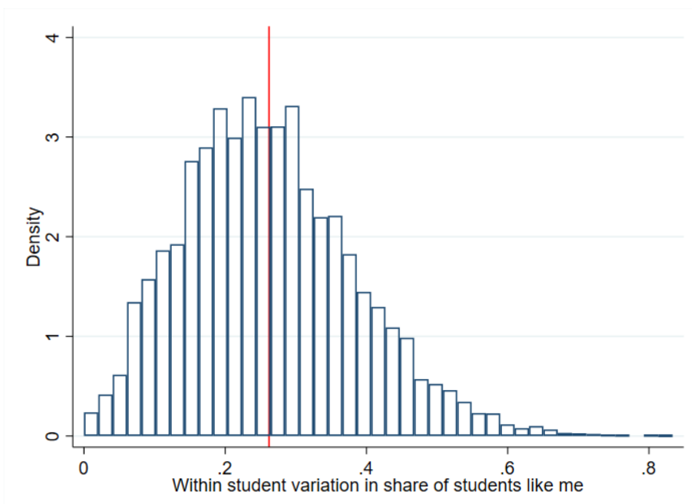
Notes: Panel A displays qualification scores at entry for 9.449 students, 90% of the students enrolled in undergraduate programs in LSE between academic years 2011/12 and 2017/18 (students for which I have information regarding qualifications at entry). Following Campbell et al. (2019), I construct a qualification score based on the best three exam results among the A-level qualification scores students declared when applying to LSE. Some students take courses/programs that are equivalents to A-Levels. In these cases I calculate their A-Level equivalence scores based on the university conversion tables for foreign students. A levels are graded on a scale of A*/A/B/C/D/E. Each A-level grade is worth 30 QCA (Qualifications and Curriculum Authority) points. Panel B displays the correlation between qualification scores at entry and course grades controlling for qualification type, individual characteristics (program of enrolment, gender, ethnicity, social background), and course fixed effects

Figure V
Within-student Variation in Share of Same-gender Classmates

Panel A

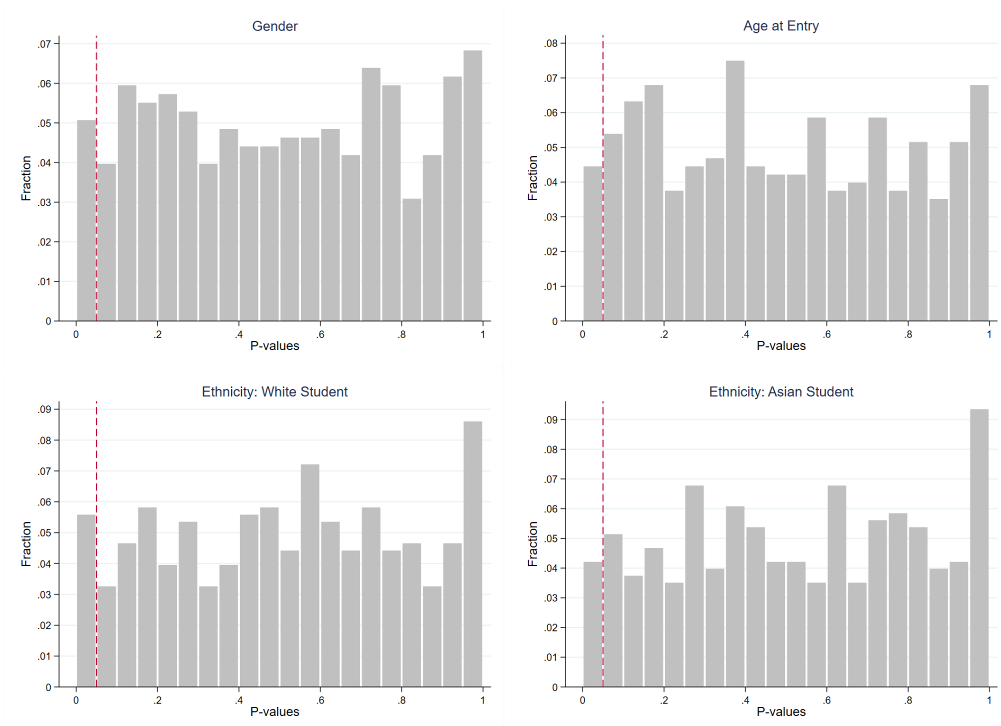


Panel B



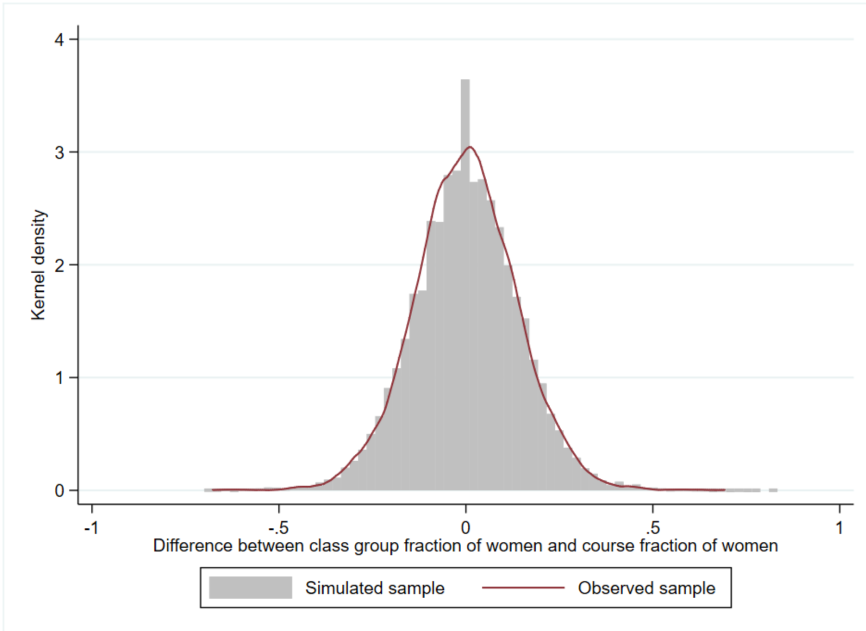
Notes: Panel A displays the distribution of the share of students like me, defined as the share of same gender classmates. The red vertical line indicates the average, while the two grey dashed lines indicate the average plus and minus one standard deviation. Panel B displays the distribution of the within student variation in the share of students like me. This is defined as the difference between the maximum and minimum share of same gender classmates experienced by each student during the first academic year. The red line indicates the average variation.

Figure VI
Identification: Observed Sample P-values



Notes: This graph show the p-values of a test of joint significance of class dummies from a regression of gender (age at entry, a dummy for being White, and a dummy for being Asian, one at the time) on class dummies and dummies for each one of the courses taken by students in the academic years in analysis. A total of 454 regressions were performed. The sample is restricted to the courses that have at least 2 classes. The test is performed on all the students that attend courses attended by the students in the sample of analysis and that did not change classes (which are the students for whom I can recover the initial class allocation). More information on the text can be found in Appendix C..

Figure VII
Identification: Allocation Simulation - Difference in Share of Female in Class and Course

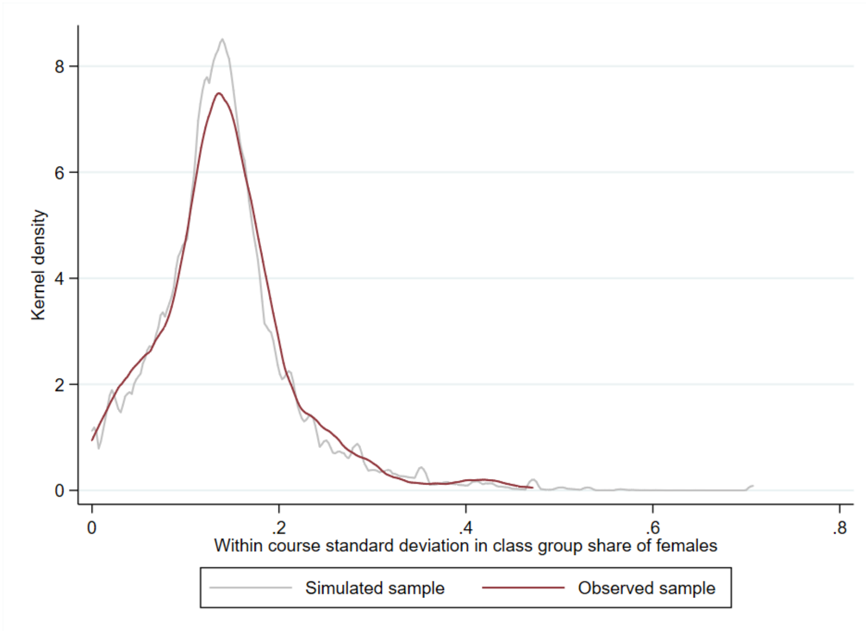


Two-sample Kolmogorov-Smirnov test

Smaller group	D	P-value
0:	0.0130	0.193
1:	-0.0116	0.271
Combined K-S:	0.0130	0.384

Notes: The Figure displays in dark grey the distribution of the difference between the share of women in each class and the share of women in the course for 1000 simulations of an unconstrained (not considering scheduling constraints) random allocation of students to classes. This is compared with the same statistics for the observed allocation (in red). The result of a two-sample Kolmogorov-Smirnov test shows that the distribution of the observed statistics is not significantly different from the simulated statistics.

Figure VIII
Identification: Allocation Simulation - Within-course Std. Dev. in Class Groups Share of Females

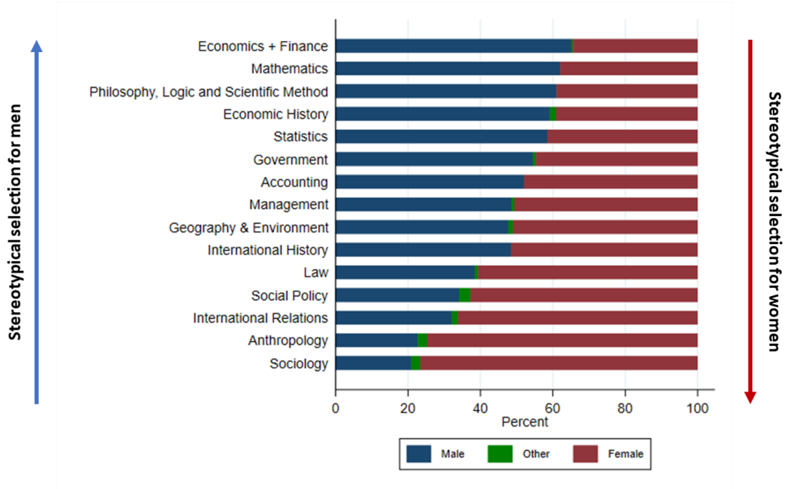


Two-sample Kolmogorov-Smirnov test

Smaller group	D	P-value
0:	0.0286	0.410
1:	-0.0272	0.446
Combined K-S:	0.0286	0.764

Notes: The Figure displays in dark grey the distribution of the within course standard deviation in the class share of females for 1000 simulations of an unconstrained (not considering scheduling constraints) random allocation of students to classes. This is compared with the same statistics for the observed allocation (in red). The result of a two-sample Kolmogorov-Smirnov test shows that the distribution of the observed statistics is not significantly different from the simulated statistics.

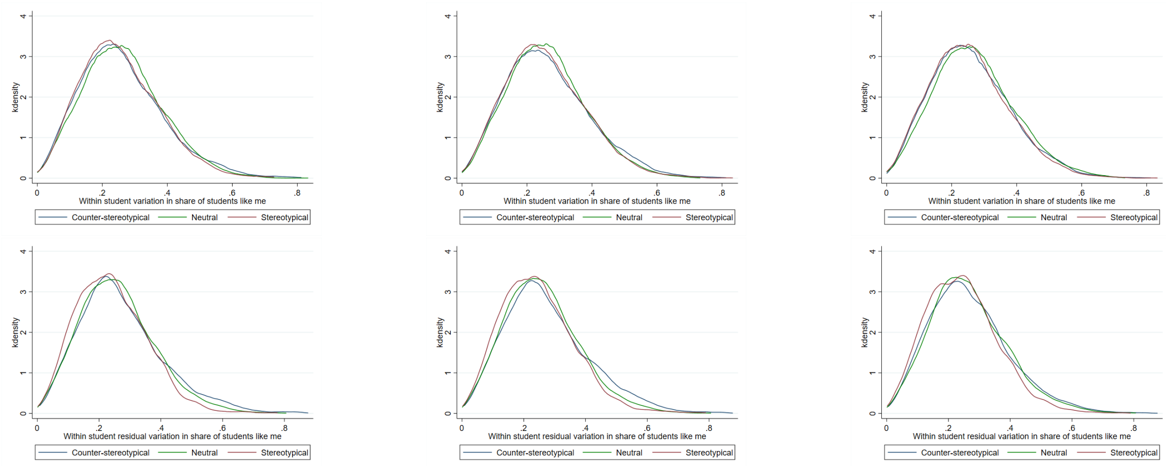
Figure IX
Stereotypical Selection - Definition



Notes: The Figure displays the distribution of men and women enrolled in the first year of undergraduate programs across departments. The statistics is the average of academic years 2008-2017. The blue and red arrows at the sides display the direction of choices of majors that would be in line with stereotypes and gender norms for men and women respectively.

Figure X
Within-student Variation in Share of Same-gender Classmates Across Departments

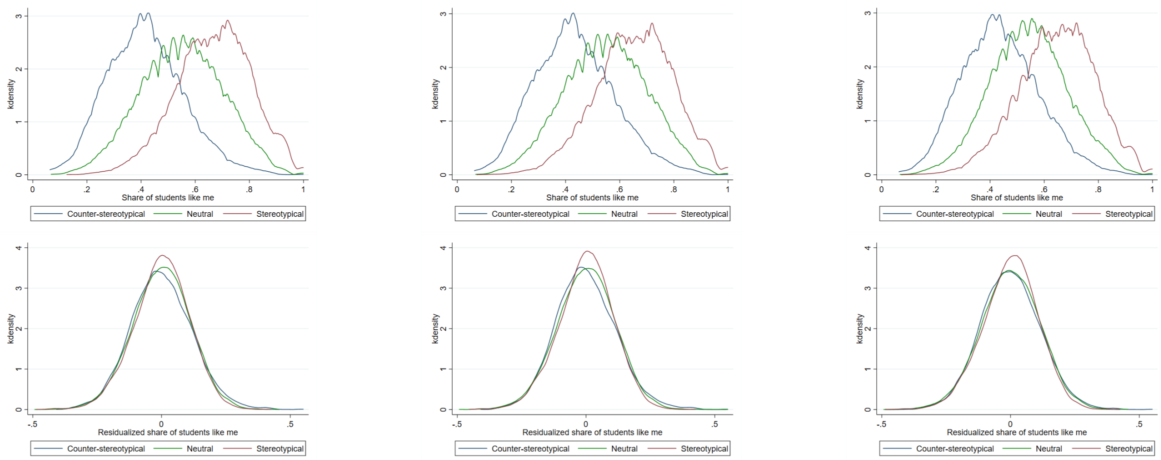
Panel A: Top & Bottom 2 | **Panel B: Top & Bottom 3** | **Panel C: Top & Bottom 5**



Notes: The figure displays the within student variation in the share of same gender classmates (top) and the within student variation in the residual share of same gender classmates (bottom) for students enrolled in stereotypical and counter-stereotypical departments, according to the three definitions used. The within student variation in the share of same gender classmates is obtained by taking the difference between student i's highest and lowest share of same gender classmates across all the classes attended during the first year. The within student variation in the residual share of same gender classmates is obtained by taking the difference between student i's highest and lowest residual share of same gender classmates across all the classes attended during the first year, obtained after regressing the share of same gender classmates on course × year fixed effects. Kolmogorov-Smirnov Test of equality in distribution for Stereotypical vs Against-stereotypes P-value: Variation - 0.526, 0.208, 0.139; Residual variation - 0.000, 0.000, 0.000.

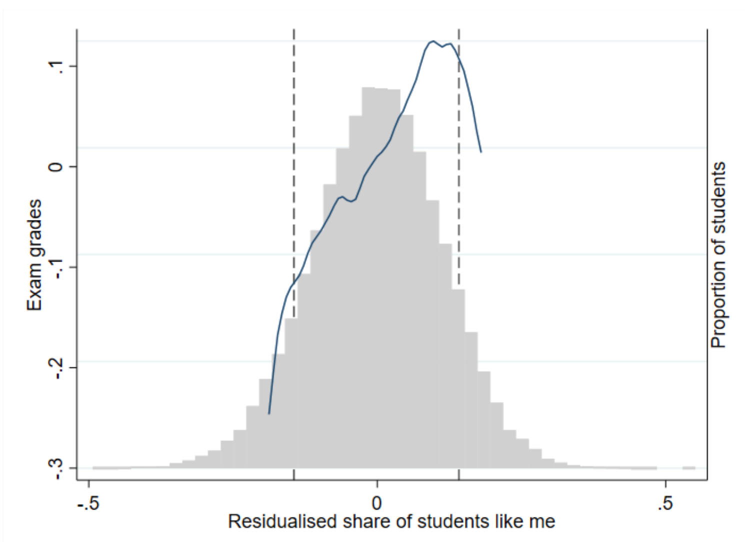
Figure XI
Share of Same-gender Classmates Across Departments

Panel A: Top & Bottom 2 **Panel B: Top & Bottom 3** **Panel C: Top & Bottom 5**



Notes: The figure displays the share of same gender classmates and the residual share of same gender classmates for students enrolled in stereotypical and counter-stereotypical departments, according to the three definitions used. The residual share of same gender classmates is obtained by taking the residuals of the regression of same gender classmates on course × year fixed effects and student fixed effects. Kolmogorov-Smirnov Test of equality in distribution for Stereotypical vs Against-stereotypes P-value: Share - 0.000, 0.000, 0.000; Residual share - 0.000, 0.000, 0.000.

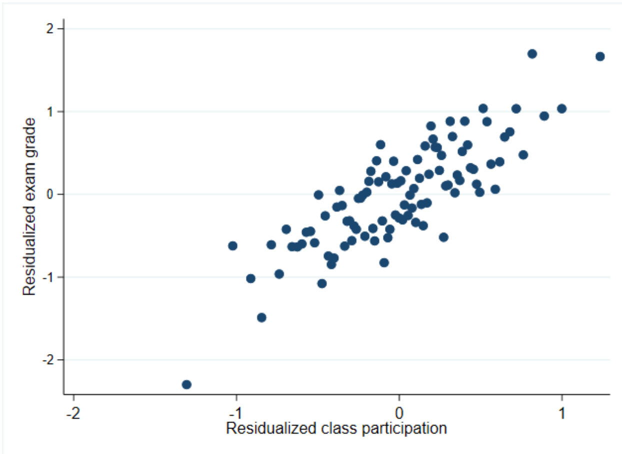
Figure XII
Alternative Mechanisms: Minority Effect - Linearity



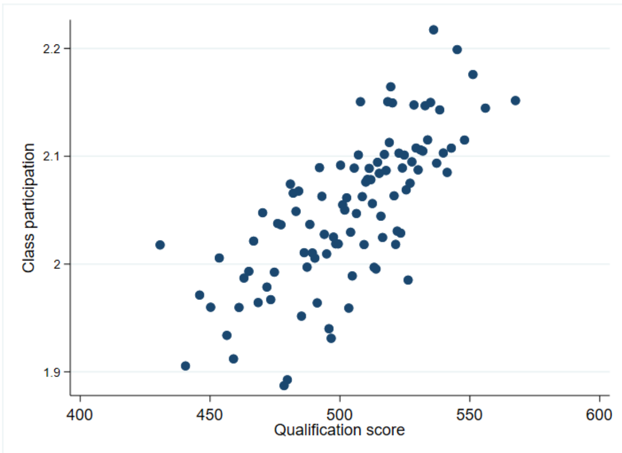
Notes: Local polynomial plot of the relationship between residualised course grades on the residualised share of same gender classmates. These are obtained by regressing course grades and the share of samge gender classmates on course and student fixed effects, respectively. The vertical lines display the 5st and 95th percentile of the share of same gender classmates. The grey histogram displays the support of the residualised share of same gender classmates.

Figure XIII
Mechanisms: Class Participation

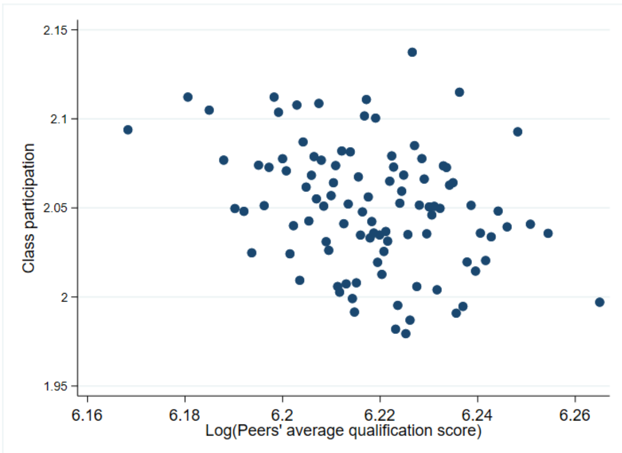
Panel A



Panel B



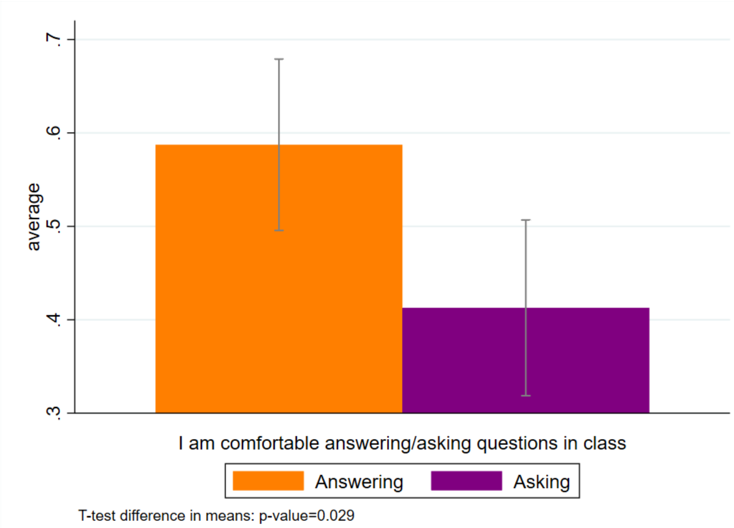
Panel C



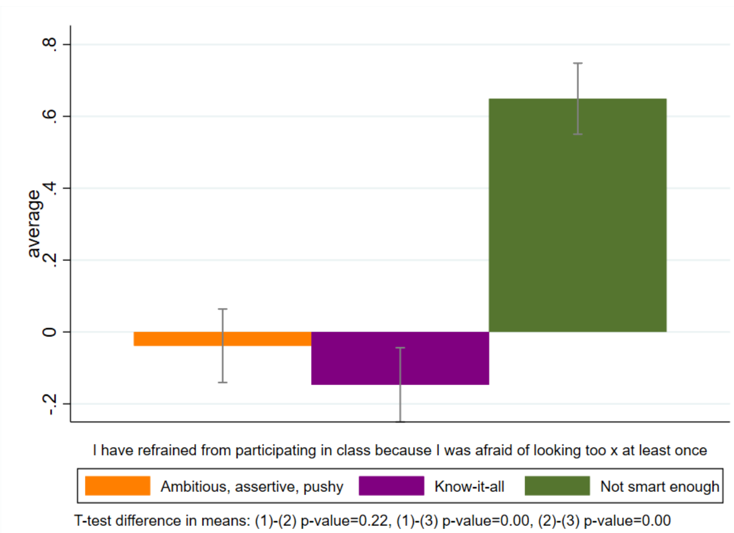
Notes: Panel A displays a binned scatterplot of residualized exam grades on residualized participation in class. Each variable is the residual of a regression on class group and students fixed effects. Panel B displays a binned scatterplot of participation grades on students' ability controlling for class group and individual characteristics (gender, program of enrollment, school of origin, ethnicity). Panel C displays a binned scatterplot of participation grades on class group peers' ability controlling for class size, course and students fixed effects.

Figure XIV
Mechanisms: Image Concerns and Class Participation

Panel A

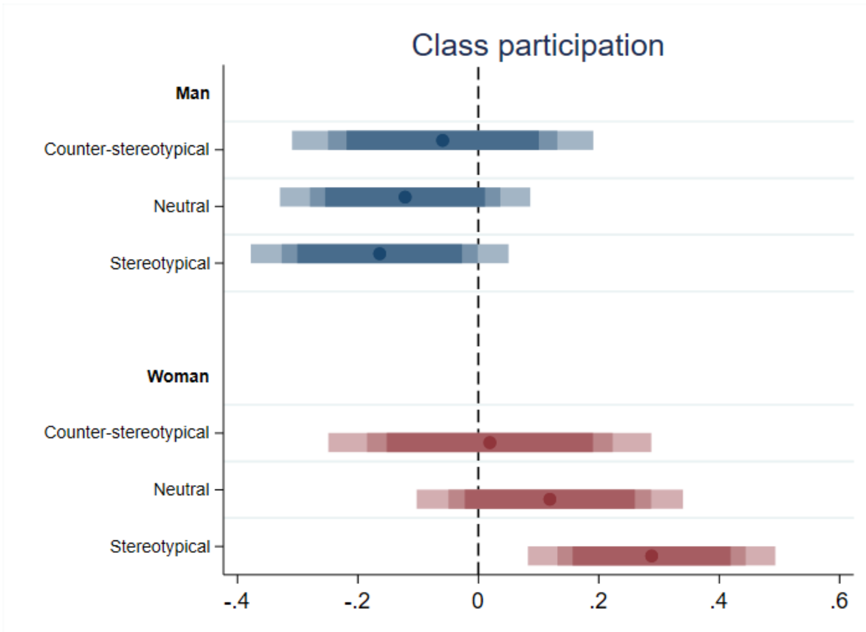


Panel B



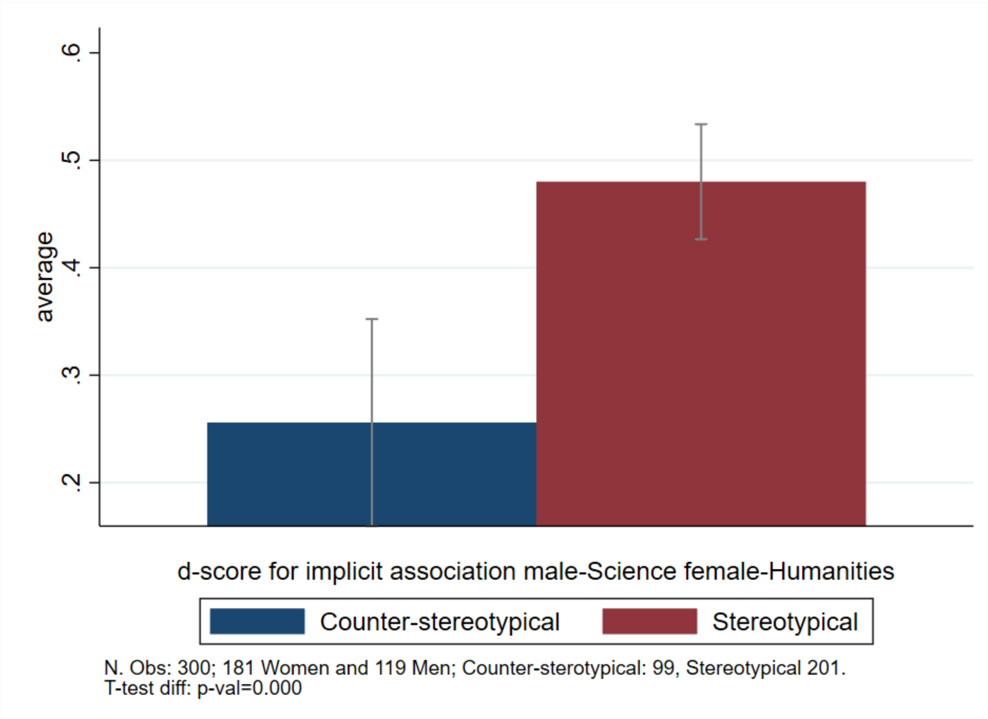
Notes: Following Bursztyn et al. (2017). Panel A displays the answers to the following questions: "Thinking about the courses you have taken in your career at LSE, how much do you agree or disagree with the following statements? (-2: strongly disagree, +2: strongly agree) (i) I am comfortable answering questions or contributing to the discussion in class; (ii) I am comfortable answering questions or contributing to the discussion in class". Panel B displays the answers to the following questions: "Thinking about the courses you have taken in your career at LSE, how much do you agree or disagree with the following statements? (-2: strongly disagree, +2: strongly agree) (i) I have refrained from participating in class because I was afraid of looking too ambitious, assertive, or pushy at least once (ii) I have refrained from participating in class because I was afraid of looking too “know-it-all” at least once (iii) I have refrained from participating in class because I was afraid of looking “not smart enough” at least once".

Figure XV
Mechanisms: Stereotypical Selection - Class Participation



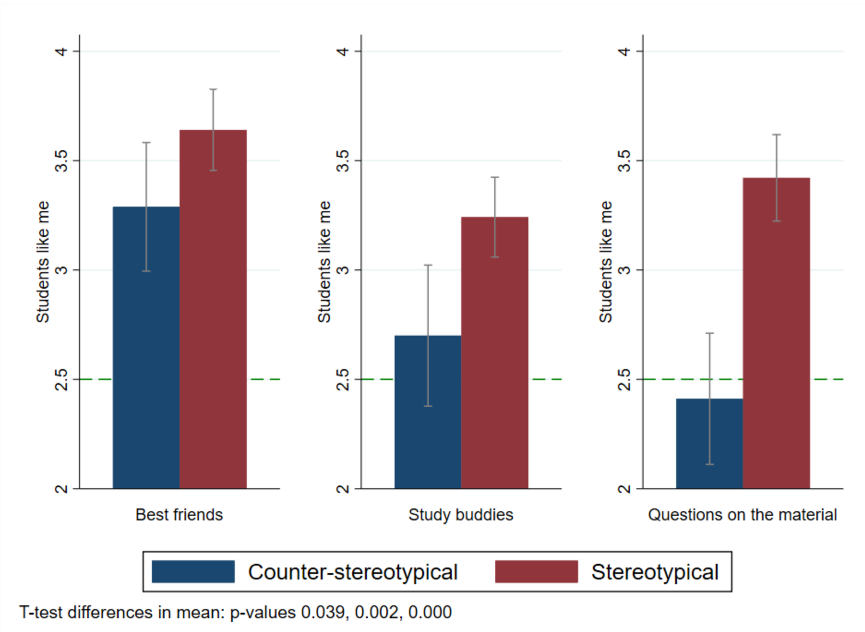
Notes: The Figure displays the results of specification 3. The outcomes variable are participation grades. Stereotypical and counter-stereotypical departments are defined based on the top and bottom 5 departments in terms of share of students like me enrolled in the department. Standard errors are clustered at the class group level.

Figure XVI
Mechanisms: Scientific-Male, Humanistic-Female IAT



Notes: Results of an implicit association test (Greenwald et al. 1998) to elicit students' associations between female-humanistic and male-scientific. A score of 0 indicates no association between male -scientific and female-humanistic; a positive score indicates that the student associates women with humanities and men with science and math; lastly a negative score indicates that the student associates men with humanities and women with science and math.

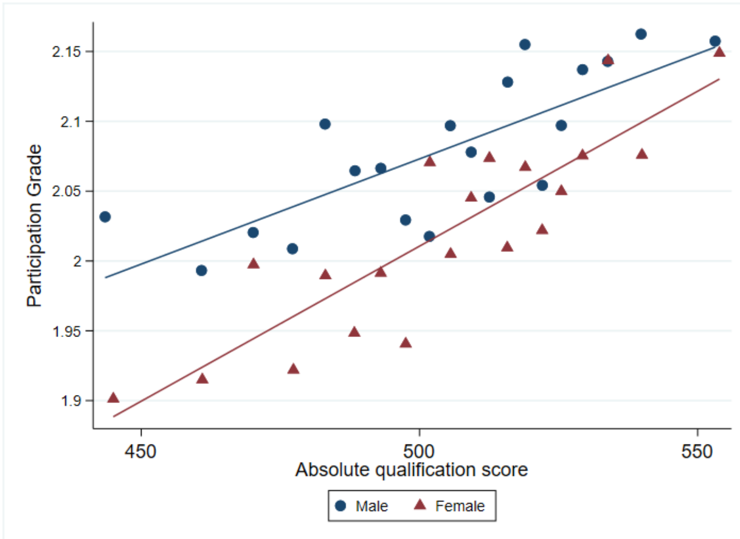
Figure XVII
Mechanisms: Social Network



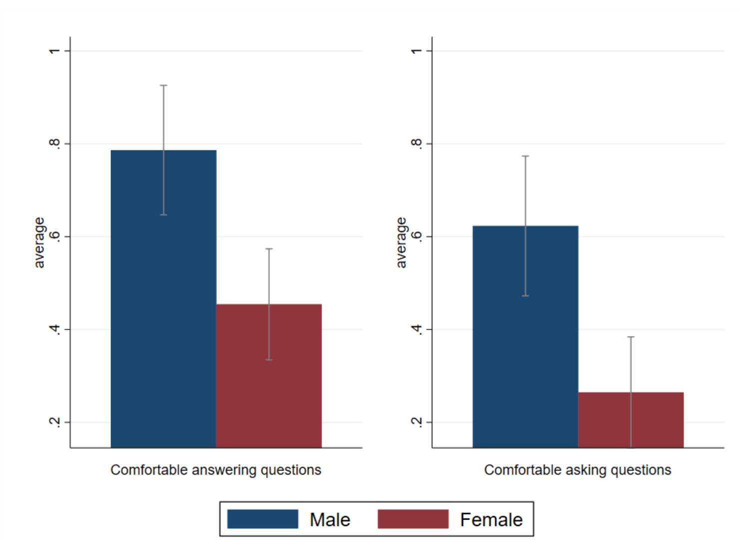
Notes: The figure displays students' replies to the following questions: "Thinking about 5 of your best friends/people you study with/people you ask questions on the material to, how many of them are women?. The answers were standardized in order to display on the y-axis the number of same gender people they nominated. An answer equal to 2.5 indicates that they interact with peers independently on their gender. The P-values for the T-test for the difference in means between Stereotypical and Counter-stereotypical are: 0.039, 0.002, 0.000. The p-value for the difference in mean between "best friends vs questions on the material" and "study buddies and questions on the material" for students enrolled in counter-stereotypical departments are 0.000 and 0.156. The p-values for the difference between the gaps for "best friends vs questions on the material" and "study buddies and questions on the material" are 0.061 and 0.008.

Figure XVIII
Mechanisms: Cost of Participating

Panel A

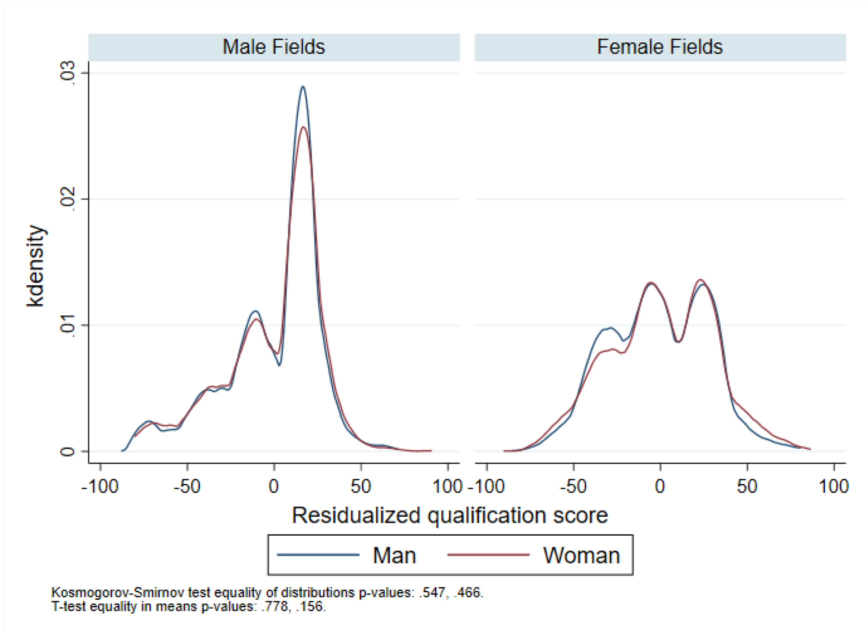


Panel B



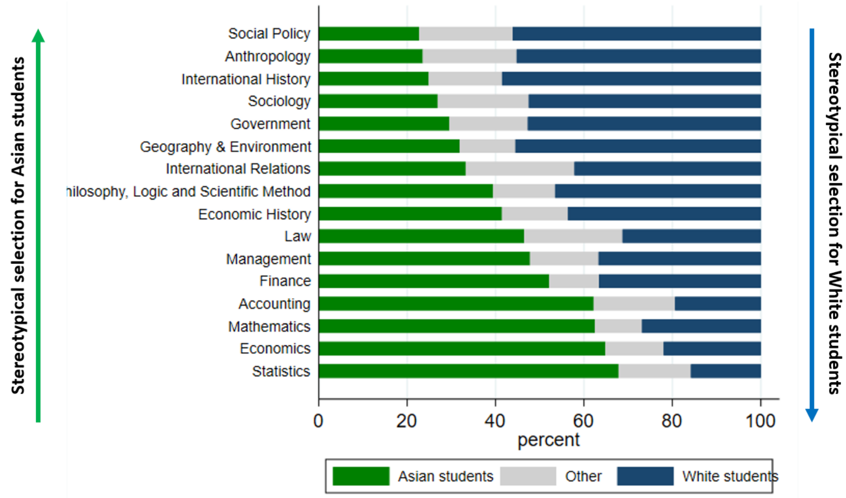
Notes: Panel A displays a binned scatterplots of class participation controlling for class group fixed effects and individual characteristics (program of enrollment, school of origin, ethnicity). Panel B displays the answers to the survey questions: "Thinking about the courses you have taken in your career at LSE, how much do you agree or disagree with the following statements? (-2: strongly disagree, +2: strongly agree) (i) I am comfortable answering questions or contributing to the discussion in class; (ii) I am comfortable answering questions or contributing to the discussion in class".

Figure XIX
Qualification Scores at Entry By Department and Gender



Notes: The graph displays the residualized qualification score, residuals of regression of qualification score on qualification type and program x year fixed effects, for men and women enrolled in Male and Female fields. Male and Female fields are the 5 departments with the highest share of men and women among undergraduates students, respectively. The statistics include the undergraduate students enrolled in LSE between 2011 and 2017 for whom I was able to reconcile information at entry with qualifications requirements at LSE: 9.449 students, 90% of the students enrolled in undergraduate programs in LSE between academic years 2011/12 and 2017/18. Following Campbell et al. (2019), I construct a measure of individual quality based on the best three exam results among the A-level qualification scores students declared when applying to LSE. Some students take courses that are equivalents to A-Levels. In these cases I calculate their A-Level equivalence scores based on the university conversion tables for foreign students. A levels are graded on a scale of A*/A/B/C/D/E. Each A-level grade is worth 30 QCA (Qualifications and Curriculum Authority) points. P-values of two-sided Kosmogorov-Smirnov tests of equality of distributions: Male Fields - 0.547, Female Fields - 0.466. P-values of ttes of equality in means: Male Fields - 0.778, Female Fields - 0.156.

Figure XX
Stereotypical Selection - Definition Ethnicity



Note: The figure displays the distribution of Asian, White, and other ethnicity students enrolled in undergraduate programs across departments. The statistics is the average of academic years 2008-2017. The green and blue arrows at the sides display the direction of choices of majors that would be in line with stereotypes and norms for Asian students and White students respectively. Other students include other ethnicity students and students who don't disclose their ethnic group. Other ethnicity students is a residual category that includes students of other ethnic groups, but also students who provided no information regarding their ethnicity.

IX. TABLES

Table I
Descriptive Statistics: Outcome Measures

	N.	Mean	SD
Panel A: Course grades			
Raw	54603	60.32	16.35
Residual grades after controlling for Course FEs	54603	0.00	15.84
Residual grades after controlling for Course and Student FEs	54603	0.00	7.94
Panel B: Participation grades			
Raw	44771	2.04	.85
Residual grades after controlling for Course FEs	44771	0.00	.82
Residual grades after controlling for Course and Student FEs	44771	0.00	.56

Notes: Panel A - The sample of the analysis on performance: students for which I can observe an exam grade in at least two courses: 54.603 course-year-class group level observations, corresponding to information on 14.313 students. The residual variation is obtained by taking the residuals of a regression of exam grades on course fixed effects, and course and student fixed effects respectively. Panel B - The sample is the sample of analysis restricted to all the students for which I can observe a participation grade in at least two class groups: 44.771 course-year-class group level observations, corresponding to information on 13.485 students. The residual variation is obtained by taking the residuals of a regression of participation grades on course fixed effects, and course and student fixed effects respectively.

Table II
Descriptive Statistics: Students' Characteristics

	N.	Mean	SD	Min	Max
Students' Characteristics:					
Females	14313	.49	.50	0	1
Age at entry	14313	18.55	1.224551	16	56
White	14313	.36	.48	0	1
Asian	14313	.47	.50	0	1
Black	14313	.04	.20	0	1
Other	14313	.04	.19	0	1
Missing	14313	.09	.28	0	1
Single sex school	9154	.33	.47	0	1
Qualification score at entry	9449	503.90	34.27	420	540
Classes and courses:					
Course size	512	138.44	163.91	12	1011
N. classes per course	512	9.36	10.77	1	69
Class size	4886	13.43	2.45	4	23
N. classes per student	14313	3.81	0.60	2	7
Class composition:					
Share of same gender class mates	54603	.56	.16	.06	1
Share of co-ethnic class mates	54603	.35	.18	.04	1
Share of same program class mates	54447	.49	.32	.04	1
Average peers qualification score at entry	35935	501.89	15.57	420	540
Outcomes:					
Course Grade	54603	60.32	16.35	0	100
Pr(complete course)	54603	.97	.17	0	1
Participation grade	44771	2.04	.85	0	3

Notes: The descriptive statistics concern the sample of the analysis on performance: students for which I can observe an exam grade in at least two courses: 54.603 course-year-class group level observations, corresponding to information on 14.313 students. The maximum number of classes for each students are 7. Students for which we observe more than 4 classes are 5%, 85% of which attend 5 courses. These are students who attend language courses or half-unit courses during Michaelmas term.

Table III
Descriptive Statistics: Variation in Share of Same-gender Classmates

	N.	Mean	SD
Raw	54603	0.56	0.162
Residual share after controlling for Course FEs	54603	0.00	0.158
Residual share after controlling for Student and Course FEs	54603	0.00	0.112

Notes: Following Olivetti et al., (2020). The table displays the share of same gender students in class and the residual share of same gender students in class obtained by taking the residuals of a regression of the class share of same gender classmates on course fixed effects, and course and student fixed effects respectively.

Table IV
Identification: Effect of Gender On The Share of Same-gender Peers

	Share of same gender peers	
	(1)	(2)
Panel A: Gender		
Female	0.003	0.001
	(0.004)	(0.004)
Course leave-out-mean	0.993***	0.994***
	(0.016)	(0.018)
Observations	54603	44771
Course fixed effects	Y	Y
Parallel course dummies	Y	Y

Notes: Standard errors are clustered at the class group level. For each course c in academic year t I test that individual characteristics cannot predict the share of same gender peers in the class g the student was assigned to. The dependent variable is the share of same gender peers. *leave-out mean*$_{ict}$ is the share of same gender peers (dependent variable) in the course for each student. Parallel course dummies are a series of dummies that are equal to one for all the students that attend the course in the same year, and zero otherwise. The omitted category are men. Column (1) sample is the sample of the course performance analysis, while column (2) restrict the sample to the analysis on participation.

Table V
Identification: Effect of Other Characteristics On The Share of Same-gender Peers

	Independent Variable: Share of females	
Dependent Variables	(1)	(2)
White	-0.005	-0.004
	(0.015)	(0.017)
Asian	0.014	0.005
	(0.016)	(0.017)
Other Ethnicity	-0.003	-0.002
	(0.009)	(0.010)
Unknown Ethnicity	-0.005	0.002
	(0.010)	(0.011)
Independent School	-0.029*	-0.036**
	(0.015)	(0.017)
State School	0.014	0.011
	(0.014)	(0.016)
Other School	0.015	0.025
	(0.017)	(0.019)
Mixed School	0.005	0.000
	(0.017)	(0.019)
Single Sex	-0.006	-0.015
	(0.014)	(0.016)
Not applicable	0.001	0.015
	(0.017)	(0.018)
Age at entry	-0.027	-0.051
	(0.040)	(0.044)
Qualification Score at entry	-0.374	-0.604
	(0.873)	(0.956)
Course Fixed Effects	Y	Y
Parallel Course Dummies	Y	Y
N	54603	44771
N. Tests performed	12	12
N. Tests significant at 1%	0	0
N. Tests significant at 5%	0	1
N. Tests significant at 10%	1	1
Share Tests significant at 1%	0	0
Share Tests significant at 5%	0	0.08
Share Tests significant at 10%	0.08	0.08
Total N. Tests performed	24	
Total Share Tests significant at 1%	0	
Total Share Tests significant at 5%	0.04	
Total Share Tests significant at 10%	0.08	

Notes: Standard errors are clustered at the class group level. For each course c in academic year t I test that individual characteristics cannot predict the share of female in the class g the student was assigned to. Column (2) restricts the sample to the sample for the analysis on participation. Each row corresponds to a separate regression where the independent variable is the share of females in the class and the dependent variable is the individual characteristic, controlling for course fixed effects, a dummy for each course the student attends during the academic year, and a female dummy. When the dependent variables are age at entry and qualification score at entry, a dummy for missing values is included in the regression.

Table VI
Main Results: Stereotypical Selection

	Course grade			
	Continuous	Top and Bottom 2	Top and Bottom 3	Top and Bottom 5
	(1)	(2)	(3)	(4)
Share of students like me	-5.937***	-3.555***	-2.884***	-1.167
	(1.701)	(1.232)	(1.042)	(0.725)
Share of students like me × Neutral		3.757***	3.154***	1.028
		(1.300)	(1.132)	(0.954)
Share of students like me × Stereotypical selection	12.390***	7.759***	6.138***	3.447***
	(3.166)	(1.555)	(1.313)	(0.944)
Course fixed effects	Y	Y	Y	Y
Student fixed effects	Y	Y	Y	Y
Observations	54603	54603	54603	54603
Mean Dependent Variable	60.320	60.320	60.320	60.320
	(16.345)	(16.345)	(16.345)	(16.345)

Notes: This table provides evidence of the results of Specification 2 in Column (1) and Specification 3 in Columns (2)-(4). In Column (1) stereotypical selection is defined as the average share of same gender students in the student i department of enrolment across academic years 2008-2017. In Columns (2) I consider as students making stereotypical choices, students who are enrolled in the top two departments with the highest average share of same gender students, while students who enrolled in the two departments with the lowest share of same gender students, as students who made a choice not in line with stereotypes regarding gender roles and skills. In Columns (3) and (4), I replicate the same analysis by defining stereotypical and counter stereotypical choices by considering the top and bottom 3 and 5 departments in terms of share of same gender students among undergraduates. The outcome variable is course grades, with incomplete courses coded as 0. Standard errors are clustered at the class group level.

Table VII
Main Results: Stereotypical Selection - Effect For Different Groups of Students

	Course grade		
	Top and Bottom 2	Top and Bottom 3	Top and Bottom 5
	(1)	(2)	(3)
Share of students like me × Counter-stereotypical Selection	-3.555***	-2.884***	-1.167
	(1.232)	(1.042)	(0.725)
Share of students like me × Neutral	0.202	0.270	-0.139
	(0.423)	(0.452)	(0.622)
Share of students like me × Stereotypical Selection	4.204***	3.254***	2.280***
	(0.901)	(0.774)	(0.587)
Course fixed effects	Y	Y	Y
Student fixed effects	Y	Y	Y
Observations	54603	54603	54603
Mean Dependent Variable	60.320	60.320	60.320
	(16.345)	(16.345)	(16.345)

Notes: This table provides evidence of the results of Specification 2 in Column (1) and Specification 3 in Columns (2)-(4). In Column (1) stereotypical selection is defined as the average share of same gender students in the student i department of enrolment across academic years 2008-2017. In Columns (2) I consider as students making stereotypical choices, students who are enrolled in the top two departments with the highest average share of same gender students, while students who enrolled in the two departments with the lowest share of same gender students, as students who made a choice not in line with stereotypes regarding gender roles and skills. In Columns (3) and (4), I replicate the same analysis by defining stereotypical and counter stereotypical choices by considering the top and bottom 3 and 5 departments in terms of share of same gender students among undergraduates. The outcome variable is course grades, with incomplete courses coded as 0. Standard errors are clustered at the class group level.

Table VIII
Main Results: Stereotypical Selection - Grades Distribution

	Pr(Drop-out)	Grade ≥ 40	Grade ≥ 50	Grade ≥ 60	Grade ≥ 70
	(1)	(2)	(3)	(4)	(5)
Panel A: Continuous definition					
Share of students like me	0.015	-0.056*	-0.131***	-0.198***	-0.089
	(0.016)	(0.031)	(0.047)	(0.066)	(0.056)
Share of students like me × Stereotypical Selection	-0.023	0.117**	0.276***	0.386***	0.207**
	(0.030)	(0.058)	(0.086)	(0.124)	(0.103)
Panel B: Top & Bottom 2:					
Share of students like me × Counter-stereotypical Selection	0.009	-0.037	-0.058*	-0.094**	-0.041
	(0.011)	(0.024)	(0.035)	(0.042)	(0.041)
Share of students like me × Neutral	0.003	-0.000	0.003	-0.014	0.018
	(0.003)	(0.008)	(0.012)	(0.017)	(0.015)
Share of students like me × Stereotypical selection	-0.000	0.053***	0.093***	0.132***	0.060*
	(0.007)	(0.016)	(0.026)	(0.034)	(0.031)
Panel B: Top & Bottom 3:					
Share of students like me × Counter-stereotypical Selection	0.009	-0.037*	-0.056*	-0.098***	-0.009
	(0.010)	(0.019)	(0.029)	(0.037)	(0.035)
Share of students like me × Neutral	0.003	0.002	0.003	-0.006	0.010
	(0.004)	(0.009)	(0.013)	(0.019)	(0.015)
Share of students like me × Stereotypical selection	0.000	0.041***	0.079***	0.091***	0.062**
	(0.006)	(0.014)	(0.022)	(0.029)	(0.027)
Panel B: Top & Bottom 5:					
Share of students like me × Counter-stereotypical Selection	0.005	-0.010	-0.022	-0.044	-0.011
	(0.006)	(0.014)	(0.021)	(0.029)	(0.025)
Share of students like me × Neutral	0.006	-0.001	-0.008	-0.025	0.012
	(0.005)	(0.011)	(0.018)	(0.024)	(0.021)
Share of students like me × Stereotypical selection	-0.001	0.022**	0.056***	0.064***	0.046**
	(0.005)	(0.011)	(0.017)	(0.023)	(0.020)
Course fixed effects	Y	Y	Y	Y	Y
Y					
Student fixed effects	Y	Y	Y	Y	Y
Y					
Observations	54603	54603	54603	54603	54603
Mean Dependent Variable	0.029	0.931	0.843	0.628	0.232
	(0.167)	(0.254)	(0.364)	(0.483)	(0.422)

Notes: This table provides evidence of the results of Specification 2. In Column (2) the dependent variable is a dummy equal to one if students drop-out from the course and zero otherwise. In Columns (3) to (6) the outcome variables are a series of dummies that are equal to one if grades are greater or equal to 40, 50, 60, and 70, respectively. Standard errors are clustered at the class group level.

Table IX
Main Results: Stereotypical Selection - Gender Differences

	Course grade			
	Continuous	Top and Bottom 2	Top and Bottom 3	Top and Bottom 5
	(1)	(2)	(3)	(4)
Female x Share of students like me	-2.432	-2.982**	-2.344*	-0.449
	(2.169)	(1.332)	(1.215)	(1.022)
Female x Share of students like me × Neutral		3.722**	2.985**	0.811
		(1.462)	(1.380)	(1.368)
Female x Share of students like me × Stereotypical Selection	5.252	5.536***	4.509***	1.186
	(3.866)	(2.120)	(1.635)	(1.196)
Male x Share of students like me	-10.508***	-5.910*	-4.348**	-2.238**
	(2.535)	(3.135)	(2.006)	(0.957)
Male x Share of students like me × Neutral		5.555*	4.233**	1.573
		(3.195)	(2.107)	(1.298)
Male x Share of students like me × Stereotypical Selection	21.279***	10.655***	8.112***	5.664***
	(4.838)	(3.311)	(2.256)	(1.319)
Course fixed effects	Y	Y	Y	Y
Student fixed effects	Y	Y	Y	Y
Observations	54603	54603	54603	54603
Mean Dependent Variable	60.320	60.320	60.320	60.320
	(16.345)	(16.345)	(16.345)	(16.345)
Gender differences in effect:				
Share of students like me	-8.008**	-2.876	-1.974	-1.790
	(3.254)	(3.346)	(2.314)	(1.390)
Share of students like me × Stereotypical Selection	15.906***	5.071	3.576	4.463**
	(6.041)	(3.710)	(2.703)	(1.760)
Effect for students who made a stereotypical selection:				
Female		2.554	2.165*	.737
		(1.633)	(1.106)	(0.639)
Male		4.745***	3.765***	3.426***
		(1.069)	(1.008)	(0.895)

Notes: This table provides evidence of the results of Specification 2 (Column 1) and Specification 3 (Column 2-4) interacting Share of students like me and Share of students like me × Stereotypical Selection by a dummy equal to one if the student is female and a dummy equal to one if the student is male. In column (1) stereotypical selection is defined as the average share of same gender students in the student's department of enrolment across academic years 2008-2017. In column (2) I consider as students making stereotypical choices, students who are enrolled in the top two departments with the highest average share of same gender students, while students who enrolled in the two departments with the lowest share of same gender students as students who made a choice not in line with stereotypes regarding gender roles and skills. In Columns (3) and (4), I replicate the same analysis by defining stereotypical and counter stereotypical choices by considering the top and bottom 3 and 5 departments in terms of share of same gender students among undergraduates. The outcome variable is course grades, with incomplete courses coded as 0. Standard errors are clustered at the class group level. The additional tests at the bottom of the table display a test of the differences between the estimated coefficient for men and women for each specification (Gender differences in effect), and the estimated effect of an increase in the share of same gender classmates for students who made a stereotypical selection: sum of the coefficient of the Share of students like me and Share of students like me interacted with Stereotypical Selection for men and women (Effect for students who made a stereotypical selection).

Table X
Alternative Mechanisms: Peers' Characteristics and TA Fixed Effects

	Course grade				
	(1)	(2)	(3)	(4)	(5)
Panel A: Continuous definition					
Share of students like me	-5.937***	-4.033**	-5.999***	-3.175*	-6.549***
	(1.701)	(1.933)	(1.701)	(1.703)	(2.094)
Share of students like me × Stereotypical Selection	12.390***	8.757**	12.515***	6.953**	13.191***
	(3.166)	(3.619)	(3.164)	(3.171)	(3.905)
Panel B: Top & Bottom 2:					
Share of students like me	-3.555***	-2.173	-3.610***	-2.420**	-4.524***
	(1.232)	(1.353)	(1.232)	(1.230)	(1.606)
Share of students like me × Stereotypical Selection	7.759***	5.570***	7.803***	5.186***	8.848***
	(1.555)	(1.765)	(1.554)	(1.552)	(2.023)
Panel B: Top & Bottom 3:					
Share of students like me	-2.884***	-2.135*	-2.926***	-1.914*	-3.626***
	(1.042)	(1.153)	(1.042)	(1.045)	(1.303)
Share of students like me × Stereotypical Selection	6.138***	4.911***	6.165***	3.962***	7.081***
	(1.313)	(1.485)	(1.312)	(1.321)	(1.663)
Panel B: Top & Bottom 5:					
Share of students like me	-1.167	-0.649	-1.204*	-0.596	-1.182
	(0.725)	(0.813)	(0.724)	(0.722)	(0.888)
Share of students like me × Stereotypical Selection	3.447***	2.447**	3.489***	2.208**	3.187***
	(0.944)	(1.077)	(0.942)	(0.940)	(1.176)
Course fixed effects	Y	Y	Y	Y	Y
Student fixed effects	Y	Y	Y	Y	Y
TA fixed effects	X	Y	X	X	X
Share of same ethnicity students	X	X	Y	Y	X
Share of same school students	X	X	Y	Y	X
Share of same program students	X	X	X	Y	X
Classmates' average qualification score at entry	X	X	X	X	Y
Observations	54603	51622	54603	54447	35883

Notes: This table provides evidence of the results of Specification 2 (Panel A) and Specification 3 (Panel B). In Panel A stereotypical selection is defined as the average share of same gender students in the student's department of enrolment across academic years 2008-2017. In Panel B I consider as students making stereotypical choices, students who are enrolled in the top two departments with the highest average share of same gender students,while students who enrolled in the two departments with the lowest share of same gender students, as students who made a choice not in line with stereotypes regarding gender roles and skills. I replicate the same analysis by defining stereotypical and counter stereotypical choices by considering the top and bottom 3 and 5 departments in terms of share of same gender students among undergraduates (Top Bottom 3, Top Bottom 5). The outcome variable is course grades, with incomplete courses coded as 0. Standard errors are clustered at the class group level. Column (1) contains the basic specification, Column (2) includes teaching assistants fixed effects, Column (3) includes controls for the share of same ethnicity and same previous school classmates, Column (4) adds the share of same program students to Column (3) controls, Column (5) includes a control for the classmates' average qualification score at entry.

Table XI
Alternative Mechanisms: Spillovers and Mechanical Effects

	Course Grade			
	(1)	(2)	(3)	(4)
Share of students like me	-5.321***	-4.598**	-9.757***	-9.308**
	(1.566)	(1.818)	(2.067)	(4.511)
Stereotypical Selection			-6.365***	-6.294
			(2.093)	(4.538)
Share of students like me × Stereotypical Selection	11.116***	9.886***	18.920***	18.321**
	(2.892)	(3.399)	(3.802)	(8.211)
Observations	54603	54603	54603	14313
Course × year FEs	Y	X	Y	Y
Student FEs	Y	Y	X	X
Parallel Courses	X	X	Y	Y

Notes: This table provides evidence of the results of Specification 2 in Column (1). Column (2) displays the results of Specification 2 with no course fixed effects. Columns (3) and (4) display the results obtained by exploiting the within-course variation: Specification 2 with no student fixed effects, but with controls for the other courses attended by each student during the academic year. Column (3) includes all the observations, while (4) exploits one observation per student. Standard errors are clustered at the class level.

Table XII:
Alternative Mechanisms: Placebo Tests

	Course grade		
	(1)	(2)	(3)
Share of students like me	-5.937***	-6.006***	-3.097*
	(1.701)	(1.701)	(1.701)
Share of students like me × Stereotypical Selection	12.390***	12.534***	6.919**
	(3.166)	(3.165)	(3.170)
Share of co-ethnic classmates		1.987	
		(1.860)	
Share of co-ethnic classmates × Stereotypical Selection		-0.537	
		(3.422)	
Share of same program classmates			4.707***
			(0.967)
Share of same program classmates × Stereotypical Selection			-2.291
			(1.770)
Course fixed effects	Y	Y	Y
Student fixed effects	Y	Y	Y
Observations	54603	54603	54447

Notes: This table provides evidence of the results of Specification 2 in Column (1). Column (2) displays the results of Specification 2 with additional controls for Stereotypical selection interacted with the share of same ethnicity classmates (Column 2), and the share of same program classmates (Column 3). Standard errors are clustered at the class level.

Table XIIIDepartments Categorization

	LSE Undergraduate Enrolment (1)	LSE Undergraduate Applications (2)	UK HESA Undergraduate Enrolment (3)	UK HESA Staff (4)
Finance	0.31	0.32	0.42	0.43*
Economics	0.35	0.35	0.33	0.30
Mathematics	0.38	0.36	0.39	0.23^{+}
Philosophy, Logic and Scientific Method	0.39	0.39	0.47	0.29
Economic History	0.39	0.33	0.53^{+}	
Statistics	0.41	0.40	0.43	0.23^{+}
Government	0.45	0.45	0.47*	0.37^{x}
Accounting	0.48	0.43	0.46	0.43*
Management	0.51	0.48	0.47	0.43*
Geography & Environment	0.51	0.53	0.56	0.40
International History	0.51	0.49	0.53^{+}	0.42
Law	0.61	0.58	0.63	0.51
Social Policy	0.66	0.62	0.69	0.65
International Relations	0.67	0.62	0.47*	0.37^{x}
Anthropology	0.76	0.73	0.74	0.51
Sociology	0.78	0.76	0.75	0.55

Notes: Columns (1)-(4) display the share of females among students enrolled in undergraduate programmes at LSE, students who applied to undergraduate programmes at LSE, students enrolled in undergraduate programmes in UK universities, staff working in UK universities respectively. In columns (1)-(3) the share is calculated as the average of the share of females in each subject across academic years 2008-2017. In column (4) the reported share is the share across academic years 2014-2018. *,+,x are symbols used to indicate that data come from aggregate statistics since the department isn't present as a separate voice.

Table XIV
Alternative Mechanisms: Alternative Measures of Stereotypical Selection

	Course grade			
	LSE Undergraduates	LSE Applications	UK HESA Undergraduates	UK HESA Staff
	(1)	(2)	(3)	(4)
Share of students like me	-5.937***	-4.890***	-7.409***	-4.125***
	(1.701)	(1.812)	(1.739)	(1.401)
Share of students like me × Stereotypical Selection	12.390***	10.410***	15.313***	8.934***
	(3.166)	(3.411)	(3.265)	(2.645)
Course fixed effects	Y	Y	Y	Y
Student fixed effects	Y	Y	Y	Y
Observations	54603	54603	54603	52623

Notes: This table provides evidence of the results of Specification 2 using the share of same gender students among students enrolled in undergraduate programmes at LSE as proxy for stereotypical selection in Column (1). Columns (2)-(4) display the results of Specification 2, using different proxies for stereotypical selection: the share of same gender students among students who applied to undergraduate programmes at LSE (Column 2), the share of same gender students among students enrolled in undergraduate programmes in UK universities (Column 3), the share of same gender staff working in UK universities in each field (Column 4). Information on the data used to create the alternative proxies of stereotypical selection can be found in Table XIII. Standard errors are clustered at the class level.

Table XV
Gender Stereotypes: Gender of Teaching Assistant

	Course Grade			
	Continuous	Top and Bottom 2	Top and Bottom 3	Top and Bottom 5
	(1)	(2)	(3)	(4)
Same gender TA	-1.309*	-0.461	-0.417	-0.002
	(0.775)	(0.558)	(0.462)	(0.376)
Same gender TA × Neutral		0.192	0.105	-0.421
		(0.594)	(0.512)	(0.457)
Same gender TA × Stereotypical selection	2.635*	2.105***	1.669***	0.664
	(1.461)	(0.764)	(0.629)	(0.538)
N	26669	26669	26669	26669
Course Fixed Effects	Y	Y	Y	Y
Student Fixed Effects	Y	Y	Y	Y

Notes: This table provides evidence of the results of Specification 2, interacting Share of students like me with a dummy equal to 1 if the teacher's gender is the same as the student's gender. The sample is restricted to classes where I have information on the gender of the teaching assistant. In Column (1) I am considering all students, in Column (2) students who are enrolled in counter-stereotypical departments, in Column (3) students who are enrolled in stereotypical departments. The definition of Stereotypical selection is based on top and bottom 5 departments in terms of share of same gender students enrolled.

Table XVI
Ethnicity Results: Stereotypical Selection

	Course grade	
	(1)	(2)
Panel A: White students:		
Share of students like me	-2.041	-2.034
	(1.978)	(1.977)
Share of students like me × Stereotypical selection	9.369**	9.366**
	(4.394)	(4.391)
Panel B: Asian students:		
Share of students like me	-5.615***	-5.618***
	(1.869)	(1.869)
Share of students like me × Stereotypical selection	13.980***	13.993***
	(3.659)	(3.659)
Panel C: Other students:		
Share of students like me	-11.798**	-11.746**
	(4.840)	(4.837)
Share of students like me × Stereotypical selection	62.967***	62.692***
	(23.808)	(23.793)
Share of same gender classmates		0.549
		(0.368)
Course fixed effects	Y	Y
Student fixed effects	Y	Y
Observations	54603	54603

Notes: This table provides evidence of the results of Specification 6. The outcome variable is course grades, with incomplete courses coded as 0. Stereotypical selection is defined as the average share of same ethnicity students in the student i department of enrolment across academic years 2008-2017. In Column (2) I additionally control for the share of same gender classmates. Standard errors are clustered at the class group level.

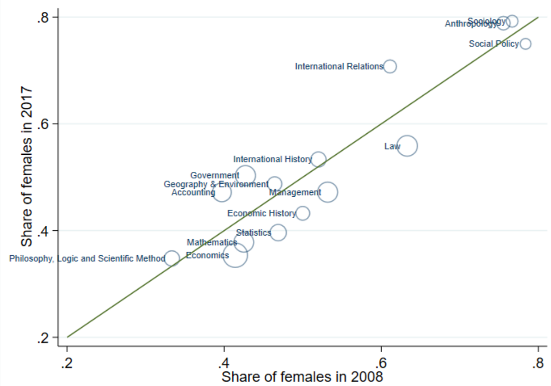
Table XVII
Gender stereotypes: Global Gender Gap Index

	Course Result	
	Counter-Stereotypical	Stereotypical
	(1)	(2)
Panel A: All		
Share of students like me	-0.758	3.463***
	(1.376)	(1.101)
GGI Tercile=2 × Share of students like me	0.633	-0.761
	(1.863)	(1.474)
GGI Tercile=3 × Share of students like me	0.918	-3.275**
	(1.984)	(1.526)
N	11290	20255
Course Fixed Effects	Y	Y
Student Fixed Effects	Y	Y

Notes: This table provides evidence of the results of Specification 1, interacting Share of students like me with a dummy equal to 1 if the student's country of origin belongs to the second tercile of the GGI distribution or the third tercile. The reference category are students whose country of origin belongs to the first tercile of the Gender Gap Index distribution (more unequal countries). The sample is restricted to students for which I have information on the Gender Gap Index of the country of origin. In Column (1) I consider students who are enrolled in counter-stereotypical departments, in Column (2) students who are enrolled in stereotypical departments. The definition of Stereotypical selection is based on top and bottom 5 departments in terms of share of same gender students enrolled. Standard errors are clustered at the class level.

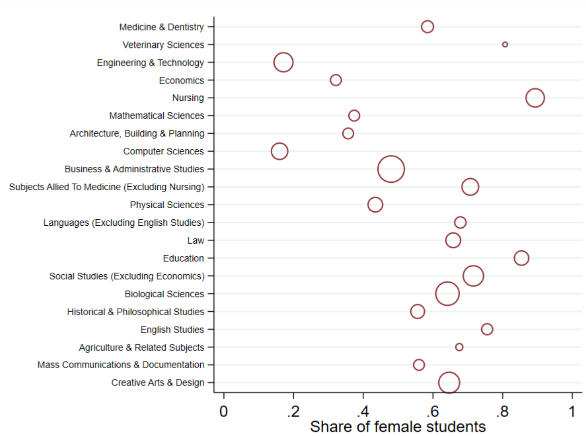
Appendix A. ADDITIONAL TABLES AND FIGURES

Figure A.1
Departments Gender Composition Over The Years



Notes: This figure illustrates the change in gender composition of LSE departments between 2008/09 and 2017/18. It is constructed based on the number of first year undergraduate students that are enrolled in bachelor programmes offered by each department in the two academic years.

Figure A.2
Enrollment UK Higher Education - Gender And Subject



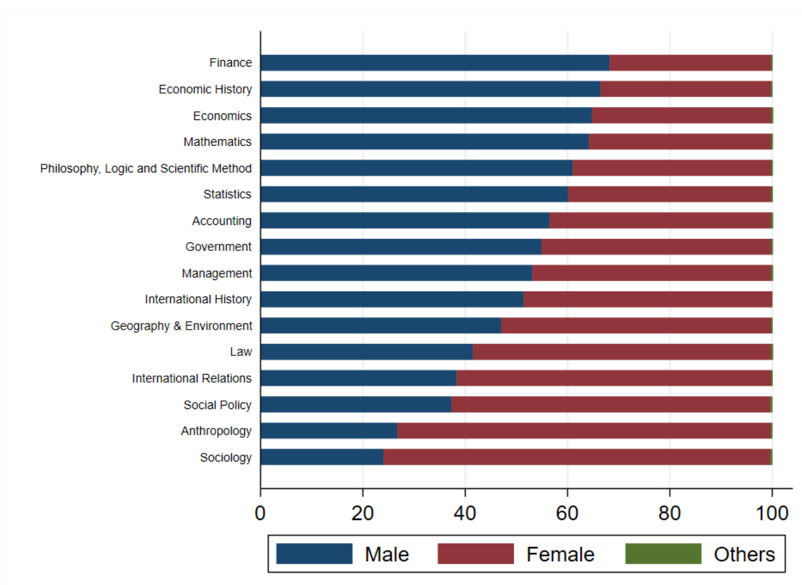
Notes: Information on the share of women among the overall population of undergraduate students enrolled in UK Universities. Source: HESA Higher Education Student Data, academic year 2018-2019.

Figure A.3
Share of Courses Outside The Department of Enrolment

	Academic year										
Department of Enrolment	**2008/09**	**2009/10**	**2010/11**	**2011/12**	**2012/13**	**2013/14**	**2014/15**	**2015/16**	**2016/17**	**2017/18**	**2018/19**
Accounting	0,75	0,75	0,75	0,75	0,75	0,80	0,80	0,80	0,80	0,80	0,75
Anthropology	0,44	0,44	0,44	0,44	0,44	0,38	0,38	0,38	0,38	0,38	0,50
Economic History	0,75	0,75	0,75	0,75	0,75	0,75	0,75	0,67	0,67	0,57	0,61
Economics	0,75	0,75	0,75	0,75	0,75	0,75	0,75	0,75	0,75	0,75	0,75
Finance									0,80	0,80	0,80
Geography & Environment	0,50	0,50	0,50	0,44	0,38	0,38	0,41	0,44	0,44	0,44	0,44
Government	0,67	0,67	0,67	0,67	0,67	0,67	0,67	0,55	0,65	0,55	0,58
International History	0,88	0,88	0,88	0,88	0,88	0,88	0,88	0,88	0,67	0,78	0,56
International Relations	0,75	0,75	0,75	0,75	0,75	0,75	0,75	0,75	0,75	0,75	0,50
Law	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Management	0,79	0,79	0,79	0,79	0,67	0,67	0,67	0,67	0,67	0,67	0,71
Mathematics	0,50	0,50	0,50	0,50	0,50	0,50	0,50	0,50	0,50	0,50	0,50
Mathematics, Statistics										0,25	0,25
Philosophy, Logic and Scientific Method	0,58	0,58	0,64	0,58	0,58	0,58	0,58	0,67	0,67	0,67	0,67
Social Policy	0,52	0,52	0,52	0,52	0,52	0,52	0,52	0,55	0,43	0,38	0,38
Sociology	0,50	0,50	0,50	0,50	0,50	0,25	0,25	0,25	0,25	0,40	0,25
Statistics	0,75	0,75	0,75	0,75	0,75	0,75	0,75	0,75	0,75	0,75	0,75

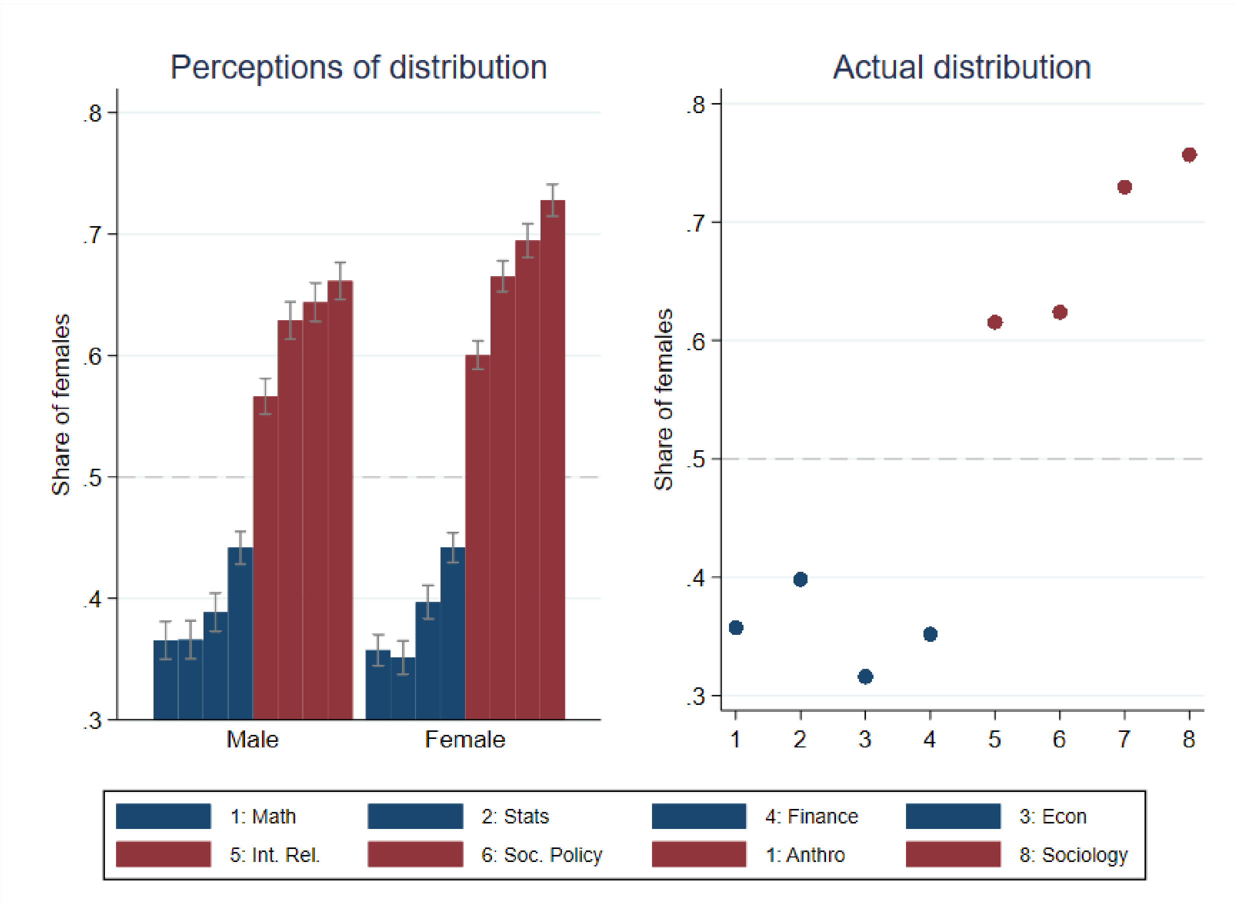
Notes: The table displays the share of courses outside the department of enrollment among courses in program for first year undergraduate students. Information are gathered from LSE Course Guides and Program Regulations for each academic year. When students can choose more than one course, the course is considered as outside the department if one of the choices is a course outside the department of enrolment. E.g. EC100 is considered outside the department for all the students who are enrolled in programs of study that do not fall under the Economics department.

Figure A.4
Gender Composition of Applicants



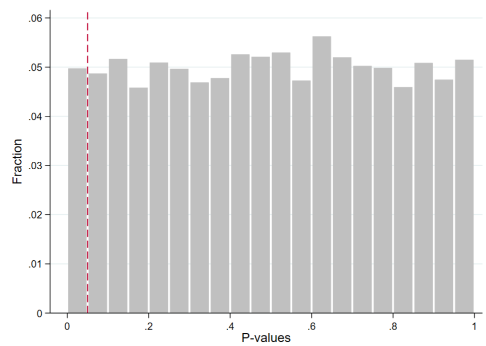
Notes: This figure illustrates the gender composition of each department. It is constructed based on the number of men and women that apply for bachelor programs at the university every year. The figure shows the average across academic years 2007/08 to 2019/20.

Figure A.5
Explicit Beliefs: Share of Females Among Applicants



Notes: Panel A displays the answers to a question regarding the application of men and women across departments at the LSE. The question asks "Consider 10 people who apply for an undergraduate program in one of the following departments at LSE. How many of them do you think are women?". Departments appeared in a random order to students in the survey. Panel B displays the actual share of women among applicants to programs in the same departments (average across academic years 2007/08 to 2019/20).

Figure A.6
Identification: Simulated Samples P-values



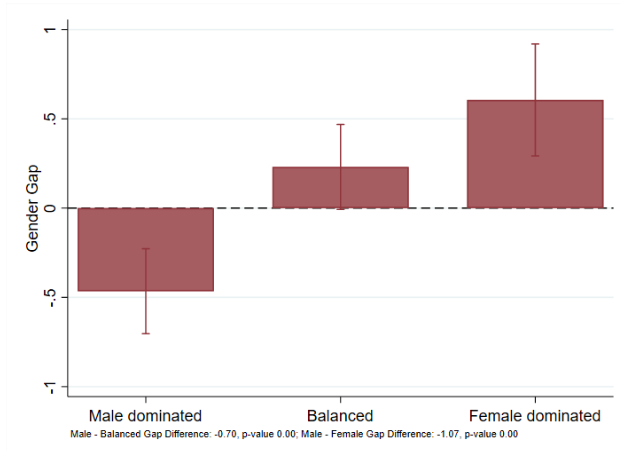
Notes: This graph show the p-values of a test of joint significance of the class dummies from a regression of gender on class dummies for each first year course in each academic year for each of the 1000 randomly simulated allocations. The sample is restricted to the courses that have at least 2 classes. The test is performed on all the students that attend first year courses and that did not change classes (which are the students for whom I can recover the initial class allocation).

Table A.1
Ethnicity Robustness: Stereotypical Selection - No Other

	Course grade	
	(1)	(2)
Panel A: White students:		
Share of students like me	-1.743	-1.731
	(1.991)	(1.989)
Share of students like me × Stereotypical selection	9.350**	9.332**
	(4.411)	(4.408)
Panel B: Asian students:		
Share of students like me	-6.195***	-6.193***
	(1.874)	(1.874)
Share of students like me × Stereotypical selection	14.374***	14.381***
	(3.659)	(3.659)
Share of students like me		0.769*
		(0.400)
Course fixed effects	Y	Y
Student fixed effects	Y	Y
Observations	45471	45471

Notes: This table provides evidence of the results of Specification 6. The outcome variable is course grades, with incomplete courses coded as 0. Stereotypical selection is defined as the average share of same ethnicity students in the student i department of enrolment across academic years 2008-2017. In Column (2) I additionally control for the share of same gender classmates. The sample is restricted to students who declared Asian or White as their ethnic group. Standard errors are clustered at the class group level.

Figure A.7
Gender gap: Course Performance



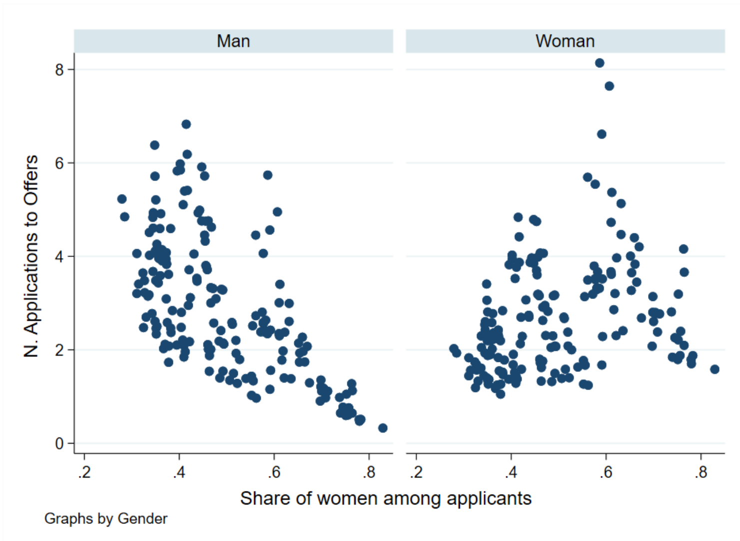
Notes: The gender gap is calculated as the difference in course performance between women and men enrolled in the same course in the same year. The sample contains all the courses attended by students enrolled in undergraduate programs in the department during the bachelor. Departments are categorized based on the top and bottom 5 definition of stereotypical selection.

Table A.2
Gender gap: Class Participation

	Class participation	
	(1)	(2)
Female	-0.089***	-0.091***
	(0.014)	(0.013)
Female × Balanced	-0.046**	-0.052***
	(0.020)	(0.019)
Female × Female Fields	-0.060***	-0.068***
	(0.021)	(0.021)
Balanced	0.023	0.006
	(0.017)	(0.017)
Female Fields	0.016	-0.009
	(0.024)	(0.025)
Observations	44771	44661
Course fixed effects	Y	X
Class group fixed effects	X	Y

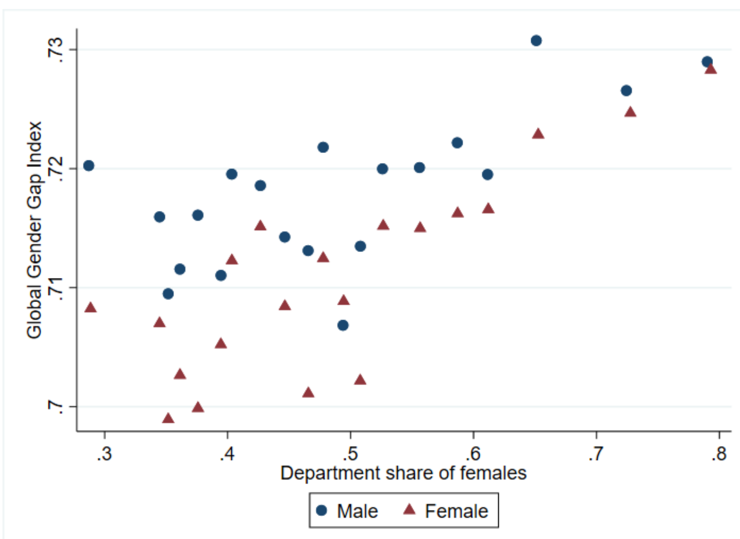
Notes: This table provides evidence of the results of a specification where class participation is regressed on course × year fixed effects (Column 1) the female dummy interacted with a dummy equal to one if the department the student is enrolled in is a balanced department and a dummy equal to one if the department the student is enrolled in is a female dominated department. The reference category are male dominated departments. The definition of female, balanced and male dominated departments is based on top and bottom 5 departments in terms of share of female students among undergraduate students. Standard errors are clustered at the class level. The outcome variable is participation grades. In Column (2) I control for class fixed effects instead of course × year fixed effects.

Figure A.8
Number of Applications By Number of Offers - Gender



Notes: Information on undergraduate students applying to LSE in academic years 2007-2018. The y-axis displays the ratio of number of applications to number of offers in each academic year for each department. On the x-axis, departments are ordered based on the share of women among applicants, from lowest to highest.

Figure A.9
Selection Into Fields: Gender Gap Index



Notes: This is a binned scatter-plot that displays the relationship between the gender gap index in student's country of origin (y-axis) and the share of females in the departments the student is enrolled into (x-axis), controlling for year fixed effects. Sample: students enrolled in the first year of undergraduate programs for which I have information on the GGI of the country of origin.

Appendix B. APPENDIX TO DATA

B.1. Students Changing Class Groups

Students have the possibility to change class group during the term. Students are not allowed to change class group whenever they fancy, but changes are allowed only under particular circumstances, i.e. if the student is not able to follow the allocated seminar due to clashes with other courses that arose during the term, or external circumstances. In order to be able to change class group, they have to submit an official request. Considering the sample of analysis, this happens 5.9% of times. If students changed class group in a systematic way, class group allocation would not be exogenous anymore. In order to test that the decision to change class group does not depend on the gender composition of the class group, and that omitting students who changed class group does not generate bias, I perform two tests.

Table A.3 and A.4 show the results of regressions where a dummy equal to one when a student changed class group is regressed on gender. Columns (2) to (5) include also class group fixed effects, while Columns (3) and (5) include a control for parallel courses, a dummy for each course the student attend in Michaelmas term. In Table A.3 the sample is restricted to all the students in analysis (first year undergraduate students who attended the courses for the first time), while Table A.4 includes all the students that are allocated to a class group where a student in analysis is allocated, i.e. all the students that will contribute in defining class group gender composition. Both tables are divided into two panels. Panel A considers all the class groups the student has been allocated to during the term. If a student changed class group once during Michaelmas term, the student will be present in the sample twice for the same course. Panel B consider only one class group per student. If a student changed class group, I randomly select one of the groups to which he has been allocated.

In order to test whether the decision to change class group is independent on the group composition, I reconstruct the initial and final class group allocation for the students who changed class group. Since students normally keep the class group to which they have been allocated for all the three terms in the academic year, for 87.68% of students I am able to identify final and initial allocation by assuming that, among the two class groups observed, the final allocation is the class group the student is allocated to in Lent term. Table A.5 shows the results of a regression on the sample of all students who change class group in Michaelmas term for which I am able to reconstruct initial and final allocation, of the share of females in the class group on a dummy equal to one when the class group corresponds to the reconstructed initial allocation and zero if it corresponds to the final allocation (Initial allocation), a dummy for whether the student is a female, and the interaction of the two.

We can see from Panel B of Table A.4 that women have a slightly higher probability of changing class group. However, this probability is very small, between 0.04 and 0.07 percentage points higher with respect to a man in the same class group[63]. Furthermore, we can see from Table that the decision to change class group is independent to the gender composition of the class group. Since I am excluding all the students who changed class group when I construct the measure of class group composition, the share of female

[63]To get a correct estimate of this probability we have to look at Panel B, since in Panel A people who changed class group are oversampled

peers will be underestimated, but in an homogeneous way and independently to class group composition.

Lastly, Panel B of Table shows that once I control for parallel courses, the female dummy coefficient becomes insignificant, and Table shows that for the student in analysis, class group changes are independent on class group composition. This means that when I exclude the observation for students who change class group, I am creating an unbalanced sample, but this shouldn't generate a bias since, once I control for parallel courses (or equivalently student fixed effects) this does not vary with the gender of the student, and this is orthogonal to class group composition.

Table A.3
Decision To Change Class Group - Analysis Sample

	Dummy=1 if student changed class group				
	(1)	(2)	(3)	(4)	(5)
Panel A: All class groups					
Female	0.008**	0.008***	0.006**	0.013***	0.010**
	(0.003)	(0.002)	(0.002)	(0.004)	(0.004)
Observations	61696	61696	61696	37584	37584
Panel B: Random class groups					
Female	0.004**	0.003*	0.003	0.006*	0.005
	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)
Observations	58593	58593	58593	34481	34481
class group Fixed Effects		x	x	x	x
Parallel Courses			x		x

Notes: The Table shows the results of a regression where the independent variable is a dummy equal to one if the student changed class group. The sample is restricted to Michaelmas term, first year undergraduate students who attended the courses for the first time. Panel A considers all the class groups the student has been allocated to during the term. If a student changed class group, the student will be present in the sample twice for the same course. Panel B consider only one class group per student. If a student changed class group, I randomly select one of the groups to which he has been allocated. Columns (1)-(3) include all class groups, columns (4) and (5) only class groups where at least one student changed class group. t statistics from standard errors clustered at class group level in parentheses * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.4
Decision To Change Class Group - All Students

	Dummy=1 if student changed class group				
	(1)	(2)	(3)	(4)	(5)
Panel A: All class groups					
Female	0.009***	0.011***	0.009***	0.018***	0.015***
	(0.002)	(0.002)	(0.002)	(0.004)	(0.004)
Observations	76280	76280	76280	47266	47266
Panel B: Random class groups					
Female	0.005**	0.005**	0.004**	0.008**	0.007**
	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)
Observations	72031	72031	72031	43017	43017
class group Fixed Effects		x	x	x	x
Parallel Courses			x		x

Notes: The Table shows the results of a regression where the independent variable is a dummy equal to one if the student changed class group. The sample consists of all the students that have been allocated to a class group where a student in analysis has been allocated. Panel A considers all the class groups the student has been allocated to during the term. If a student changed class group, the student will be present in the sample twice for the same course. Panel B consider only one class group per student. If a student changed class group, I randomly select one of the groups to which he has been allocated. Columns (1)-(3) include all class groups, columns (4) and (5) only class groups where at least one student changed class group. t statistics from standard errors clustered at class group level in parentheses * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.5
Decision To Change Class Group - Initial Allocation

	class group share of females					
	(1)	(2)	(3)	(4)	(5)	(6)
Initial allocation	0.007	0.007	0.007	0.010	0.010	0.010
	(0.009)	(0.008)	(0.008)	(0.008)	(0.007)	(0.007)
Initial allocation × Female	0.005	0.005	0.005	0.003	0.003	0.003
	(0.011)	(0.008)	(0.008)	(0.010)	(0.007)	(0.007)
Female	0.127***	0.072***	0.069***	0.125***	0.067***	0.064***
	(0.008)	(0.006)	(0.006)	(0.007)	(0.005)	(0.005)
Course Fixed Effects		x	x		x	x
Parallel courses			x			x
Included Students	Analysis	Analysis	Analysis	All	All	All
Observations	5222	5222	5222	7040	7040	7040

Notes: The Table shows the results of a regression where the share of females in the class group is regressed on "Initial allocation", a dummy equal to one when the class group corresponds to the reconstructed initial allocation and zero if it corresponds to the final allocation, a dummy for whether the student is a female, and the interaction of the two. The sample consists of all the students who changed class group once in Michaelmas term, for whom I was able to reconstruct the initial and final class group allocation (87.68% of cases). Since students normally keep the class group to which they have been allocated for all the three terms in the academic year, the final allocation is assumed to be the class group the student is allocated to in Lent term. t statistics from standard errors clustered at class group level in parentheses * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

B.2. Participation grade attrition checks

Teaching assistants are required to give a performance and participation grade to all the students in the seminar. However, even if assessment is in principle compulsory, not all the teaching assistants give a feedback to the students. In Michaelmas term, participation is Missing for 16.75% of course-year-class group level observations, thus, the dataset is an unbalanced panel. However, the reason why class group participation is Missing is because the teaching assistant didn't give a grade to anybody in the class group.

As it can be seen in A.6 in 75% of cases, participation is Missing for all the students in the class group, while in the remaining of the cases, it is Missing because the student dropped the course or changed class group before the term ended, or the student never attended the seminar. In Table A.7 we can see that women have a slightly lower probability (1 percentage point) of having a participation grade, but this effect disappears when we control for course fixed effects and class group fixed effects. This indicates that women don't have a lower probability of having a participation grade, but the probability that teaching assistants give a participation grade to students is lower in courses where there are more women. The same is true for ethnicity. Regarding age of entry, even controlling for class group fixed effects, an additional year of age at entry decreases the probability of having a participation grade by 0.1 percentage points. Table A.8 confirms the fact that the probability of having a participation grade does not depend on the share of females or the share of students from particular ethnic groups in the class group.

Table A.6
Attrition Checks: Participation Grade - Reasons

Fraction of class groups based on presence of participation information

	MT	LT	ST
None or only general course students (compulsory)	0.34	0.19	0.95
Whole class	0.40	0.28	0.01
Whole class, except students who changed class group	0.05	0.13	0.00
Whole class, except students who were never present	0.04	0.06	0.02
Whole class, except students who dropped the course	0.01	0.06	0.00
Whole class, except combination of the above	0.16	0.28	0.02

Table A.7
Attrition Checks: Participation Grade - Individual Characteristics

	Pr(Participation grade)		
	(1)	(2)	(3)
Female	-0.011**	-0.003	-0.001
	(0.004)	(0.003)	(0.001)
Age on entry	-0.001	-0.001	-0.002*
	(0.001)	(0.001)	(0.001)
Arab	-0.010	0.016	0.012
	(0.022)	(0.018)	(0.008)
Asian	-0.023*	-0.010	-0.003
	(0.014)	(0.011)	(0.005)
Black	-0.020	-0.005	0.005
	(0.016)	(0.013)	(0.006)
Chinese	-0.028**	-0.002	0.002
	(0.014)	(0.011)	(0.005)
Missing	0.017	-0.008	-0.002
	(0.015)	(0.011)	(0.005)
Mixed	-0.035**	-0.017	-0.000
	(0.015)	(0.012)	(0.005)
White	-0.032**	-0.011	-0.005
	(0.014)	(0.011)	(0.005)
Observations	54603	54603	54603
Course Fixed Effects	X	Y	X
Class group Fixed Effects	X	X	Y

Notes: The Table shows the results of a regression where the independent variable is a dummy equal to one when the student has a participation grade and zero otherwise. The independent variables are background students' individual characteristics. The omitted category for ethnicity is Other. t statistics from standard errors clustered at class group level in parentheses * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

B.3. Robustness: Qualification Score At Entry

Table A.9
Correlation Between Different Measures of Ex-ante Ability

	(1)	(2)	(3)	(4)
Best 3	1			
Best 3 Not-excluded	0.995***	1		
Best 3 Preferred	0.983***	0.987***	1	
Best 3 Most common	0.951***	0.954***	0.954***	1
N.	9449			

Table A.8
Attrition Checks: Participation Grade - Class Group Composition

	Pr(Participation grade)		
	(1)	(2)	(3)
Panel A: Gender			
Share of females	-0.097***	-0.025	-0.003
	(0.031)	(0.031)	(0.027)
Panel B: Ethnicity			
Share of Indian	0.095*	-0.017	-0.020
	(0.055)	(0.053)	(0.046)
Share of Chinese	-0.027	0.016	0.007
	(0.037)	(0.042)	(0.038)
Share of Black	0.024	-0.093	-0.115
	(0.104)	(0.091)	(0.079)
Share of Other Asian	0.024	-0.008	-0.004
	(0.052)	(0.046)	(0.042)
Share of Other	0.031	-0.023	-0.059
	(0.096)	(0.078)	(0.069)
Share of Missing	0.218***	0.010	-0.013
	(0.041)	(0.052)	(0.046)
Observations	54603	54603	54603
Course Fixed Effects	X	Y	Y
Student Fixed Effects	X	X	Y

Notes: The Table shows the results of a regression where the independent variable is a dummy equal to one when the student has a participation grade and zero otherwise. In Panel A the independent variable is the share of females in the class group. In Panel B the independent variables are the share of students from different ethnic groups in the class group, and the omitted category is the share of white students. t statistics from standard errors clustered at class group level in parentheses * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Appendix C. APPENDIX TO IDENTIFICATION STRATEGY

C.1. class group allocation does not predict students' characteristics

Following Feld and Zölitz (2017) and Braga et al. (2016), I perform the following regression:

$$y_{itcg} = \sum_{g=1}^{n_c} \alpha_g * \mathbb{1}(\textit{i's group = g}) + \epsilon_{itcg},\ \forall t, c \tag{19}$$

where the dependent variable is a dummy equal to one if student i enrolled in course c in year t allocated to class group g has characteristic t (female, asian, white, age at enrolment), and the independent variables are dummies for each class group g in course c in year t. The dummy for class group g is equal to one if student i is assigned to class group g and zero otherwise. I run one regression for each combination of course c x academic year t to cover all the courses that first year undergraduate students attend in the academic years in sample [64]. The sample for each regression consists in all the students enrolled in course c in the same academic year t that didn't change class group during Michaelmas term [65]. Furthermore, the sample is restricted to all the courses that have at least 2 class groups.

I test that class group dummies are jointly significantly different from zero:

$$H_0 : \alpha_g = 0,\ \forall g = \{1; n_c\}$$

Table A.10 shows the results of the tests performed on the observed sample. The total number of combinations of courses x academic year are 576, 523 observations belong to first year courses, while 53 to second year courses. Column 'N.' displays the number of performed regressions, column 'P<0.05' displays the proportion of tests with a p-value smaller than 0.05, column 'P<0.10' displays the proportion of tests with a p-value smaller than 0.10. Columns (2) - (4) show the results of regression 19; columns (5) - (7) show the results of regression 19 in which I include a dummy for all the mandatory courses that students attend in the same academic year; columns (8) - (10) show the results of regression 19 with dummies for all the elective courses that students attend in the same academic year; and columns (11) - (13) show the results of regression 19 controlling for dummies for all the mandatory and elective courses that students attend in the same academic year. We can see that when I add the parallel course dummies to control for clashes the number of performed regressions decrease. This is due to the fact that these specifications are identified out of students that are allocated to the same class and attend the same combination of courses during the year. The more courses I include in the parallel course dummies, the higher becomes the number of courses and class groups in which I don't have enough students that attend the same combination of courses. In particular, when I control for clashes with mandatory and elective courses, I lose 99 observations (combination of courses $\times$ academic years). 85 of these courses correspond to language courses in different academic years. This makes sense if we think that language courses are not part of a bachelor program, but

[64] All first year courses plus 15 second year courses, among which 7 are language courses. Students can choose second year courses as elective courses. These correspond to 0.57% of the observations

[65] I am excluding students that changed class group since I can't observe their initial allocation. Tests of exogeneity of the decision to change class group can be found in the appendix.

can be attended by students that come from different programs. In the last 3 columns (14-16), I report the results of the tests obtained from the regressions in which I included dummies for mandatory and elective courses (columns 11-13), to which I add the results of the tests for the unconstrained regressions for all the Missing 109 combinations of courses × academic years.

We can see from the Table that the proportion of tests with a p-value smaller than 0.05 and 0.10 are respectively 5.8% and 12.1% in the unconstrained regressions, and remain slightly above the threshold (5% and 10%) also when I control for mandatory or elective courses separately. When I control for all the courses that the students contemporaneously attend during the year (mandatory and electives), the results of the tests seem to become consistent with random allocation of students into class groups, indicating that the exogenous allocation is conditional on course clashes. As a matter of fact, the fraction of p-values smaller than 5% is exactly 5% and the fraction of p-values smaller than 10% is 9%. In other words, students who take the same combination of courses are allocated to class groups independently from their gender. Lastly, the proportion of p-values below the thresholds remain below 5% and 0% also in columns (13) - (15).

Table A.10
Identification: Observed Sample P-values - Gender

	Unconstrained			Mandatory			Elective			Mandatory + Elective			Combination		
	N.	P<0.05	P<0.10	N.	P<0.05	P<0.10	N.	P<0.05	P<0.10	N.	P<0.05	P<0.10	N.	P<0.05	P<0.10
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
First year	505	0.061	0.125	505	0.071	0.119	501	0.064	0.113	441	0.050	0.091	505	0.048	0.085
Second year	48	0.02	0.083	48	0.042	0.104	48	0.042	0.063	13	0.077	0.077	48	0.021	0.0625
Total	553	0.0579	0.1211	553	0.0687	0.1175	549	0.0619	0.1093	454	0.0507	0.0903	553	0.0452	0.0832

Notes: The table shows the results of the tests of joint significance of the class group dummies in specification 19. The sample consists of all the students that are enrolled in courses attended by first year students in the academic years in analysis. The total number of combinations of courses x academic year are 553, 505 observations belong to first year courses, while 48 to second year courses. Column 'N.' displays the number of performed regressions, column 'P<0.05' displays the proportion of tests with a p-value smaller than 0.05, column 'P<0.10' displays the proportion of tests with a p-value smaller than 0.10. Columns (2) - (4) show the results of specification 19; columns (5) - (7) show the results of specification 19 in which I include a dummy for all the mandatory courses that students attend in the same academic year; columns (8) - (10) show the results of specification 19 with dummies for all the elective courses that students attend in the same academic year; and columns (11) - (13) show the results of specification 19 controlling for dummies for all the mandatory and elective courses that students attend in the same academic year. In the last 3 columns (14-16), I report the results of the tests obtained from the regressions in which I included dummies for mandatory and elective courses (columns 11-13), to which I add the results of the tests for the unconstrained regressions for all the missing 99 combinations of courses × academic years.

Table A.11 shows the result of specification 19 controlling for clashes (mandatory and electives) when the dependent variable is age on entry, a dummy equal to 1 if the student is Asian, and a dummy equal to one if the student is white respectively. The results of the tests are consistent with random allocation of students into class groups, indicating that the exogenous allocation is conditional on course clashes. In other words, students who take the same combination of courses are allocated to class groups independently from their age on entry or ethnicity.

Table A.11
Identification: Observed Sample P-values - Other Characteristics

	Mandatory + Elective								
	Age on Entry			Asian			White		
	N.	P<0.05	P<0.10	N.	P<0.05	P<0.10	N.	P<0.05	P<0.10
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
First year	405	0.040	0.094	407	0.044	0.096	409	0.056	0.090
Second year	22	0.136	0.182	21	0.00	0.048	21	0.048	0.048
Total	427	0.045	0.098	428	0.042	0.093	430	0.056	0.088

Notes: The table shows the results of the tests of joint significance of the class group dummies in specification 19. The sample consists of all the students that are enrolled in courses attended by first year students in the academic years in analysis. Column 'N.' displays the number of performed regressions, column 'P<0.05' displays the proportion of tests with a p-value smaller than 0.05, column 'P<0.10' displays the proportion of tests with a p-value smaller than 0.10. Columns (1) - (9) show the results of specification 19 controlling for dummies for all the mandatory and elective courses that students attend in the same academic year. The number of performed regressions are 427, 428 and 430 for Age on entry, Asian dummy and White dummy respectively.

C.2. Same gender peers

We might be worried that class group gender composition is defined using the full mean rather than the leave-out mean (Angrist, 2014). Here I replicate the results defining class gender composition as the share of peers like me, i.e. the share of same gender students among classmates excluding myself. We can see that this matters little as the estimated coefficients become slightly smaller in magnitude, but the results are unchanged.

Table A.12
Robustness: Share of Same-gender Peers

	Course grade			
	Continuous	Top and Bottom 2	Top and Bottom 3	Top and Bottom 5
	(1)	(2)	(3)	(4)
Panel A: Full mean				
Share of students like me	-5.937***	-3.555***	-2.884***	-1.167
	(1.701)	(1.232)	(1.042)	(0.725)
Share of students like me × Neutral		3.757***	3.154***	1.028
		(1.300)	(1.132)	(0.954)
Share of students like me × Stereotypical selection	12.390***	7.759***	6.138***	3.447***
	(3.166)	(1.555)	(1.313)	(0.944)
Panel B: Leave-out mean				
Share of peers like me	-5.413***	-3.287***	-2.681***	-1.085
	(1.587)	(1.149)	(0.967)	(0.674)
Share of peers like me × Neutral		3.468***	2.935***	0.978
		(1.211)	(1.050)	(0.884)
Share of peers like me × Stereotypical selection	11.303***	7.193***	5.680***	3.171***
	(2.949)	(1.449)	(1.219)	(0.874)
Course fixed effects	Y	Y	Y	Y
Student fixed effects	Y	Y	Y	Y
Observations	54603	54603	54603	54603

Notes: This table provides evidence of the results of Specification 2 in Column (1) and Specification 3 in Columns (2)-(4). The outcome variable is course grades, with incomplete courses coded as 0. Standard errors are clustered at the class group level. In Panel A, class group composition is defined as the share of same gender classmates, while in Panel B as the share of same gender peers, except me.

Appendix D. APPENDIX TO THEORETICAL FRAMEWORK

D.1. Definition and derivation of beliefs on ability

Following the framework proposed by Bordalo et al. (2019), individual i's belief on individual j and k's abilities are defined by the following expressions, assuming that individual j belongs to group G, while individual k belongs to group $-G$

$$a_j^b = a_{i,j}^b = a_{i\to j}^b = A_g + \mu_j + \theta_i \sigma(s_g)(A_g - A_{-g})$$
$$a_k^b = a_{i,k}^b = a_{i\to k}^b = A_{-g} + \mu_k + \theta_i \sigma_g(A_{-g} - A_g)$$

The strength of stereotypical distortions depends on how much stereotypes are salient for individual i, and on the strength of stereotypical associations for individual i ($\theta_i \sigma_g$).

Let us assume that there is a share s_g of type G students and a share $s_{-g} = (1 - s_g)$ of type $-G$ students in class. Student i's belief on the ability of the peers in the class $(-i)$ is defined by the following expression

$$a_{-i}^b = a_{i,-i}^b = E(a_{i\to -i}^b) = A_g s_g + A_{-g}(1 - s_g) + \theta_i \sigma(s_g)(2s_g - 1)(A_g - A_{-g})$$

I assume that $E_{j\in -i}(\mu_j) = 0$ and $E_{k\in -i}(\mu_k) = 0$. This implies that in absence of stereotypical distortions, i's beliefs regarding the average ability of type G students in class is equal to A_G, the average ability of individuals belonging to group G, and i's beliefs regarding the average ability of type $-G$ students in class is equal to A_{-G}, the average ability of individuals belonging to group $-G$.

Student i's belief on their relative ability will be equal to the difference between their belief regarding their own ability and the ability of the peers in the class

$$a_i^b - a_{-i}^b = (A_g - A_{-g})(1 - s_g)(1 + 2\theta_i \sigma(s_g))$$

Appendix E. APPENDIX TO ANALYSIS ALONG ETHNIC LINES

E.1. Validity of the identification strategy

In order provide evidence that students belonging to different ethnic groups are not systematically assigned to particular classes, I replicate the tests performed for gender. The first test I perform is discussed in Section. Following Feld and Zölitz (2017) and Braga et al. (2016), I test that class group allocation does not predict students' individual characteristics by regressing individual characteristics on class dummies for each course in each year. I then test that class group dummies are jointly significantly different from zero. Figure VI shows the p-values obtained from the tests of joint significance of the class group dummies for the observed sample. Regarding the White dummy, slightly more than 5% of tests display a p-value smaller than 0.05, but less than 10% of tests display a p-value smaller than 0.10. Regarding the Asian dummy, slightly less than 5% of tests display a p-value smaller than 0.05, and less than 10% of tests display a p-value smaller than 0.10. More details on the exogeneity tests and the simulation performed can be found in appendix.

In the second test, I test that students are not systematically assigned to class groups where there are more (less) students of their own ethnic group, conditional on having a certain share of same ethnicity peers among the students enrolled in a course. The specification used is the following:

$$\textit{Share of same ethnicity peers}_{iacg} = \alpha_{ca} + \beta \times \mathbb{1}(\text{Ethnic group}_i = e) + \gamma \times \textit{leave-out mean}_{ica} + \sum_{a=1}^{n_a} \sum_{p=1}^{n_{p,a}} \delta_{p,a} \times \mathbb{1}(\textit{i took course p}) + \epsilon_{iacg} \quad (20)$$

For each course c in academic year a I test that individual characteristics (Ethnic group) do not predict the composition of peers in the class group g they are assigned to. I control for the share of same ethnicity peers enrolled in the course (*leave-out mean*$_{ica}$). Lastly, I control for a series of dummies that are equal to one for all the students that attend the course in the same year, and zero otherwise ($\sum_{a=1}^{n_a} \sum_{p=1}^{n_{p,a}} \delta_{p,a} \times \mathbb{1}(\textit{i took course p})$). This is because the allocation is constrained by the fact that students attend more than one course in the same term and they cannot attend two classes at the same time. I perform this test for all the courses where there are at least two class groups and standard errors are clustered at the class group level.

Table A.13 reports the result of the above regression. Students are not assigned to class group systematically based on their ethnic group. As a matter of fact, Chinese, Asian, White, Black and other minority students do not have a higher probability of being assigned to a class group with more students of their ethnic group.

Lastly, I produce an array of "balancing tests" to study whether the variation in the share of ethnic classmates a student is allocated to is related to the variation in a number of predetermined student characteristics: gender, previous school characteristics, age at entry, and qualification score at entry. As shown in Table A.14, only two of the estimated correlations appear to be significantly different from zero for the sample of analysis. They concern the characteristics of the previous attended school.

Table A.13
Identification: Effect of Ethnicity On Share of Same-ethnicity Peers

	Share of same ethnicity peers	
	(1)	(2)
Chinese	0.006	0.007
	(0.004)	(0.004)
Other	-0.004	-0.006
	(0.004)	(0.004)
White	0.003	0.005
	(0.004)	(0.004)
Missing	-0.003	-0.002
	(0.005)	(0.005)
Course level leave-out-mean	0.992***	0.991***
	(0.012)	(0.013)
Observations	54603	44771
Course fixed effects	Y	Y
Parallel course dummies	Y	Y

Notes: Standard errors are clustered at the class group level. For each course c in academic year t I test that individual characteristics cannot predict the share of same ethnicity peers in the class group g (y_{itcg}) the student was assigned to. The dependent variable is the share of same ethnicity peers. Parallel course dummies are a series of dummies that are equal to one for all the students that attend the course in the same year, and zero otherwise. The omitted category is Other Asian. This sample includes also students that are not in their first year of a bachelor program at the university. I am excluding students that changed class group since I can't observe their initial allocation. Column (2) restricts the sample to students with no missing participation grade.

Table A.14
Identification: Effect of Other Characteristics On The Share of Same-ethnicity Peers

Dependent Variables	Independent Variable: Share of Asian (1)	(2)	Independent Variable: Share of White (3)	(4)
Female	0.001	-0.004	0.016	0.013
	(0.017)	(0.019)	(0.017)	(0.019)
Independent School	0.032**	0.029*	-0.039**	-0.027
	(0.015)	(0.017)	(0.016)	(0.017)
State School	-0.026*	-0.029*	0.036**	0.035**
	(0.014)	(0.016)	(0.014)	(0.016)
Other School	0.003	0.001	-0.000	-0.002
	(0.010)	(0.011)	(0.010)	(0.012)
Mixed School	-0.008	-0.001	-0.016	-0.007
	(0.016)	(0.017)	(0.017)	(0.019)
Single Sex	0.021	0.015	0.002	0.002
	(0.014)	(0.016)	(0.014)	(0.016)
Not applicable	-0.015	-0.007	0.014	0.005
	(0.016)	(0.017)	(0.016)	(0.017)
Age at entry	0.006	0.002	-0.059	-0.047
	(0.042)	(0.047)	(0.043)	(0.049)
Qualification Score at entry	0.438	0.024	-1.197	-1.361
	(0.871)	(0.938)	(0.916)	(0.989)
Ethnicity dummies	Y	Y	Y	Y
Course Fixed Effects	Y	Y	Y	Y
Parallel Course Dummies	Y	Y	Y	Y
Program of enrollment Fixed Effects	Y	Y	Y	Y
N	54603	44771	54603	44771

Notes: Standard errors are clustered at the class group level. For each course c in academic year t I test that individual characteristics cannot predict the share of Asian or White students in the class g (y_{itcg}) the student was assigned to. Column (2) and (4) restrict the sample to the sample for the analysis on participation. Each row corresponds to a separate regression where the independent variable is the share of Asian or White classmates and the dependent variable is the individual characteristic, controlling for course fixed effects, a dummy for each course the student attends during the academic year, ethnic group dummies, and a dummy for the program of enrollment (not included in the balance checks for gender). When the dependent variables are age at entry and qualification score at entry, a dummy for missing values is included in the regression.

E.2. Effect of changes in ethnic composition of the class on students' performance

I estimate the effect of a change in class composition along ethnic lines on students' performance estimating the following specification:

$$y_{iacg} = \alpha_{ac} + \alpha_i + \sum_{e=1}^{n_{e-1}} [\beta_{1,e} \textit{Share of ethnic group e students}_{iacg} + \\ + \beta_{2,e} \times \textit{Share of ethnic group e students}_{iacg} \times \mathbb{1}(\text{Ethnic group}_i = e)] + \epsilon_{iacg} \quad (21)$$

where the unit of analysis is student i belonging to ethnic group e, who attends course c, in the academic year a, and who is assigned to class g, and the independent variables y_{iacg} are course grades. $\mathbb{1}(\text{Ethnic group}_i = e)$ is a dummy equal to one if the student belongs to ethnic group e. Students are divided in seven groups: White, Chinese, Indian, Other Asian, Black, other ethnic minorities, and students with missing ethnic group. This is the finest possible split that could be performed. Students with missing ethnic information are included in the regression so that the omitted category are White students and the coefficients estimate the effect of an increase in the share of Chinese, Indian, Black, Other Asian classmates with respect to having additional White classmates. *Share of ethnic group e students*$_{iacg}$ is the share of classmates belonging to ethnic group e, thus $\beta_{1,e}$ will capture the effect of an increase in the share of ethnic group e classmates for students in the omitted (reference) group, White students, while $\beta_{2,e}$ will capture the effect of an increase in the share of type e classmates on the gap in course grades between students belonging to ethnic group e and the reference group (White students). The specification follows the main empirical strategy of the paper, thus I include course fixed effects, α_{acg}, and student fixed effects, α_i. The assumptions underlying the identification strategy are the same as for the main empirical strategy. The validity is discussed in the previous section.

Table A.15 display the results of Specification 21 with and without individual fixed effects (Column 2 and 3 respectively. Minority status affecting performance is not a phenomenon specific to gender, since it is able to explain part of the performance gap also for ethnic minorities. Chinese students' are the group for which performance appears to be more affected by class group composition, and the effect is robust to all the specifications. In particular, the gap in performance between Chinese and White students increases by 0.999 points when the share of Chinese students in the class group increases by 10%. This represents 22% of the raw gap between Chinese and White students in first year exams. Interestingly, Chinese do not benefit only from having more Chinese in the class, but also from having more Indian and Other Asian students with respect to White students in the class group. The effect is qualitatively the same for Indian and Other Asian students: their average performance gap with respect to white students decreases if the proportion of Asian students in the class group is higher. However, the effect is not significant. This is not very surprising for Other Asian students given that the category is a spurious category, including all Asian students who are not Chinese or Indian. Interestingly, the effect for white students mirrors exactly the effect for Asian students: white students' performance is significantly lower in class groups where there is a higher proportion of Asian students. In particular, a 10% increase in the share of Indian students decreases White students' performance by 0.368 points, significant at 1%, a 10% increase in the share of Chinese students in class decreases White students' performance by 0.240 points, significant at 1%, and a 10% increase in the share of Other Asian students decreases White students' performance by 0.192 points, significant at 5%. The proportion of Black students is very low, so it is not possible to estimate precisely an effect for this group.

Given that the effects of increasing the share of Chinese, Indian and Other Asian students are very homogeneous both for Asian students and White students, in the main analyses I aggregate students of Asian background in a composite category to increase the power of the estimates.

Table A.15
Ethnicity Results: Minority Effect

	Course grades (1)	(2)
Share of Indian students	-1.121	-3.680***
	(1.451)	(1.141)
Share of Chinese students	-1.711	-2.403***
	(1.119)	(0.876)
Share of Other Asian students	-0.331	-1.912**
	(1.280)	(0.923)
Share of Black students	1.606	1.262
	(2.105)	(1.493)
Indian	0.235	
	(0.933)	
Indian × Share of Indian students	-4.166*	3.145
	(2.437)	(2.091)
Indian × Share of Chinese students	2.810	0.621
	(1.758)	(1.556)
Indian × Share of Other Asian students	-2.210	0.564
	(2.501)	(1.979)
Indian × Share of Black students	0.705	-4.848
	(4.556)	(3.531)
Chinese	-1.447*	
	(0.749)	
Chinese × Share of Indian students	7.815***	9.166***
	(1.991)	(1.830)
Chinese × Share of Chinese students	10.146***	7.526***
	(1.531)	(1.345)
Chinese × Share of Other Asian students	6.707***	6.595***
	(1.977)	(1.593)
Chinese × Share of Black students	3.897	3.545
	(3.415)	(2.812)
Other Asian	-1.349*	
	(0.818)	
Other Asian × Share of Indian students	-0.029	4.629**
	(2.362)	(1.983)
Other Asian × Share of Chinese students	0.316	1.342
	(1.652)	(1.507)
Other Asian × Share of Other Asian students	1.100	1.310
	(2.199)	(1.708)
Other Asian × Share of Black students	-2.772	-2.885
	(3.809)	(2.779)
Black	-1.280	
	(1.344)	
Black × Share of Indian students	-0.631	1.090
	(3.749)	(3.295)
Black × Share of Chinese students	-4.292*	-3.000
	(2.585)	(2.572)
Black × Share of Other Asian students	-2.981	-0.393
	(3.699)	(3.003)
Black × Share of Black students	-0.940	-2.983
	(4.995)	(4.233)
N.	54603	54603
Class size	Y	Y
Course fixed effects	Y	Y
Student fixed effects	X	Y

Notes: The table displays the results of Specification 21 in Column (2). In Column (1), the results of the same specification without student fixed effects are presented. The outcome variable is course grades, with non-takers coded as 0. The regressions also contain the share of students with missing information on ethnicity and the share of other ethnic minorities, a dummy for missing information and other residual ethnic minorities, and their interactions with all the other ethnic dummies and share. This is done so that the omitted category are white students, and the increase in the share of each ethnic minority can be read as an increase in the share of the ethnic minorities with respect to white students. Standard errors are clustered at the class group level.

Appendix F. SURVEY

During the academic year 2020/2021 I administered a complementary online survey to students (June-July 2021). The survey was designed using Qualtrics and was sent to students to their institutional email address through the university system. The survey included 4 sections.

In the first section, I asked students some questions regarding their experience at the university, disregarding as much as they could the last year of remote teaching. Following Bursztyn et al. (2019), I asked students questions regarding the extent to which they think image is important, and what elements contribute to being popular. Moreover, following Bursztyn et al. (2017), I asked students questions regarding how much they felt comfortable participating in class. This section was meant to be used to micro-found the theoretical model.

In the second section, I asked students to perform a double-target Implicit Association Test (Greenwald et al., 1998). I followed Carlana (2019) to design a Gender-Scientific implicit association test to elicit the extent to which students automatically associate Scientific disciplines with men and Humanistic disciplines with women.

Subjects are presented with two sets of stimuli. The first set of stimuli are female and male names. Given the multicultural environment, in order to be neutral with respect to language differences and use names that could be identified by every student as clearly referring to male or female, I followed the approached used in the Gender-Science IAT designed by Project Implicit [66]. I used as female and male names words such as Man, Son, Father, Boy, Uncle, Grandpa, Husband, Male, Mother, Wife, Aunt, Woman, Girl, Female, Grandma, Daughter. The second set are words related to Scientific disciplines (e.g., Math, Physics, Engineering, etc.) and Humanistic disciplines (e.g. Literature, History, Humanities, etc.).

The IAT is composed of seven rounds. Round 1 and 2 are practice rounds of only female and male names and only Humanistic and Scientific disciplines respectively. In the following five rounds, one word at a time (either a female or male name, or a word associated with Humanistic or Scientific in a random fashion) appears on the screen and individuals are instructed to categorize it to the left or the right according to different labels displayed on the top of the screen. In "hypothesis-inconsistent" ("hypothesis-consistent") rounds individuals categorize to one side of the screen - Humanistic Male (Humanistic Female) and to the opposite side of the screen - Scientific Female (Scientific Male). The order of the two types of rounds was randomized at the individual level.

The blocks used to calculate the IAT score (d-score) are rounds 3, 4, 6, and 7. The number of words that need to be categorized is 20 in blocks 3 and 6, and 40 in blocks 4 and 7, as in the standard IAT 7-blocks (Greenwald et al. 2003). The measure of implicit association between gender and Scientific is given by the standardized mean difference score in four types of rounds. The intuition is that people with a greater implicit association of Scientific with men and Humanistic with women take longer to correctly categorize names in the "hypothesis-inconsistent pairings". Thus, the higher and more positive the d-score the stronger is the association between the two concepts. The order of the four types of blocks was randomized at the individual level. The IAT was incorporated in Qualtrics using the ad-hoc approach designed by Carpenter et

[66] implicit.harvard.edu. More details on the organization can be found below.

al. (2021)[67].

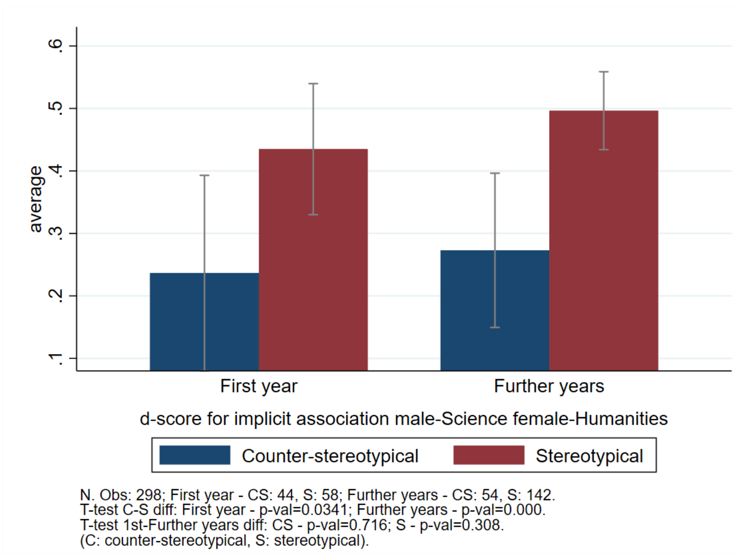
The third section of the survey consists in questions regarding explicit associations. Following Delfino (2021) and Carlana (2019), I asked questions concerning their beliefs regarding the distribution of men and women and the performance of women compared to men in different fields.

Finally, the last section of the survey contained questions regarding demographic information and the students' social network.

F.1. Gender-Scientific IAT - Evidence for first year students

You can find below the results of the implicit association test when the sample is restricted to first year students.

Figure A.10
Mechanisms: Scientific-Male, Humanistic-Female IAT



Notes: Results of an implicit association test (Greenwald et al. 1998) to elicit students' associations between female-humanistic and male-scientific. A score of 0 indicates no association between male -scientific and female-humanistic; a positive score indicates that the student associates women with humanities and men with science and math; lastly a negative score indicates that the student associates men with humanities and women with science and math. Sample restricted to respondents who are enrolled in first year undergraduate programs.

F.2. Gender-Scientific IAT - Evidence from Harvard Project Implicit

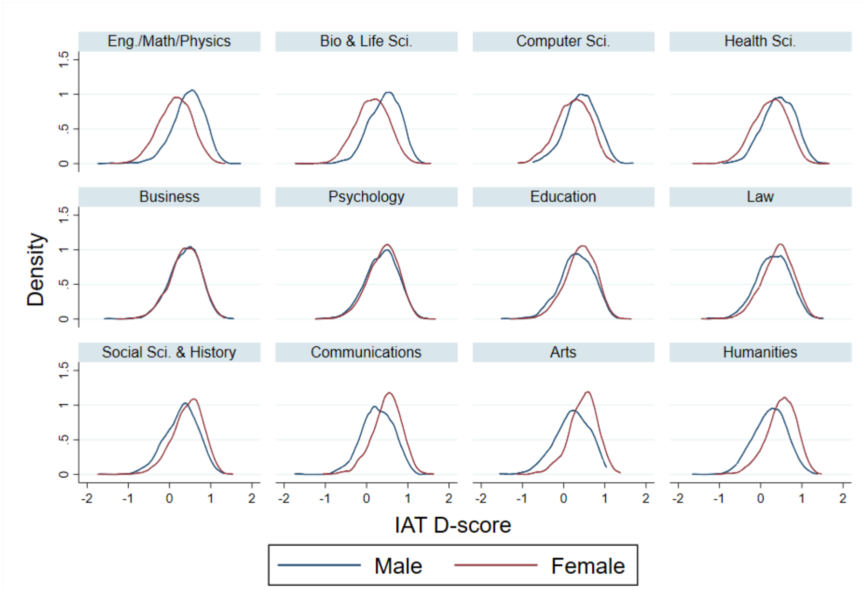
The results of the implicit association test carried out by Project Implicit can be found below. Project Implicit is a non-profit organization and international collaborative of researchers who are interested in implicit social cognition. It was founded in 1998 by three scientists – Dr. Tony Greenwald (University of Washington), Dr. Mahzarin Banaji (Harvard University), and Dr. Brian Nosek (University of Virginia). The mission of the organization is to educate the public about bias and to provide a "virtual laboratory" for collecting data on

[67] An explanation of the approach can be found at this link: https://iatgen.wordpress.com/

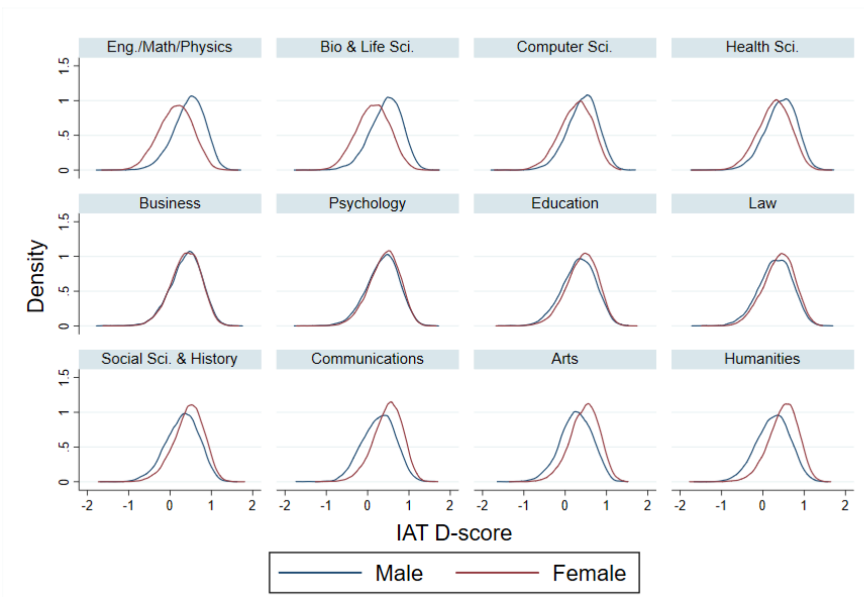
the internet. Everybody is allowed to take a test and the results are collected by the organization and made publicly available for researchers (https://osf.io/y9hiq/). The analysed sample pertains the results of the Gender-Science IAT collected from 2003 to 2020. The sample includes (a) respondents who are currently students - 63.648 observations, and (b) all the respondents that declare their university major - 421.641 respondents.

Exactly as for the survey carried out on the students enrolled at the institution, although the respondents might be a selected sample of people not representative of the overall population, this selection bias would be problematic for the interpretation of the results only if it is systematically correlated with choices of major and implicit attitudes. In other words, for the selection bias to lead to an overestimation of the effects, people who uphold stronger stereotypes who selected into stereotypical fields (e.g. men in math) should be significantly more inclined to take the test than people of the same gender who uphold weaker stereotypes and who chose an occupation/education in the same fields. At the same time, it should be the case that the sample of respondents is significantly skewed towards people who uphold weaker stereotypes who selected into counter-stereotypical fields (e.g. women in math) compared to people of the same gender who chose an occupation/education in the same fields.

Figure A.11
Project Implicit: Gender-Science IAT



(a) Currently studying



(b) All respondents

Notes: Results of an implicit association test (Greenwald et al. 1998) to elicit students' associations between female-humanistic and male-scientific. A score of 0 indicates no association between male -scientific and female-humanistic; a positive score indicates that the student associates women with humanities and men with science and math; lastly a negative score indicates that the student associates men with humanities and women with science and math. The analysed sample pertains the results of the Gender-Science IAT collected from 2003 to 2020 ((https://osf.io/y9hiq/). The sample includes (a) respondents who are currently students - 63.648 observations, and (b) all the respondents that declare their university major - 421.641 respondents.