Does the Gender Ratio at Colleges Affect High School Students' College Choices?*

Chihiro Inoue   Asumi Saito   Yuki Takahashi
Job Market Paper

January 24, 2026

[Link to the latest version]

Abstract

Female students may not choose STEM programs in college, even when they excel in mathematics, because these programs are male-dominated and would make them a minority. Using an incentivized discrete choice experiment with high school students, we show that the gender ratio affects both female and male students' college choices: they prefer gender-balanced programs and prefer being a majority to a minority, primarily because they anticipate difficulty fitting in as minorities. Importantly, these preferences differ by field: female students with high mathematics ability show stronger minority avoidance in STEM than non-STEM, while male students show weaker minority avoidance in STEM. These asymmetric preferences, together with the low female share in STEM, lead to talent misallocation: male students with low mathematics ability crowd out female students with high mathematics ability. We validate experimental choices against actual choices several months later. These findings suggest that the low female share in STEM deters female students from entering these fields.

JEL Classification: J16, J24, I24

Keywords: STEM Gender Gap, College Choice, Gender Ratio, Preference Elicitation, Discrete Choice Experiment

*Inoue: Faculty of Economics, Kobe University, email: inoue@econ.kobe-u.ac.jp. Saito: Center for Geospatial Analytics, North Carolina State University, email: asaito2@ncsu.edu. Takahashi (*corresponding author*): Department of Economics, Tilburg University. Waradelaan 2, Koopmans Building K333, 5037 AB Tilburg, The Nethelands. Phone: +31-6-4145-8415; email: y.takahashi@tilburguniversity.edu.

We thank Michihito Ando, Sonia Bhalotra, Maria Bigoni, Jochem de Bresser, Caterina Calsamiglia, Damian Clarke, Lucia Corno, Dean Hyslop, Sota Ichiba, Clément Imbert, Shoya Ishimaru, Daiji Kawaguchi, Daiki Kishishita, Annalisa Loviglio, Akira Matsushita, Francesca Miserocchi, Takeshi Murooka, Johannes Rincke, Anna Salomons, Daan van Soest, Teodora Tsankova, Kiyotaka Yageta, Atsushi Yamagishi, Basit Zafar, and participants at EEA-ESEM Rotterdam, the Trans Pacific Labor Seminar, SWET, and seminars at Comenius University, Kobe University, Maastricht University, Tilburg University, the University of Florence, the University of Gothenburg, and the University of Osaka for their valuable feedback. We are deeply grateful to NPO Waffle for their valuable insights and support at the early stages of this project. Keigo Furukawa, Mifuyu Kira, Yoshino Kodama, and Keita Yamada provided excellent research assistance. We gratefully acknowledge financial support from JSPS KAKENHI (24K22636) and the Suntory Foundation (Torii Fellowship). The experiment was approved by the Institutional Review Board of the Tilburg School of Economics and Management (approval no. IRB FUL 2023-009) and preregistered on the AEA Registry under AEARCTR-0012577.

1 Introduction

Although the gender gap is negligible in mathematics and sciences at age 15 in OECD countries (OECD 2019), the gap in Science, Technology, Engineering, and Mathematics (STEM) majors at colleges is substantial: female students are 30.8 percentage points less likely than male students to major in STEM.[1] This situation causes several social issues, such as suboptimal allocation of talent, gender-biased research topics (Truffa and Wong 2025), and gender-biased product innovation (Einiö, Feng, and Jaravel 2025; Koning, Samila, and Ferguson 2021). It can also lead to the so-called "gender data gap," where men's data is primarily used to establish medical and industry safety standards, making medical treatments and industry tools less suitable for women (Perez 2019).

One potential reason that has received little attention in the literature is that STEM programs are male-dominated and thus make female students a gender minority.[2] As a result, female students may avoid such programs in anticipation of the cost of being a minority. Indeed, the literature documents various costs, such as experiencing a worse classroom environment, having less influence in a study group, having limited peer networking opportunities, and having a higher likelihood of dropping out (Bostwick and Weinberg 2022; Gong, Lu, and Song 2021; Hampole, Truffa, and Wong 2024; Karpowitz et al. 2024).

This paper studies whether the gender ratio at colleges affects high school students' college choices and whether the low female share in STEM contributes to talent misallocation. Answering these questions with observational data poses two challenges: the gender ratio correlates with other program attributes, making it difficult to isolate its effect; and most STEM programs are male-dominated, leaving little variation to identify how higher female shares would affect choices.

We address these challenges using an incentivized discrete choice experiment with high school students. The experiment elicits preferences over independently varied program attributes – including the gender ratio across its full range – thereby isolating the effect of gender composition from confounding attributes. We incentivize truthful responses using the incentivized resume rating method (Kessler, Low, and Sullivan 2019), providing students with career advice tailored to their choices. We validate experimental choices against actual choices several months later, achieving 78% prediction accuracy.

We find that the gender ratio affects both female and male students' college choices. Both prefer gender-balanced programs over male- or female-majority programs, and both prefer being a majority to being a minority. Relative to a gender-balanced program (50% female), being a minority reduces choice probability by 11 percentage points for female students and 9 percentage points for male students; being a strong majority reduces choice probability by 6.5 percentage points for both genders. A decomposition reveals that students avoid being a minority primarily because they anticipate difficulty fitting in.

1. Based on the number of bachelor program graduates in 2023 from OECD Data Explorer, "Number of enrolled students, graduates and new entrants by field of education" dataset (accessed November 6, 2025).

2. A notable exception is Ersoy and Speer (2025), who show that providing non-job-related major information to students, including student gender composition, changes their major choices.

We also find that preferences for the gender ratio differ between STEM and non-STEM programs in ways that matter for talent allocation. Female students in the top half of mathematics ability exhibit stronger minority avoidance in STEM than in non-STEM: relative to a gender-balanced program, being a minority (25% female) reduces their choice probability by 14 percentage points in STEM versus 7 percentage points in non-STEM – a twofold difference. Male students show the opposite pattern: they are less sensitive to gender composition in STEM than in non-STEM. Being a minority reduces male students' choice probability by 8 percentage points in STEM versus 12 percentage points in non-STEM; being a strong majority reduces it by 6 percentage points in STEM versus 8 percentage points in non-STEM. This asymmetry – high-ability female students more deterred by male-dominated STEM, male students less deterred – implies possible talent misallocation.

We then quantify talent misallocation by comparing two scenarios: one where students have preferences over gender composition, and one where these preferences are removed.[3] The asymmetric preferences documented above – stronger minority avoidance among high-ability female students in STEM, weaker sensitivity among male students – generate sizable misallocation. In our exercise, the low female share in STEM results in male students with mathematics ability 0.46 standard deviations below the median to crowd out female students with mathematics ability 0.54 standard deviations above the median.[4]

Related Literature Our paper contributes to several strands of literature. First, we identify a novel source of the STEM gender gap in higher education: the gender composition itself. Prior work has identified numerous contributing factors, which we group into three categories. Individual-level factors include male students' stronger preferences for competition (Buser, Niederle, and Oosterbeek 2014) and female students' comparative advantage in non-STEM (Breda and Napp 2019; Goulas, Griselda, and Megalokonomou 2024). Environmental factors include gender-biased teachers (Carlana 2019; Miserocchi 2024), parental expectations (Carlana and Corno 2024), lack of role models (Breda et al. 2023; Carrell, Page, and West 2010; Riise, Willage, and Willén 2022; Riley 2024), non-cooperative classroom dynamics (Di Tommaso et al. 2024), and insufficient emphasis on STEM's social relevance (Long and Takahashi 2025). Anticipated returns include preferences for job amenities (Burbano, Padilla, and Meier 2024; Zafar 2013; Wiswall and Zafar 2018) and marriage-market considerations (Wiswall and Zafar 2021). Peer effects also play a role (Bechichi and Kenedi 2024; Valdebenito 2023; Mouganie and Wang 2020; Fischer 2017).[5] Our effect size –

3. By misallocation we mean the skewed gender ratio in STEM programs leads to deviation from students' intrinsic interests and underutilization of talent. This usage is in the spirit of Hsieh et al. (2019), who show that women's and black men's occupational choices were distorted in 1960 in the US due to (i) labor market discrimination, (ii) barriers to access to education, and (iii) social norms.

4. Because we are holding students' preferences fixed between the two scenarios, our results are not due to female students' comparative advantage in non-STEM.

5. Bechichi and Kenedi (2024) show that recent graduates' college choices affect high school students' college choices, and that female students respond more strongly to female graduates' choices. Valdebenito (2023) show that female classmates' college choices affect female students' college choices. Mouganie and Wang (2020) show that having high-performing female peers in the same high school class increases female students' science-track choices, whereas

an 11 percentage points increase in female students' choice probability when moving from 25% to 50% female – is economically meaningful: it corresponds to roughly one-third of the baseline STEM choice probability among female students.[6]

Our study is closest to Carlana and Corno (2025), who show that female junior high students are less likely to choose counter-stereotypical tasks when they expect male-dominated peer groups, while male students are unaffected. We extend this work in two ways. First, we elicit preferences over actual college attributes rather than experimental tasks. Second, students' responses in our experiment are private – unobserved by peers, parents, or teachers – thereby ruling out social image concerns as a driver of our results. Our paper also relates to Ersoy and Speer (2025), who show that providing gender composition information changes students' major choices. We go further by recovering preferences over the full range of female shares and identifying the mechanisms underlying these preferences.

Second, we extend the emerging literature on preferences for gender diversity from workplace to educational settings. Existing work shows that workers – especially women – prefer gender-diverse workplaces, and that this preference contributes to occupational segregation (Schuh 2024; Högn et al. 2025; Funk, Iriberri, and Savio 2024). We document analogous preferences among high school students choosing colleges, suggesting these preferences form before labor market entry. This implies that interventions targeting educational choices may be effective in facilitating STEM-related human capital accumulation among women, compared to those targeting workplace that may only address job sorting. Our finding that students perceive gender-balanced programs as easier to fit in aligns with documented benefits of gender diversity,[7] including reduced harassment risk (Folke and Rickne 2022), better recognition of ideas and talent (Koffi 2025; Bello, Casarico, and Nozza 2025), and improved team performance (Hoogendoorn, Oosterbeek, and van Praag 2013).[8]

Finally, our findings inform policies aimed at closing the STEM gender gap. Existing interventions – role model exposure (Breda et al. 2023; Riise, Willage, and Willén 2022; Riley 2024),[9] mentorship (Canaan and Mouganie 2023; Carrell, Page, and West 2010), single-sex instruction (Booth, Cardona-Sosa, and Nolen 2018), and pedagogical changes (Di Tommaso et al. 2024; Long and Takahashi 2025)[10] – implicitly assume that increasing female representation will attract more female students in the long run. Our results validate this assumption: the gender composition itself affects entry decisions. This provides an efficiency-based rationale for such policies, as talent misallocation imposes costs regardless of one's views on equity.[11]

having high-performing male peers decreases female students' science-track choices. Fischer (2017) finds similar results as Mouganie and Wang, but in a college setting and regardless of the gender of the peers.

6. The baseline STEM choice probability among female students in our sample is approximately 34.8%.

7. On the negative side, Shan (2024) finds that students adopt more traditional gender roles when assigned to a mixed-gender study group in college.

8. In the context of family socioeconomic status (SES), Tadjfar and Vira (2025) find that students from low-SES backgrounds are reluctant to apply for elite colleges dominated by students from high-SES backgrounds due to concerns about fitting in.

9. Porter and Serra (2020) show the effectiveness of female role models in inspiring female college students to pursue an economics major.

10. Avery et al. (2024) and Owen and Hagstrom (2021) also find similar results in economics curricula.

11. While not in education, Baltrunaite et al. (2014) and Besley et al. (2017) show that a gender quota for political

The remainder of the paper is structured as follows. Section 2 explains the institutional background. Section 3 details the experimental design. Section 4 describes the summary statistics of the experimental data. Section 5 presents the main results. Section 6 quantifies the degree of talent misallocation. Section 7 concludes.

2 Institutional Background

High School In Japan, high school runs from grades 10 to 12, typically from age 15 to 18. Though not compulsory, nearly 99% of junior high school graduates attend high school (Ministry of Education, Culture, Sports, Science and Technology 2021). College enrollment rates are also high, at nearly 60%.[12] However, colleges vary widely in academic orientation: while some are research-focused, many function more like vocational institutions. We refer to the former as selective colleges and the latter as non-selective colleges. Because STEM fields require strong mathematical preparation, research-oriented STEM programs are concentrated at selective colleges. As such, academic high schools are the primary pathway for students aiming to enter selective colleges or research-oriented STEM programs.

Students at academic high schools choose a track at the end of grade 10, which determines the subjects they study in grades 11 and 12. There are two tracks: humanities and sciences. In the humanities track, students study advanced reading (advanced Japanese), English, history/social studies, and mathematics. In the sciences track, students study reading (Japanese), English, sciences (biology, chemistry, and/or physics), and advanced mathematics. A similar track system exists in many European countries.[13]

Students in these academic high schools regularly take mock exams to prepare for college entrance exams. On the score reports for these mock exams, students can see the likelihood of admission to their preferred programs, calculated based on their exam scores and the program's selectivity, which they use to make and adjust their study plan. To help students select programs, nearly all academic high schools maintain counselor offices with information about each program, including selectivity, gender ratio, and other attributes used in this paper's experiment.

College Application As in Europe but unlike in the US (Bordon and Fu 2015), students apply to specific college programs and cannot change their majors later. Programs differ in their attributes, such as major, selectivity, tuition, whether they are public or private, and location, among others. Since living alone can be costly and some parents prefer their children not to live alone, many college students live with their parents. Among various locations in Japan, the greater Tokyo area

candidates improved elected politicians' competence. We expect similar effects for women in STEM in higher education.

12. As of 2023: https://www.ipss.go.jp/syoushika/tohkei/Data/Popular2024/T11-03.htm (accessed June 9, 2025).

13. In Italy, students choose their track between humanities and sciences at the end of grade 9 (see, for example, Carlana and Corno 2025). In France, students choose their track between humanities, social sciences, and sciences at the end of grade 10 (see, for example, Breda et al. 2023). In the Netherlands, students in academic secondary schools (VWO) choose their track between science, health, social sciences, and humanities at the end of grade 9 (see, for example, Buser, Niederle, and Oosterbeek 2014).

offers the greatest variety of college attributes: nearly 29% of all colleges are located there, where about 41% of all college students study.[14]

Most selective college programs employ an exam-based, meritocratic admissions system. They rank applicants by exam score and make offers from the top of the list. However, each program requires exams in different subjects. Humanities and social sciences programs usually require exams in advanced reading, English, history/social studies, and mathematics. On the other hand, science, engineering, and medicine programs typically require exams in advanced mathematics, English, sciences, and reading. Since students who choose the science track in high school do not study advanced reading and history/social studies, they are effectively constrained from applying to most humanities and social sciences programs. Similarly, students who choose the humanities track in high school do not study advanced mathematics and sciences, making it difficult to apply to science, engineering, and medicine programs. In this way, high school track choice largely determines the set of college majors students can realistically pursue.

Each college program is assigned a single number called the "selectivity index" by commercial college entrance exam preparation companies. The index represents the difficulty of admission and the selectivity (or prestige) of the program and is calculated based on admitted students' mock exam performance. The index is expressed as a z-score rescaled to have a mean of 50 and a standard deviation of 10. The selectivity of the program from which a student graduates significantly affects the quality of the first job they obtain (Nakajima 2018) as well as promotions in the first few years after starting that job (Araki, Kawaguchi, and Onozuka 2016). Given the rigid labor market and limited job mobility in Japan (Moriguchi 2014), the quality of the first job affects students' career prospects more than in other OECD countries (Genda, Kondo, and Ohta 2010).[15]

Talent Misallocation Before describing our experiment, we document the extent of the talent misallocation and the persistence of gender ratios that our experiment aims to explain.

Figure 1 plots the share of female students in each of the 10 deciles of the PISA mathematics (dark dashed), science (dark dotted), and reading (dark dash-dotted) score distributions among 15-year-old students in 2018 who plan to attend a four-year college, as well as the share of female students in STEM (orange) and non-STEM (green) college programs in 2023. The figure shows a significant degree of talent misallocation, as in other OECD countries: although there are no significant gender differences in mathematics and science scores at age 15, significantly fewer female students choose STEM programs in college. While female students excel in reading, it is not large enough to attribute the STEM gender gap to female students' comparative advantage in reading.[16]

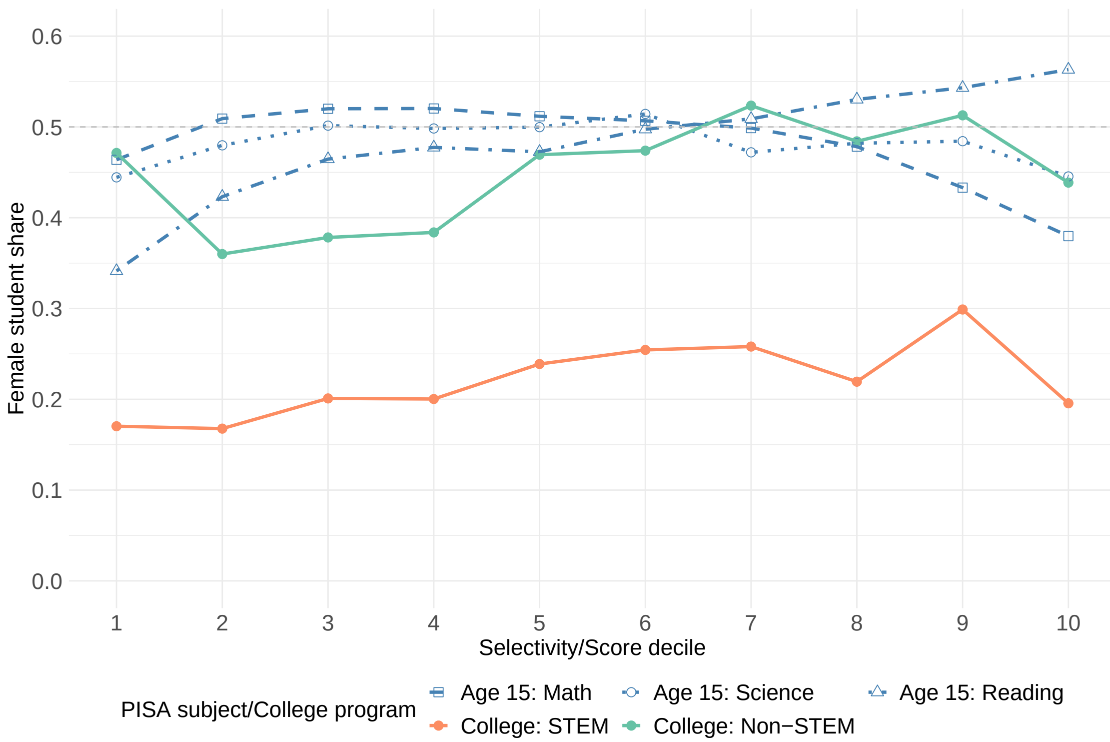
Despite such inefficiency, however, gender ratios across college programs are surprisingly persistent. Figure 2 plots binned averages of the female share among newly enrolled students (Panel A) and the

14. From an article by an educational materials and news publishing company Obunsha: https://eic.obunsha.co.jp/file/educational_info/2024/1022.pdf (accessed July 31, 2025).

15. Genda, Kondo, and Ohta (2010) find that those who entered the labor market during a recession suffer worse employment conditions in terms of unemployment and earnings in Japan than in the US.

16. While the differences between female and male students are less pronounced in Japan than in the OECD average, the general patterns are the same; see Appendix Figure A1.

Figure 1: Talent Misallocation



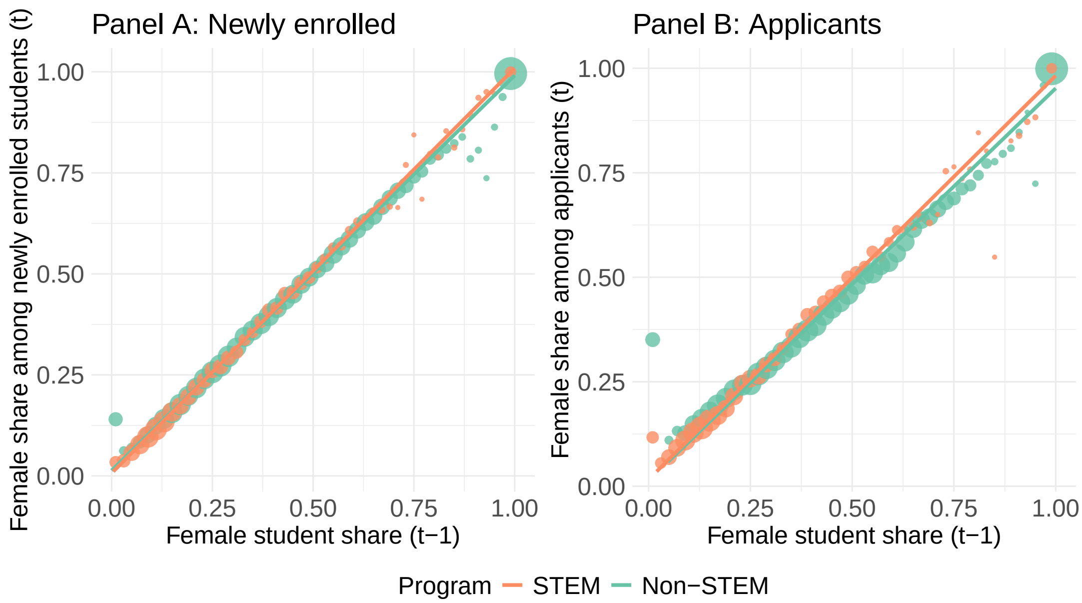
Notes: This figure plots the share of female students in each of the 10 deciles of the PISA mathematics (dark dashed), science (dark dotted), and reading (dark dash-dotted) score distributions among 15-year-old students in 2018 who plan to attend a four-year college, as well as the share of female students in STEM (orange) and non-STEM (green) college programs in 2023. *Sources:* OECD (2018), School Basic Survey.

female share among applicants (Panel B) against the total female student share in the same program in the previous year. Panel A uses the universe of STEM (orange, N=594) and non-STEM (green, N=1409) college programs in Japan observed over 21 years (2003–2023). Panel B uses applicant data available from 2015 onward (9 years, N=507 for STEM, N=1212 for non-STEM).

The figure shows that programs attract about the same share of female applicants and new students as the previous year's female student share, making the female student share highly persistent over time. This pattern holds for both STEM and non-STEM programs. However, there are two competing explanations for this persistence: taste heterogeneity and state dependence (see, for example, Heckman 1981; Hyslop 1999). Taste heterogeneity refers to the possibility that female students are inherently less inclined to pursue STEM or male students are inherently less inclined to pursue non-STEM, resulting in persistent gender ratios regardless of the current composition. State dependence refers to the possibility that female students avoid programs where they expect to be a minority, meaning the current gender composition causally affects future compositionWhile observationally equivalent, these explanations have starkly different policy implications. Under taste heterogeneity, interventions to shift the female share will not break persistence. Under state dependence, they will.

Our experiment is designed to distinguish between these two competing explanations. Our

Figure 2: Persistence of the Gender Ratio Across College Programs



Notes: This figure plots binned averages of the female share among newly enrolled students (Panel A) and the female share among applicants (Panel B) against the total female student share in the same program in the previous year. Panel A uses the universe of STEM (orange, N=594) and non-STEM (green, N=1409) college programs in Japan observed over 21 years (2003-2023). Panel B uses applicant data available from 2015 onward (9 years, N=507 for STEM, N=1212 for non-STEM). *Source:* School Basic Survey.

experiment focuses on students in the top deciles shown in Figure 1.

3 Experimental Design

To investigate whether the gender ratios at colleges affect high school students' college choices, we conducted an incentivized discrete choice experiment at four selective academic high schools in the greater Tokyo area in Japan. We integrated this experiment as a "career planning module" within the 10th-grade curriculum of the participating high schools from December 2023 to July 2024. The experiment was conducted in person at three high schools and asynchronously online at one high school.

The experiment lasted about 40 minutes on average, including distributing the participation gifts. A total of 628 students took part, with 619 providing valid responses (311 female, 298 male, 10 non-binary). Since this study focuses on binary gender, we excluded responses from non-binary students, resulting in 609 responses with 15 choices each, for a total of 9135 choices.

3.1 Sample Selection

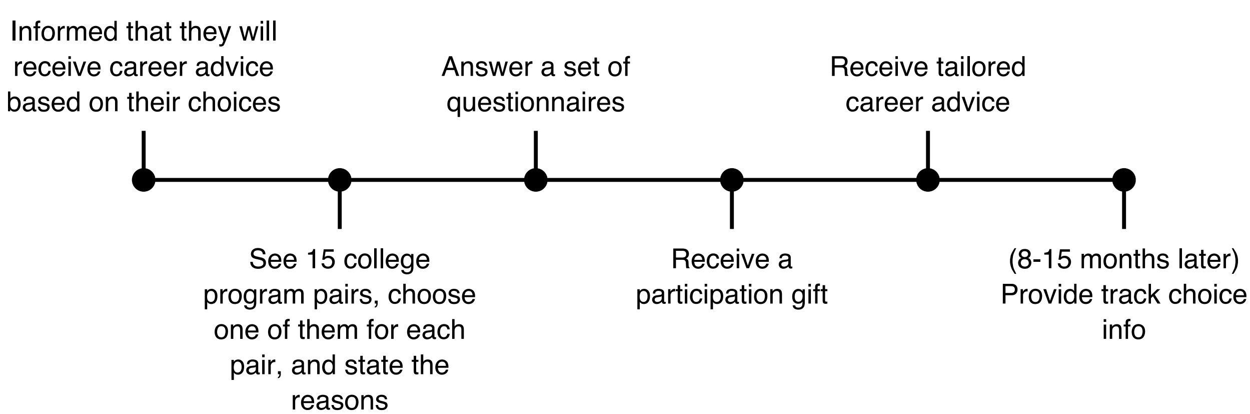
Schools We contacted teachers at academic high schools in the greater Tokyo area through our network and obtained their consent to conduct the experiment as part of their school curriculum. We restricted our potential sample to academic high schools in the greater Tokyo area for three reasons.

First, we wanted to ensure the experimental content was relevant to students: the experiment focused on college choices, and students needed to be planning to attend college. Based on the schools' placement records, over 95% of recent graduates attended college. Second, we wanted to include students whose mathematics skills did not constrain their major choices: as discussed in Section 2, most STEM programs require good mathematics abilities, and students from academic high schools are the primary group prepared for STEM. Third, we wanted to prevent students from implicitly considering potential location and financial constraints when making their choices: as discussed in Section 2, the greater Tokyo area offers the widest variety of college programs, making such constraints less relevant there. We also wanted the attributes of the hypothetical college programs to appear natural to students, and the greater Tokyo area was suitable for this purpose as well.

Students The teachers at the participating high schools distributed the information letter and consent form to guardians of all 10th-grade students, except at one school where only one class participated. We restricted the sample to 10th-grade students before their track choice because the track choice restricted the college programs they could apply for, as discussed in Section 2. The information letter did not mention that the experiment was about gender ratios or STEM to minimize the experimenter demand effect (Zizzo 2010). Instead, we explained that students would evaluate 15 hypothetical college programs, answer a short questionnaire, and receive a tailored career advice sheet based on their evaluations. We also clarified that the data from the experiment would be used for academic research to improve education policy. Nearly 90% of the guardians and students provided consent and participated.

3.2 Flow of the Experiment

Figure 3: Flow of the Experiment



Notes: This figure shows the flow of the experiment.

Students were first told that they would receive a career advice sheet based on their choices in the experiment. We promised them that their responses and the career advice sheet would not be

Figure 4: Hypothetical Program Pair

Pair 4/15

AB College Dept. of Literature		AX College Dept. of Engineering
	Dept. Characteristics	
57.5	Selectivity index	62.5
700	Cohort size	600
35% male, 65% female	Student gender ratio	55% male, 45% female
	College Characteristics	
Yes	Intl exchange program	No
65%	Club participation rate	45%

Which program would you like to attend?

AB College, Dept. of Literature

AX College, Dept. of Engineering

Which program do you feel these statements apply to more?

	AB College, Dept. of Literature	AX College, Dept. of Engineering
I can do well in my studies		
I can find a career I want to pursue		
I can fit in		
I can meet inspiring seniors		

Notes: This figure shows an example of a hypothetical program pair students would see during the experiment.

shared with anyone, including guardians, teachers, or peers, to minimize their potential influence, as previous studies suggest they can affect students' study choices (Carlana 2019; Carlana and Corno 2024; Giustinelli 2016; Miserocchi 2024; Müller 2024; Valdebenito 2023). Students then saw 15 program pairs one by one and chose the one they would prefer to attend; see Figure 4 for an example of a hypothetical program pair. Additionally, they saw four statements that many students consider important when choosing college programs, selected from answers to the open-ended questions in a pre-test. For each pair, students indicated which program each statement applied to better.

After making choices for the 15 program pairs, students completed a questionnaire about their demographics, academic abilities, behavioral traits, and beliefs about the gender ratios in average college programs in Japan. Appendix Section B presents the questionnaire. Afterward, students received a participation gift (a set of decorative, functional pens) equivalent to 500 JPY (approx. 5.27 USD in 2022 PPP) for their participation.[17]

Several weeks later, each student received a career advice sheet we created. The sheet had two parts: a tailored part and a non-tailored part. The tailored part included the top three attributes that students cared about most, along with the top reason they prioritized when choosing a college program, based on their choices in the hypothetical programs. The non-tailored part contained non-individualized information useful for most high school students, such as tips for choosing college programs, college admissions, financing, studying abroad, attending graduate school, and finding jobs.

Finally, 8-15 months after the experiment, we requested students' track choice data from schools. Some schools provided data directly from their records; others provided it through a follow-up survey completed by students.[18] In total, we obtained valid responses from 461 out of 609 students (or 75.7% of students).

3.3 Attributes

We randomly assigned attributes to each program, including college name, department, department selectivity index, department cohort size, department student gender ratio, whether the college has an international exchange program, and college club activity participation rate. Our main interests are (i) student gender ratio and (ii) department, which indicates whether the program is STEM or non-STEM. We included other attributes to make the programs appear more realistic to students and selected attribute value ranges that are plausible for students in our sample to reduce hypothetical bias (List and Shogren 1998; List, Sinha, and Taylor 2006). We asked them to assume that attributes not shown were identical between the programs.

College names consist of two uppercase letters and we draw them without replacement for each program in a pair from a list ranging from AA to BD to make sure they were unrelated to the actual college names. The department was drawn from a list of 12 popular departments

17. USD to JPY PPP was 94.93 in 2022: https://www.oecd.org/en/data/indicators/purchasing-power-parities-ppp.html (accessed November 18, 2024).

18. It is 15 months because one school had their track choice at the end of the 11th grade.

among college students, of which 6 were STEM and 6 were non-STEM.[19] First, we randomly assigned either STEM or non-STEM to one program in the pair. If STEM was selected, then the other program was assigned non-STEM with 75% probability and STEM with 25% probability to reduce the chances that both programs are in the same category. Next, we selected a specific department within the relevant list. STEM departments included Physics, Chemistry, Biology, Engineering, Information Technology, and Agriculture. Non-STEM departments included Literature, Law, Business, Economics, Sociology, and Foreign Language. We excluded Medicine and Education, as both are popular but lead to specialized careers such as doctors, nurses, pharmacists, and teachers, which differ from most student career paths, and the attributes used may not be very relevant for these programs.

Other attributes include selectivity index, which ranges from 55 to 72.5 with an increment of 2.5; cohort size, which varies from 200 to 900 with an increment of 50; student gender ratio, which spans from 5% to 95% for females and males but sums to 100%; whether the college has an international exchange program, which has an 80% chance of being "Yes" and a 20% chance of being "No"; and club participation rate, which ranges from 40% to 85% with an increment of 5%.[20] These attribute values were drawn with replacement for each program. Appendix Table A1 shows the possible values for each attribute.

We used a hypothetical choice experiment because it allowed us to elicit students' preferences over attributes that were varied independently. This design was crucial for three reasons: most STEM programs are male-dominated, leaving little variation in gender ratios; the gender ratio correlates with other program attributes; and using actual college names would lead students to infer attributes not shown.

3.4 Incentives

One concern with hypothetical choice experiments is that students may lack incentives to state their true preferences without real consequences. Although Hainmueller, Hangartner, and Yamamoto (2015) show that choices in hypothetical vignettes and actual behaviors are highly correlated, we address this concern using the incentivized resume rating method (Kessler, Low, and Sullivan 2019), which involves providing career advice based on students' choices.[21] Assuming students believed

19. We first referred to the Ministry of Education, Culture, Sports, Science and Technology (MEXT) department-classification list (https://www.mext.go.jp/component/b_menu/other/__icsFiles/afieldfile/2018/08/02/1407357_4.pdf, accessed June 30, 2023) to compile a comprehensive set of programs. We then consulted a college-choice guide for high school students (https://shingakunet.com/journal/column/20210415000002/, February 21, 2023 version) and selected 12 popular programs.

20. We set the probability that a program has an international exchange program to 80% because most colleges in Japan have one.

21. Several studies employed the incentivized resume rating method to elicit preferences for attributes that are hard to obtain from revealed preferences. Low (2024) elicited heterosexual adults' preferences for dating partners by providing dating advice from a dating coach based on their ratings of hypothetical opposite-gender partner profiles. Macchi (2023) elicited loan officers' preferences for borrowers by offering referrals to loan clients based on their ratings of hypothetical borrower profiles. Gallen and Wasserman (2023) elicited college students' mentor preferences by providing mentor characteristics that students care most about. Chan (2024) elicited patients' preferences for doctors by offering booking options based on patients' choices.

that we researchers had potentially valuable new information, the method is incentive compatible: the expected value of the advice increases with the truthfulness of their choices. Because students were from academic high schools and planned to attend selective colleges, we assume most believed the researchers had valuable academic and career information.[22]

Specifically, we provided the following information in the information letter and at the beginning of the experimental instructions. This wording closely followed the original incentivized resume rating studies (Kessler, Low, and Sullivan 2019; Low 2024):

> *Through this module we will give you information relevant for your career choice. You will complete it on the internet using a laptop or a tablet. It is expected to last for 25 minutes and consists of two parts:*
>
> - *Evaluation of the hypothetical 15 program pairs*
> - *A short questionnaire*
>
> *We will send you a career advice sheet created based on your evaluation.*

As an additional check, in Section 4 we validate that these incentives elicited truthful preferences by comparing predictions to actual track choices.

4 Data

4.1 Variable Construction

Academic Abilities We convert students' academic abilities, obtained through a post-experimental questionnaire, into population z-scores to make them comparable across schools and interpretable within the entire student pool. To do this, we use the latest placement records of graduates from each participating high school, assign the selectivity index to each college in the records, and rank the placements by the selectivity index. We then assign each student a selectivity index based on their academic rank within their high school in specific subjects (reading, mathematics, English, and total). The selectivity index for each program is obtained from the list prepared by Kawaijuku in 2024, one of the most popular commercial college exam preparation companies in Japan.[23] Since public colleges require a larger number of subjects in the entrance exam, we add 0.5 to the index of public colleges (equivalent to 5 points in the raw index), following Araki, Kawaguchi, and Onozuka (2016).

Behavioral Traits We elicit students' behavioral traits through a post-experimental questionnaire: confidence in reading, mathematics, and English; competitiveness; and risk-taking, all rated on a 5-point Likert scale with 3 being neutral.[24] For confidence questions, we ask how accurately their

22. At the time of the experiment, Inoue obtained a PhD from one of Japan's most prestigious colleges, Saito earned a master's degree in the US and has rich industry experience, and Takahashi earned a PhD from a European university.

23. https://www.keinet.ne.jp/exam/ranking/index.html (accessed December 18, 2024)

24. The questionnaire-based competitiveness measure is adapted from Buser, Niederle, and Oosterbeek (2024), and the risk-taking measure is adapted from Dohmen et al. (2011).

recent exam scores, entered on the previous page, reflected their ability. We convert these 5-point scales to the range [-1, 1] for better interpretability, with 0 being neutral.

4.2 Summary Statistics

Table 1: Summary Statistics of Students in the Final Sample

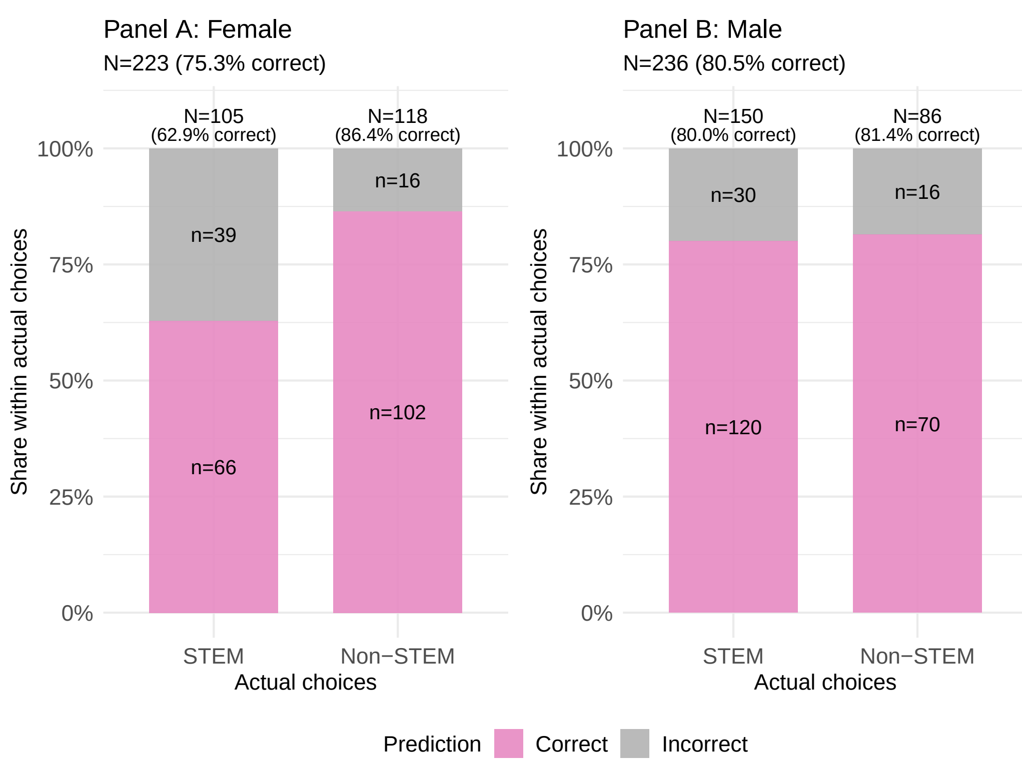
	Female (N=311)	Male (N=298)	Difference (M – F)
Panel A: Demographics			
Mother bachelor or above	0.59	0.56	-0.03
	(0.49)	(0.50)	(0.04)
Father bachelor or above	0.80	0.77	-0.03
	(0.40)	(0.42)	(0.04)
Both bachelor or above	0.50	0.48	-0.02
	(0.50)	(0.50)	(0.04)
Extra schooling (no. days/week)	0.94	0.92	-0.02
	(1.06)	(1.04)	(0.09)
Panel B: Academic abilities (population z-score)			
Reading score	0.82	0.59	-0.23***
	(0.74)	(0.75)	(0.06)
Math score	0.65	0.77	0.12**
	(0.75)	(0.75)	(0.06)
English score	0.82	0.58	-0.25***
	(0.69)	(0.78)	(0.06)
Total score	0.71	0.61	-0.11*
	(0.74)	(0.76)	(0.06)
Panel C: Behavioral traits			
Reading confidence [-1,1]	-0.01	0.05	0.06
	(0.39)	(0.49)	(0.04)
Math confidence [-1,1]	-0.02	0.18	0.20***
	(0.39)	(0.50)	(0.04)
English confidence [-1,1]	-0.00	0.10	0.10***
	(0.40)	(0.45)	(0.03)
Competitiveness [-1,1]	0.03	0.11	0.08
	(0.67)	(0.68)	(0.05)
Risk-taking [-1,1]	-0.36	-0.18	0.18***
	(0.61)	(0.70)	(0.05)

Notes: This table presents summary statistics of 311 female and 298 male students in the final sample (609 students in total) as well as their differences. Standard deviations and standard errors are in parentheses. Significance levels: * 10%, ** 5%, and *** 1%.

Table 1 presents summary statistics for 311 female and 298 male students in the final sample (totaling 609 students), along with their differences. Panel A presents students' demographics and indicates no differences in parents' education levels or parental investments (proxied by extra schooling days per week) between female and male students. Panel B presents students' academic abilities and shows that female students outperform male students in reading and English, while male students excel in mathematics; female students have a slight edge in overall scores.[25] As expected given our sample of academic high schools, students perform approximately 0.6 to 0.7 standard deviations above the national average. Panel C presents students' behavioral traits, showing that male students are more confident in their mathematics and English abilities – despite lower English performance – and are less risk-averse than female students, consistent with existing literature on gender differences in preferences (Croson and Gneezy 2009). Although statistically insignificant, the gender difference in competitiveness is also consistent with existing literature, with male students more competitive than female students.

4.3 Consistency with the Track Choice Data

Figure 5: Predictions vs. Actual Track Choices



Notes: This figure compares the predicted STEM choices obtained from the mixed logit estimates in Section 6 and the actual track choices obtained from the schools. We denote sciences track as STEM and humanities track as non-STEM. Overall, the experimental data correctly predict 78.0% of the actual choices.

Figure 5 compares the predicted STEM choices from the data and the actual track choices obtained from the schools.[26] We denote sciences track as STEM and humanities track as non-STEM.

25. Appendix Figure A2 shows the distribution of abilities for female and male students.

26. The predicted STEM choice is defined as $\mathbb{1}[\hat{\beta}_i^{STEM} > 0]$ where $\hat{\beta}_i^{STEM}$ is mixed logit estimate for STEM dummy of student i in Section 6.

First, for students whose track choice data is available, the experimental data correctly predict 78.0% of the actual choices. Second, the data predict male students' choices more accurately (80.5%) than female students (75.3%), primarily due to predictions for STEM choices. Because the predictions do not take into account the constraints students face when they make choices, such as parents' and teachers' suggestions and peers' influences, they do not need to perfectly coincide with the actual choices. Still, the lower prediction accuracy for female students' STEM choices – compared to male students' STEM choices or either gender's non-STEM choices – is consistent with female students facing stronger obstacles when choosing STEM (Carlana 2019; Carlana and Corno 2024; Giustinelli 2016; Miserocchi 2024).

5 Gender Ratio Affects Students' College Choices

5.1 Econometric Framework

Estimation of Preferences To estimate students' preferences for program attributes, we assume that student i of gender g's preferences over program d with attributes X in pair j are represented by a linear indirect utility function:

$$V_{ijd} = X'_{jd}\beta^g + \varepsilon_{ijd} \tag{1}$$

The probability that student i chooses the right program r over left l in choice pair j is then:

$$P(V_{ijr} > V_{ijl}|X, g) = F((X_{jr} - X_{jl})'\beta^g) \tag{2}$$

where F is the cumulative distribution function (CDF) of $\varepsilon_{ijr} - \varepsilon_{ijl}$. We assume an identity function for the CDF, $F(x) = I(x) = x$, and estimate the model using OLS for ease of interpretation and decomposition. The empirical specification will thus be:

$$C^r_{ij} = \alpha^g + \zeta^g(FShare_{jr} - FShare_{jl}) + (W_{jr} - W_{jl})'\omega^g + \epsilon_{ij} \tag{3}$$

where C^r_{ij} is an indicator variable equal to 1 if student i chooses the right program in choice pair j, $FShare_{jd}$ is the share of female students in program d within pair j, $W_{jd} \equiv X_{jd} \setminus \{FShare_{jd}\}$ is a vector of program d's attributes in pair j excluding the female student share, and α^g is the intercept for the right program.

We also present the results with logit (assume $F(x) = \Lambda(x)$) as a robustness check.

Decomposition of the Choices To investigate the underlying reasons for students' program choices, we treat the four reasons we elicited in the experiment as mediators: fit in, role model, studies, and career. We then decompose the treatment effects of the female student share on the choices into these four reasons, following Gelbach (2016) and Gong, Lu, and Song (2021).

Replace C^r_{ij} in equation 3 with the four reasons:

$$M^m_{ij} = \kappa^{m,g} + \xi_{m,g}(FShare_{jr} - FShare_{jl}) + (W_{jr} - W_{jl})'\psi^{m,g} + \nu^m_{ij} \tag{4}$$

where M^m_{ij} (for $m = 1, 2, 3, 4$) is an indicator variable equal to 1 if student i indicated that reason m better applies to the right program in the choice pair j.

Finally, include all the M^m_{ij}s in equation 3:

$$C^r_{ij} = \alpha^{g,full} + \zeta^{g,full}(FShare_{jr} - FShare_{jl}) + (W_{jr} - W_{jl})'\omega^{g,full} + \sum_{m=1}^{4} \eta^{m,g} M^m_{ij} + \epsilon^{full}_{ij} \tag{5}$$

Gelbach (2016) shows that:

$$\hat{\zeta}^g = \hat{\zeta}^{g,full} + \sum_{m=1}^{4} \hat{\eta}^{m,g}\hat{\xi}^{m,g} \quad \forall g \tag{6}$$

where $\hat{\eta}^{m,g}\hat{\xi}^{m,g}$ is the part of the treatment effects $\hat{\zeta}^g$ explained by reason M^m_{ij}, and $\hat{\zeta}^{g,full}$ is the part of the treatment effects $\hat{\zeta}^g$ unexplained by any of the four reasons.

We discretize the female student share as discussed later, so we perform this decomposition for each bin of the share. We first present the results of the preference estimations (Section 5.2), then apply this decomposition to understand underlying reasons (Section 5.3).

5.2 Preferences for the Gender Ratio

Table 2 presents preferences for program attributes.[27] First, female students are 7.9 percentage points less likely to choose STEM programs, while male students are 3.8 percentage points more likely to do so, consistent with the literature. Second, both female and male students prefer programs with higher selectivity indices, but male students show a slightly stronger preference: a one-standard-deviation increase in selectivity raises female students' choice probability by 4.7 percentage points and male students' by 7.5 percentage points – a 2.8 percentage points gender gap.

Third, students also favor the social aspects of the programs, consistent with the evidence that students value non-academic amenities in college (Ersoy and Speer 2025; Gong et al. 2021; Jacob, McCall, and Stange 2018). Both female and male students prefer programs that have an international exchange program: having an international exchange program increases female students' choice probability by 6.0 percentage points and male students' choice probability by 3.4 percentage points. Club participation rates also affect students' choices: a 10 percentage points increase in club participation rate increases female students' choice probability by 2.0 percentage points and male students' choice probability by 1.3 percentage points. The effect of the cohort size is quantitatively minimal: an increase of a program size by 100 students increases male students' choice probability by 0.6 percentage points but has no statistically significant effect on female

27. Appendix Table A2 presents the same specifications but with indicator variables for the four reasons instead of choice as the outcome variables.

Table 2: Preferences for Program Attributes

Sample:	Female		Male		All	
Outcome:	Choose the program (0/1)					
	(1)	(2)	(3)	(4)	(5)	(6)
STEM	-0.079***	-0.079***	0.038**	0.038**	0.038**	0.038**
	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)
Female student share	0.090***	0.821***	-0.036	0.818***	-0.036	0.818***
	(0.022)	(0.088)	(0.024)	(0.089)	(0.024)	(0.089)
Female student share squared		-0.745***		-0.875***		-0.875***
		(0.088)		(0.090)		(0.090)
Selectivity index (population SD)	0.046***	0.047***	0.075***	0.075***	0.075***	0.075***
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
Cohort size/100	0.003	0.003	0.006**	0.006**	0.006**	0.006**
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Intl exchange program	0.059***	0.060***	0.033**	0.034**	0.033**	0.034**
	(0.014)	(0.014)	(0.013)	(0.013)	(0.013)	(0.013)
Club participation rate	0.193***	0.196***	0.122***	0.130***	0.122***	0.130***
	(0.042)	(0.042)	(0.041)	(0.040)	(0.041)	(0.040)
Female					0.016	0.015
					(0.011)	(0.011)
STEM x Female					-0.117***	-0.116***
					(0.024)	(0.024)
Female student share x Female					0.126***	0.003
					(0.032)	(0.126)
Female student share squared x Female						0.131
						(0.126)
Selectivity index (population SD) x Female					-0.029*	-0.028*
					(0.016)	(0.016)
Cohort size/100 x Female					-0.003	-0.003
					(0.004)	(0.004)
Intl exchange program x Female					0.027	0.026
					(0.019)	(0.019)
Club participation rate x Female					0.071	0.066
					(0.058)	(0.058)
Constant	0.503***	0.503***	0.488***	0.487***	0.488***	0.487***
	(0.007)	(0.007)	(0.008)	(0.008)	(0.008)	(0.008)
Adj. R-squared	0.036	0.054	0.021	0.045	0.029	0.050
No. observations	4649	4649	4451	4451	9100	9100
No. students	310	310	297	297	607	607

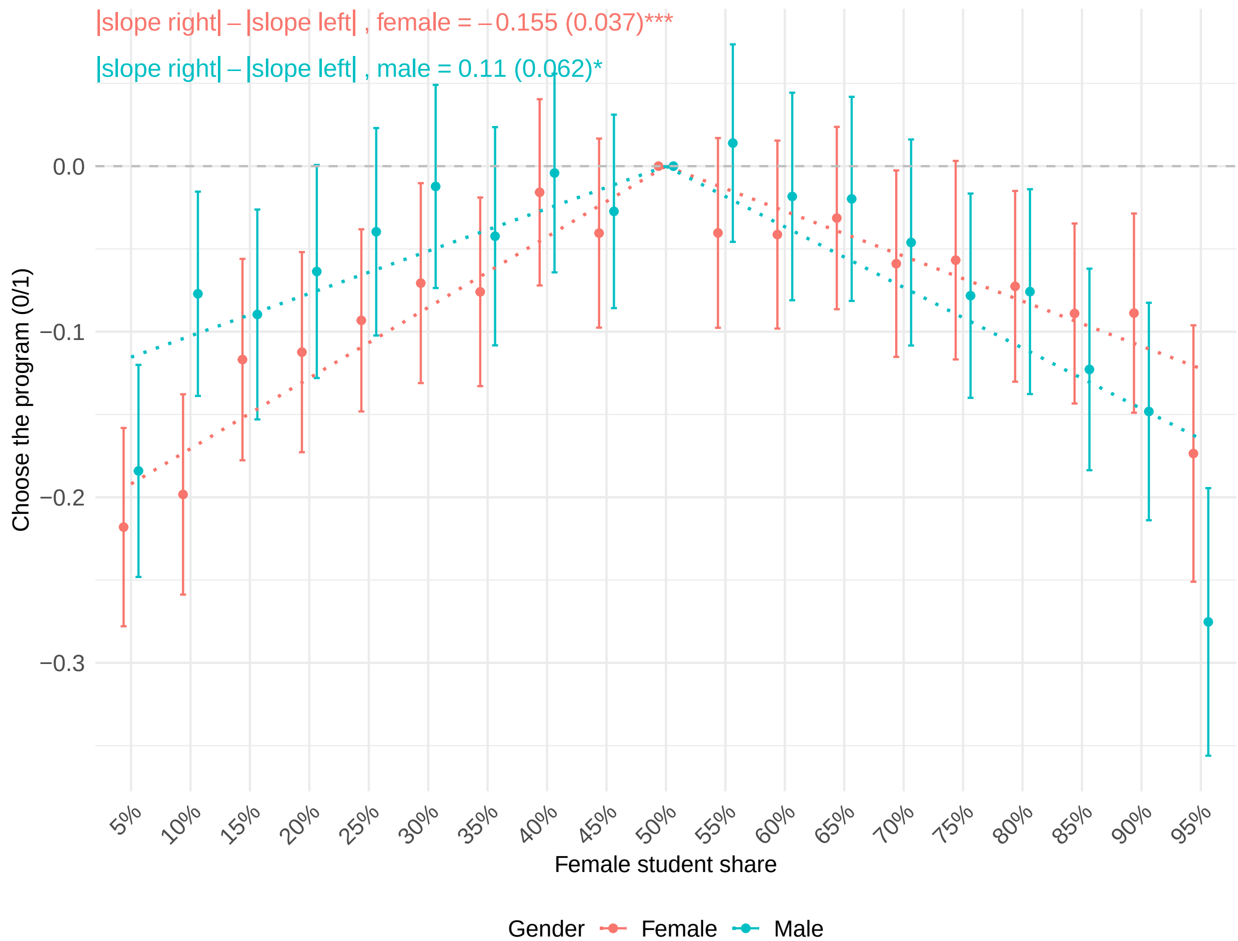
Notes: This table presents the OLS coefficient estimates on the program attributes with choice as the dependent variable. Columns 1 and 2 present estimates for female students, columns 3 and 4 present estimates for male students, and columns 5 and 6 present estimates for differences between female and male students. Standard errors are clustered at the student level. Significance levels: * 10%, ** 5%, and *** 1%.

students. Logit estimates in Appendix Table A3, where we convert the coefficient estimates into average marginal effects show essentially the same results as those with OLS.

The effect of female student share is highly non-linear: the coefficient on the squared term is significant for both genders (columns 2 and 4). We therefore discretize the female share into 19 bins and re-estimate with 50% as the baseline.

Figure 6 reveals three patterns. First, both genders prefer gender-balanced programs over those

Figure 6: Preferences for the Gender Ratio



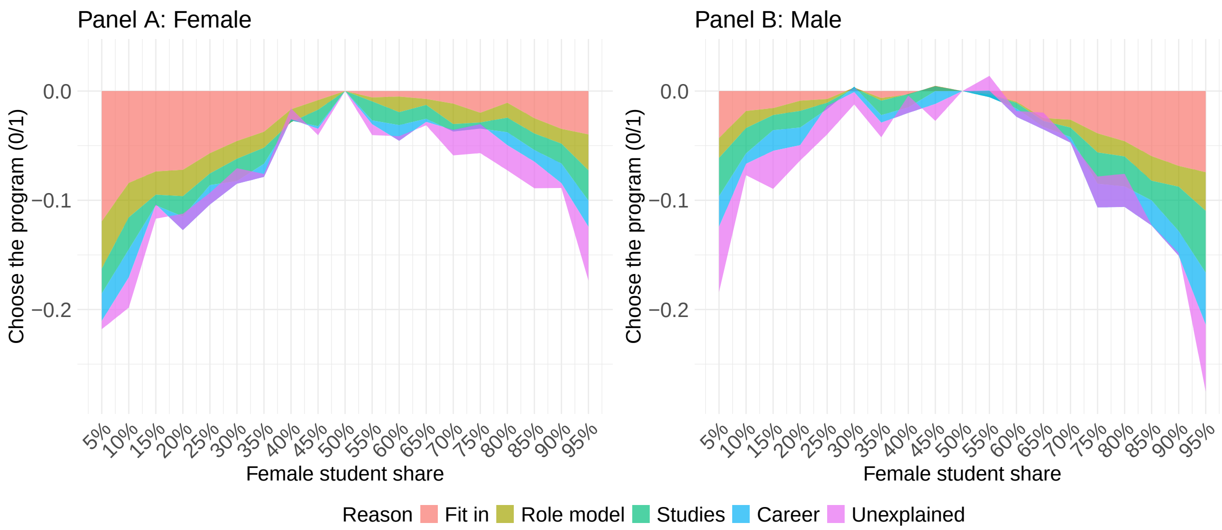
Notes: This figure plots the coefficient estimates and the 95% confidence intervals of female student share discretized into 19 equally spaced bins, with 50% as the baseline, separately for female students (red) and male students (blue). Standard errors are clustered at the student level. The dotted lines are weighted least squares linear fits of each point, with inverse variance weighting, for female and male students on both sides (below 50% and above 50%). We imposed the constraint that the lines pass through the 50% point.

where they would be a minority: relative to 50% female, being a minority (25% own gender) reduces choice probability by 11 percentage points for female students and 9 percentage points for male students. Second, both genders also prefer balance over being a strong majority: being 75% own gender reduces choice probability by 6.5 percentage points for both. Third, despite this preference for balance, students still prefer majority to minority status – the fitted slopes are steeper on the minority side than the majority side, indicating stronger aversion to minority status than to majority status.

5.3 Underlying Reasons

Figure 7 decomposes the effect of gender composition into four stated reasons – fitting in, finding a role model, doing well in studies, and finding a career – plus an unexplained residual.

Figure 7: Decomposition of Preferences for the Gender Ratio



Notes: This figure plots $\hat{\eta}^{m,g}\hat{\xi}^{m,g}$ $(m = 1, 2, 3, 4)$ and $\hat{\zeta}^{g,full}$ from equation 6 estimated for each of the 19 bins separately for female (Panel A) and male (Panel B) students. The red area shows fitting in, the yellow area shows finding a role model, the green area shows doing well in studies, the blue area shows finding a career to pursue, and the purple area shows reasons other than these four.

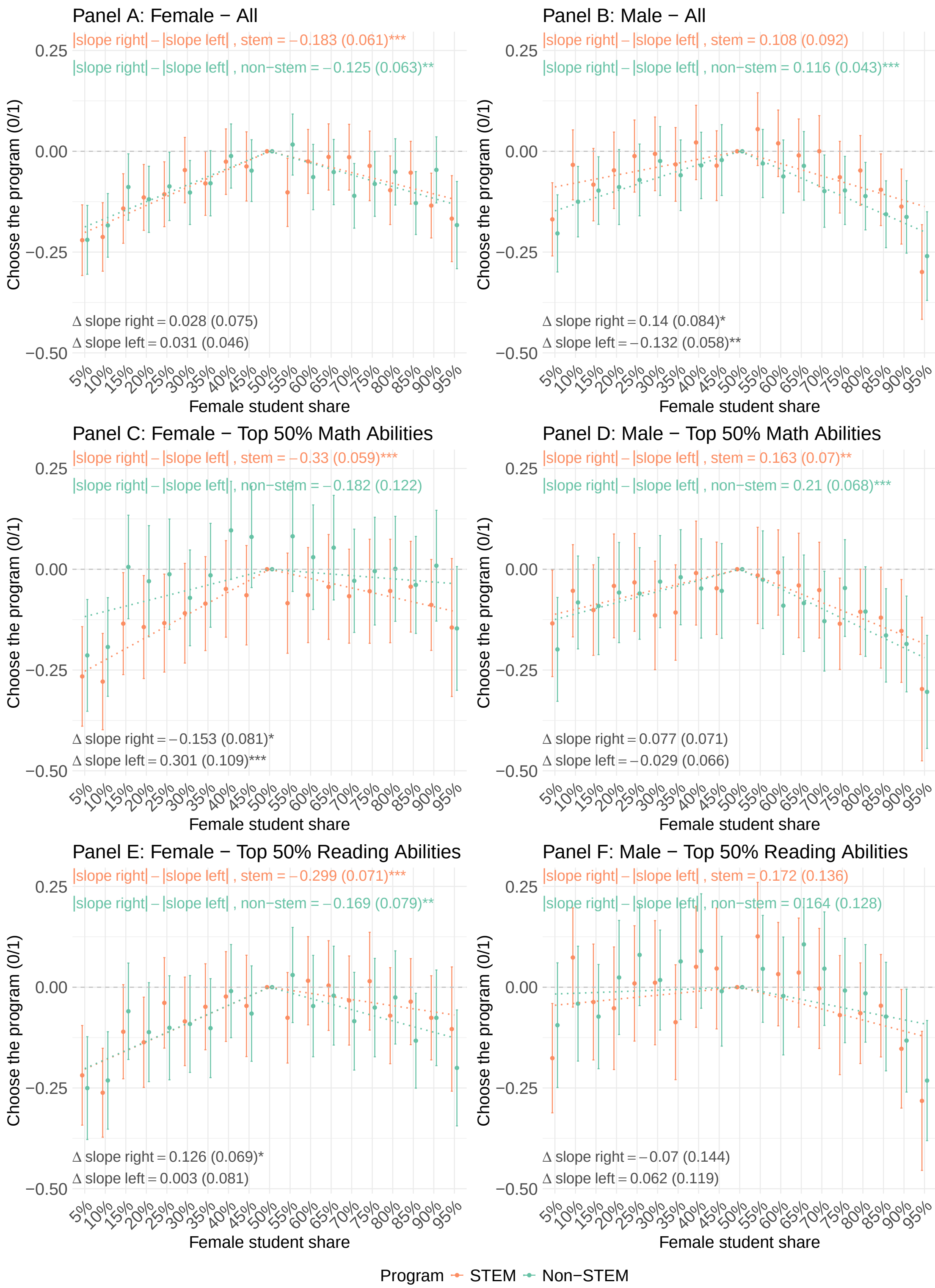
Panel A shows that the main reason the gender ratio affects female students' program choices varies significantly depending on whether most students in the program are male or female. When the majority of students are male, the gender ratio affects female students' program choices primarily due to their concerns about fitting in. On the other hand, when the majority of students are female, multiple factors – not only concerns about fitting in, but also concerns about finding a role model, doing well in studies, finding a career, and unexplained reasons – all contribute to how the gender ratio affects their choices. One notable observation is that the unexplained part is larger when the majority of students are female. It may reflect marriage market concerns, as students often find their future partners in the field of study within their college (Artmann et al. 2021; Kirkebøen et al. 2025; Pestel 2021), especially in fields with a larger fraction of opposite-gender peers (Artmann et al. 2018).

Interestingly, Panel B shows that the patterns are very similar for male students: when most students are female, the gender ratio affects male students' program choices primarily through the expected difficulty in fitting in. When most students are male, however, all four reasons, along with other reasons beyond these four, explain their choices.

5.4 Heterogeneity of Preferences

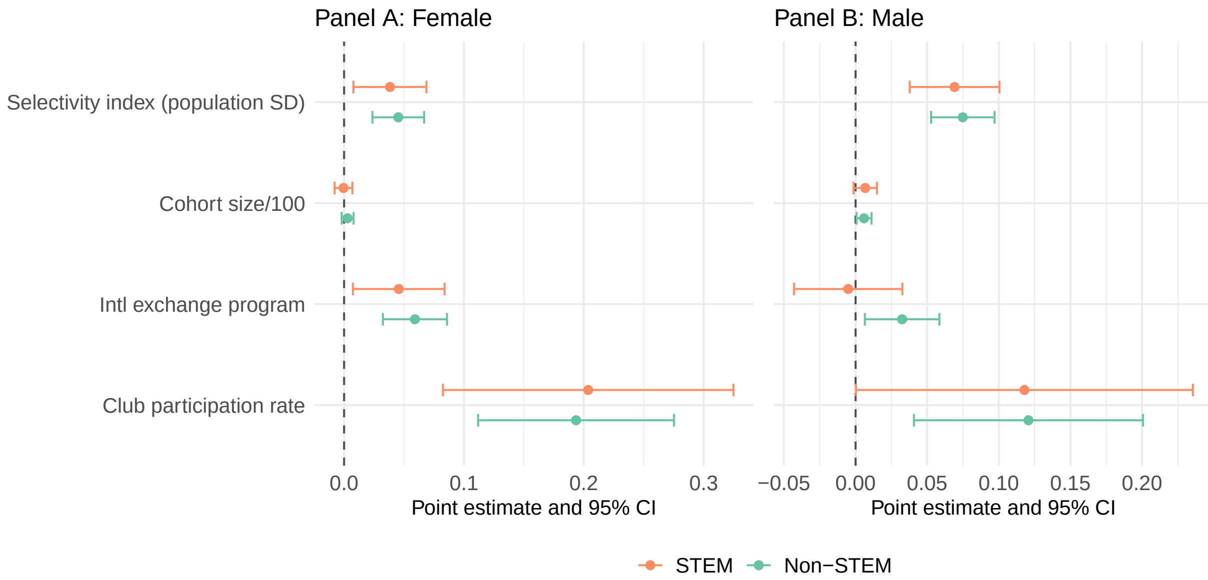
Figure 8 plots coefficient estimates and the 95% confidence intervals for female and male students, where we interact the attributes with the STEM dummy. Panels A and B show all students, Panels C and D restrict the sample to students among the top 50% in mathematics abilities in the sample, and Panels E and F restrict the sample to students among the top 50% in reading abilities in the

Figure 8: Heterogeneity of Preferences for the Gender Ratio – STEM vs. Non-STEM



Notes: This figure plots coefficient estimates and the 95% confidence intervals for female and male students, where we interact the attributes with the STEM dummy. Panels A and B plot all female and male students' preferences, Panels C and D female and male students in the top 50% of mathematics ability in the sample, and Panels E and F female and male students in the top 50% of reading ability in the sample. Standard errors are clustered at the student level. The dotted lines are weighted least squares linear fits of each point, with inverse variance weighting, for each side (below 50% and above 50%) for two groups. We imposed the constraint that the lines pass through the 50% point. The differences in the slopes between STEM and non-STEM are calculated using the weighted least squares standard errors. Significance levels: * 10%, ** 5%, and *** 1%.

Figure 9: Heterogeneity of Preferences for Other Attributes – STEM vs. Non-STEM



Notes: This figure plots coefficient estimates and the 95% confidence intervals of program attributes for female (Panel A) and male students (Panel B) for STEM and non-STEM programs. Female student share is included in the estimation but omitted from this figure for brevity. Standard errors are clustered at the student level.

sample.[28]

Panel A shows that female students' preferences for gender composition are similar in STEM and non-STEM programs. Panel B reveals a different pattern for male students: they are less sensitive to gender composition in STEM than in non-STEM. Being a minority reduces male students' choice probability by 8 percentage points in STEM versus 12 percentage points in non-STEM; being a strong majority reduces it by 6 percentage points in STEM versus 8 percentage points in non-STEM. Panels C and D restrict the sample to students in the top half of mathematics ability. Here, female students exhibit the opposite pattern from males: they show stronger minority avoidance in STEM than non-STEM. Relative to a gender-balanced program, being a minority (25% female) reduces high-math-ability female students' choice probability by 14 percentage points in STEM versus 7 percentage points in non-STEM—a twofold difference. High-math-ability male students show no significant STEM/non-STEM heterogeneity (Panel D).[29]

Panels E and F examine students in the top half of reading ability; neither gender shows significant heterogeneity by field.[30] Preferences for other attributes also do not differ between STEM and non-STEM (Figure 9).[31]

28. There is no significant heterogeneity by other program attributes; see Appendix Figures A3 and A4.
29. Appendix Figure A6 presents differences in slopes between students in the top and bottom 50% of mathematics and reading abilities. These students also differ in their preferences for other attributes, as shown in Appendix Figure A6.
30. ...
31. These ability differences are not coming from differences in parental education level, as shown in Appendix Figure A7.

Thus, both genders prefer balance to majority and majority to minority. But sensitivity to gender composition varies by field: high-ability female students are more deterred by minority status in STEM, while male students are less sensitive to composition in STEM overall. One interpretation is that students perceive STEM as a male domain where male students feel they belong regardless of the gender ratio, whereas female students – especially those accustomed to male-dominated academic environments – recognize that fitting in requires a critical mass of same-gender peers. This asymmetry implies talent misallocation, which we quantify in Section 6.

6 Evidence of Talent Misallocation

The heterogeneous preferences documented above – stronger minority avoidance among high-math-ability females in STEM, weaker sensitivity among males – suggest the low female share may lead to talent misallocation. To quantify this misallocation, we compare two scenarios: one where students have preferences over gender composition, and one where these preferences are removed. Here, the misallocation refers to the fact that the skewed gender ratio in STEM programs leads to deviation from students' intrinsic interests and underutilization of talent, in the spirit of Hsieh et al. (2019).[32]

Because we do not know the constraints students face when they make choices, we hold factors other than preferences for the gender ratio constant across the two scenarios and focus on the differences that come solely from preferences over the gender ratio.

6.1 Econometric Framework

Estimation of Individual-Level Preferences and STEM Choice Probabilities To assess talent misallocation, we estimate each student's choice probability and compare their STEM choice probability across the two scenarios. To do so, we first need to estimate individual students' preferences for program attributes. Rewrite equation 2 as follows:

$$P_{ijd}^{g} = \frac{\exp(X_{jd}'\beta^{g})}{\sum_{k\in\{r,l\}}\exp(X_{jk}'\beta^{g})} \tag{7}$$

where P_{ijd}^{g} is student i of gender g's choice probability of program $d \in \{r, l\}$ with attributes X in pair j. Note that this expression is equivalent to equation 2 in Section 5 (see, for example, Train 2009, Section 3.1).

Now, assume β_i^g is a random variable with density $f(\beta^g|\theta^g)$, where θ^g are parameters of this distribution for gender g. Then the choice probability can be written in a mixed logit form:

$$P_{ijd}^{g}|\theta^{g} = \int \frac{\exp(X_{jd}'\beta_i^{g})}{\sum_{k\in\{r,l\}}\exp(X_{jk}'\beta_i^{g})} f(\beta^{g}|\theta^{g})d\beta^{g} \tag{8}$$

32. Hsieh et al. (2019) show that women's and black men's occupational choices were distorted in 1960 relative to 2010 in the US due to (i) labor market discrimination, (ii) barriers to access to education, and (iii) social norms.

where we allow arbitrary correlations among elements in β_i^g.[33] We assume that the mixing distribution f follows a normal distribution as is standard in the literature (Train 2009). We use a quadratic functional form for the female student share to keep the number of model parameters adequate for the sample size as well as to define the choice probability over the continuous female share. Appendix Table A4 presents the mixed logit parameter estimates. While the scales are different, the sign and the relative magnitude of the coefficients between female and male students are quantitatively the same as in Table 2.

We recover individual-specific parameters β_i^g using the conditional distribution of β_i^g, following Train (2009):

$$h(\beta^g | y, X, g, \theta^g) \tag{9}$$

where y is a vector of choices across the 15 pairs.

We then apply the individual parameter estimates to the two hypothetical college programs to estimate the STEM choice probabilities for each student. The two programs are identical except that one is STEM and the other is non-STEM, and their female student shares are the actual shares of respective programs in Japan in 2024 – 22.7% in STEM and 45.2% in non-STEM.[34] Other attributes are set to the median values that a hypothetical college program in the experiment can take and are the same for both programs.[35]

To examine the substitution patterns among students with different abilities under the two scenarios, we need to apply a capacity constraint to the STEM program. To do so, we subtract the following shadow price from the STEM indirect utility:

$$\text{Shadow price}_i = \lambda(\max_{k \in I}(\text{ability}_k) - \text{ability}_i) \tag{10}$$

where ability_i is student i's academic ability and I is the set of all students in the sample. The shadow price λ increases as more students place a higher value on the STEM program. When λ is positive, it implies that entering a STEM program is more costly for students with lower ability than for students with higher ability. This shadow price can be interpreted as reflecting the effort required to prepare for the entrance exam, the effort needed to catch up with the class after entering, or the risk of not being admitted into the program.

We calibrate the λ so that the average probability of choosing STEM programs across all students in the sample is equal to 31.7%, the actual fraction of students in STEM programs in Japan in 2024.

33. Mixed logit relaxes the following standard logit assumptions: (i) no random preference variation among individuals, (ii) independence of irrelevant alternatives, and (iii) no correlation in unobserved factors over time (Train 2009, Section 6). Our logit results in Table 2 are still valid as there are only two alternatives in each pair and no notable time-varying unobserved factors exist in the experiment. We use mixed logit to relax the first assumption to estimate individual-level preference parameters.

34. From the School Basic Survey: https://www.e-stat.go.jp/stat-search/files?stat_infid=000040230298 (accessed June 5, 2025).

35. 63.75 for selectivity index, 550 for cohort size, 1 for international exchange program, and 62.5% for club participation rate.

6.2 STEM Choice Probabilities Over the Ability Distribution

Figure 10 plots STEM choice probabilities over the ability distribution under two scenarios: actual preferences (solid lines) and counterfactual no-gender-preferences (dashed lines). Panels A and B use mathematics ability in the shadow price; Panels C and D use total score.

Panel A shows that higher mathematics ability predicts higher STEM choice probability for both genders in both scenarios. However, low female share generates misallocation. Relative to the no-gender-preferences scenario, the actual scenario has 4 percentage points fewer female students in STEM, and these female students have mathematics ability 0.54 standard deviations above the median. Conversely, the actual scenario has 2.7 percentage points more male students in STEM, and these male students have mathematics ability 0.46 standard deviations below the median. Both differences are statistically significant at the 1% level (bootstrapped standard errors with 2,000 draws).

Panel B shows similar patterns using reading ability, though the ability-STEM relationship is weaker – likely reflecting the positive correlation between mathematics and reading scores.[36] Panels C and D confirm these patterns using total scores in the shadow price.

In summary, the low female share in STEM leads to talent misallocation: male students with mathematics ability 0.46 standard deviations below the median crowd out female students with mathematics ability 0.54 standard deviations above the median.

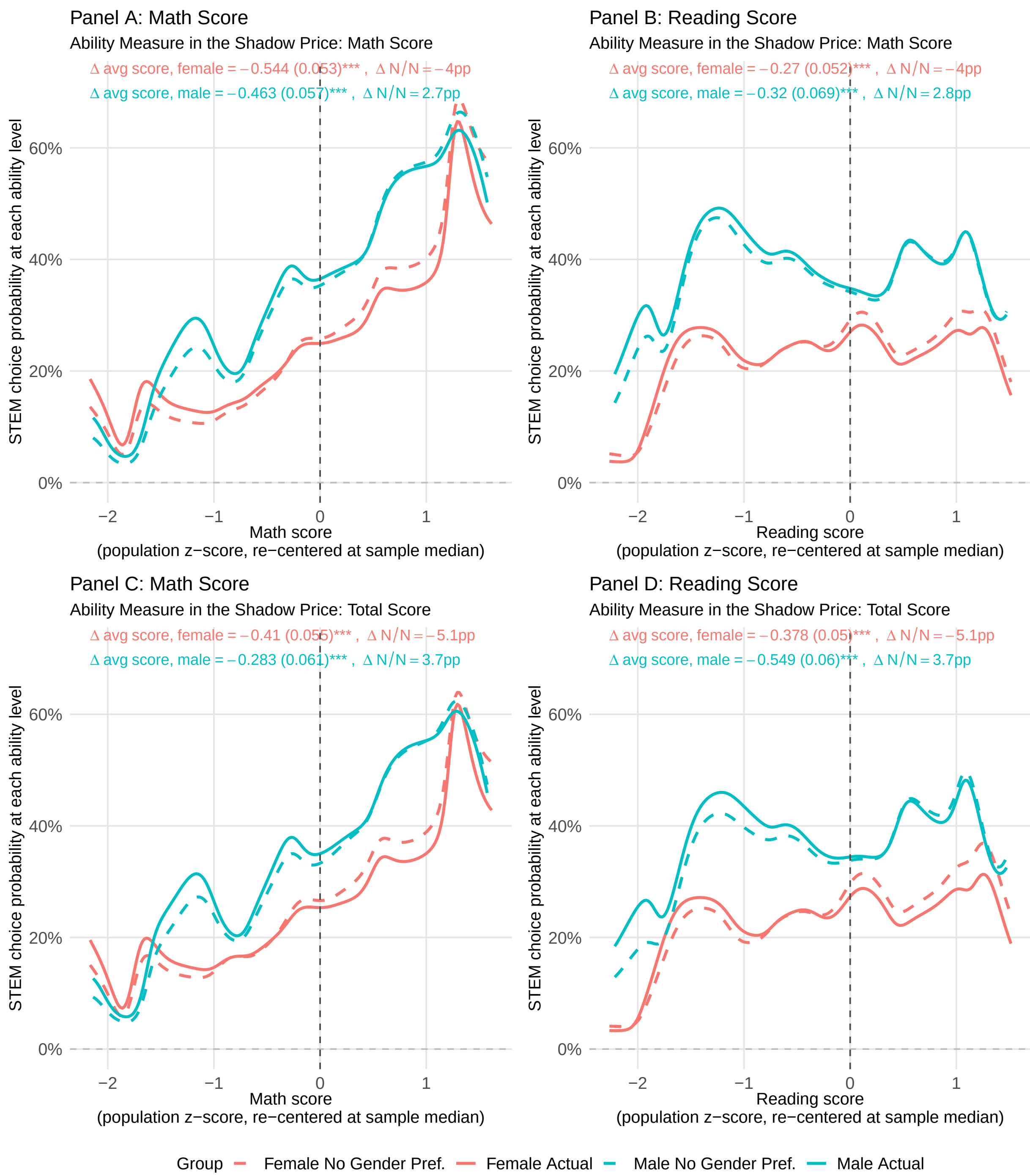
7 Discussion and Conclusion

Female students are significantly less likely to pursue STEM fields at colleges in OECD countries, despite the negligible gender gap in mathematics and sciences at age 15. One potential reason that has received little attention in the literature is that STEM programs are male-dominated, which may discourage female students from pursuing STEM in anticipation of several disadvantages of being a gender minority. In this paper, we examine whether the gender ratio affects students' college choices and whether the low female share in STEM contributes to talent misallocation using an incentivized discrete choice experiment with high school students.

We find that the gender ratio affects the college choices of both female and male students. Both genders prefer programs with a balanced gender ratio over those dominated by either gender, and both prefer being a majority to being a minority. A decomposition reveals that students avoid being a minority primarily because they anticipate difficulty fitting into such environments. Importantly, these preferences vary by context: female students with high mathematics ability are particularly averse to being a minority in STEM programs, while male students are generally less sensitive to the gender ratio in STEM than in non-STEM programs. A quantification exercise shows that these preferences generate talent misallocation – male students with lower mathematics ability crowd out female students with higher mathematics ability in STEM programs.

36. The correlation coefficient is 0.36 overall, with female students exhibiting a stronger correlation (0.48) than male students (0.27).

Figure 10: STEM Choice Probabilities Over the Whole Ability Distribution



Notes: This figure plots STEM choice probabilities over the ability distribution under actual (solid) and no-gender-preferences (dashed) scenarios for female (red) and male (blue) students. Panels A and B use the mathematics score, while Panels C and D use the total score, as the ability measure in the shadow price (equation 10). The changes in the average scores are calculated as an expected marginal change in average student's test score s_i within each gender between the two scenarios: $\sum_i (p_i^{\text{actual}} - p_i^{\text{no-gender-pref.}}) s_i \big/ \sum_i |p_i^{\text{actual}} - p_i^{\text{no-gender-pref.}}|$. Statistical significance is based on bootstrapped standard errors with 2,000 draws. The expected change in the number of students between scenarios is in parentheses. Significance levels: * 10%, ** 5%, and *** 1%.

Our findings carry several implications. First, they suggest that the low female share in STEM is self-perpetuating: the current gender imbalance deters the very students who might otherwise enter, reinforcing the imbalance. This creates a coordination problem where individual students' rational responses to the existing composition collectively sustain an inefficient equilibrium. Second, our results provide an efficiency-based rationale for policies aimed at increasing female representation in STEM. Interventions such as efforts to create more welcoming environments for female students can be justified not only on equity grounds but also because they may improve the allocation of talent.

Third, the finding that anticipated difficulty fitting in drives minority avoidance points toward specific mechanisms that interventions might target: for instance, providing information about peer support or highlighting successful integration of minority students. These initiatives would benefit society beyond efficient talent allocations. They can help address gender differences in occupational sorting and close the gender wage gap, moving the social norms toward more gender-equal ones in the long run. More broadly, more gender-equal societies are beneficial for all members: they tend to be less hostile to same-sex marriage (Baranov, De Haas, and Grosjean 2023) and have a higher share of cross-gender friendships (Bailey et al. 2025).

Several limitations warrant discussion. Our sample consists of high-achieving students at selective high schools in the Tokyo metropolitan area, who face fewer constraints on their college choices than students elsewhere in Japan or in other countries. Whether our findings generalize to students who have weaker mathematics ability or who face tighter geographic constraints remains an open question. Additionally, while our incentivized choice experiment provides clean identification of preferences, it cannot capture all factors that influence actual college decisions, including parental pressure, teacher recommendations, and peer influences. The gap between our experimental predictions and actual track choices, particularly for female students choosing STEM, suggests these external factors may disproportionately constrain female students' choices.

Future research could extend this work in several directions. Examining whether preferences for gender balance vary across countries with different cultural norms around gender would help assess the generalizability of our findings. Investigating whether information interventions – such as highlighting successful female STEM students or providing realistic previews of the social environment – can attenuate minority avoidance would have direct policy relevance. Finally, tracking students longitudinally to examine whether those who enter gender-imbalanced programs experience the difficulties they anticipate would shed light on whether students' beliefs are calibrated accurately.

In conclusion, our findings suggest that the gender composition of STEM programs is not merely a symptom of other factors driving female students away, but is itself a barrier to entry. Breaking this cycle may require coordinated efforts to shift the equilibrium. Small increases in female representation could reduce the cost of entry for subsequent cohorts, potentially generating momentum toward greater balance. More broadly, our results underscore how compositional features of educational environments can shape who pursues which fields, with consequences for both individual careers and the efficient allocation of talent in society.

References

Araki, Shota, Daiji Kawaguchi, and Yuki Onozuka. 2016. "University Prestige, Performance Evaluation, and Promotion: Estimating the Employer Learning Model Using Personnel Datasets." *Labour Economics,* SOLE/EALE Conference Issue 2015, 41:135–148.

Artmann, Elisabeth, Nadine Ketel, Hessel Oosterbeek, and Bas van der Klaauw. 2018. *Field of Study and Family Outcomes.* Working Paper 11658. Institute of Labor Economics (IZA).

———. 2021. "Field of Study and Partner Choice." *Economics of Education Review* 84:102149.

Avery, Mallory, Jane Caldwell, Christian D. Schunn, and Katherine Wolfe. 2024. "Improving Introductory Economics Course Content and Delivery Improves Outcomes for Women." *The Journal of Economic Education* 55 (3): 216–231.

Bailey, Michael, Drew Johnston, Theresa Kuchler, Ayush Kumar, and Johannes Stroebel. 2025. "Cross-Gender Social Ties around the World." *AEA Papers and Proceedings* 115:132–138.

Baltrunaite, Audinga, Piera Bello, Alessandra Casarico, and Paola Profeta. 2014. "Gender Quotas and the Quality of Politicians." *Journal of Public Economics* 118:62–74.

Baranov, Victoria, Ralph De Haas, and Pauline Grosjean. 2023. "Men. Male-biased Sex Ratios and Masculinity Norms: Evidence from Australia's Colonial Past." *Journal of Economic Growth* 28 (3): 339–396.

Bechichi, Nagui, and Gustave Kenedi. 2024. *Older Schoolmate Spillovers on Higher Education Choices.* Working Paper.

Bello, Piera, Alessandra Casarico, and Debora Nozza. 2025. "Research Similarity and Women in Academia." *The Economic Journal,* ueaf113.

Besley, Timothy, Olle Folke, Torsten Persson, and Johanna Rickne. 2017. "Gender Quotas and the Crisis of the Mediocre Man: Theory and Evidence from Sweden." *American Economic Review* 107 (8): 2204–2242.

Booth, Alison L., Lina Cardona-Sosa, and Patrick Nolen. 2018. "Do Single-Sex Classes Affect Academic Achievement? An Experiment in a Coeducational University." *Journal of Public Economics* 168:109–126.

Bordon, Paola, and Chao Fu. 2015. "College-Major Choice to College-Then-Major Choice." *The Review of Economic Studies* 82 (4): 1247–1288.

Bostwick, Valerie K., and Bruce A. Weinberg. 2022. "Nevertheless She Persisted? Gender Peer Effects in Doctoral STEM Programs." *Journal of Labor Economics* 40 (2): 397–436.

Breda, Thomas, Julien Grenet, Marion Monnet, and Clémentine Van Effenterre. 2023. "How Effective Are Female Role Models in Steering Girls Towards STEM? Evidence from French High Schools." *The Economic Journal* 133 (653): 1773–1809.

Breda, Thomas, and Clotilde Napp. 2019. "Girls' Comparative Advantage in Reading Can Largely Explain the Gender Gap in Math-Related Fields." *Proceedings of the National Academy of Sciences* 116 (31): 15435–15440.

Burbano, Vanessa, Nicolas Padilla, and Stephan Meier. 2024. "Gender Differences in Preferences for Meaning at Work." *American Economic Journal: Economic Policy* 16 (3): 61–94.

Buser, Thomas, Muriel Niederle, and Hessel Oosterbeek. 2014. "Gender, Competitiveness, and Career Choices." *The Quarterly Journal of Economics* 129 (3): 1409–1447.

———. 2024. "Can Competitiveness Predict Education and Labor Market Outcomes? Evidence from Incentivized Choice and Survey Measures." *The Review of Economics and Statistics,* 1–45.

Canaan, Serena, and Pierre Mouganie. 2023. "The Impact of Advisor Gender on Female Students' STEM Enrollment and Persistence." *Journal of Human Resources* 58 (2): 593–632.

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

Carlana, Michela, and Lucia Corno. 2024. "Thinking about Parents: Gender and Field of Study." *AEA Papers and Proceedings* 114:254–258.

———. 2025. *Peer Influence in Educational Choices: Social Image Concerns and Same-Gender Interactions.* Working Paper.

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

Chan, Alex. 2024. *Discrimination Against Doctors: A Field Experiment.* Working Paper.

Croson, Rachel, and Uri Gneezy. 2009. "Gender Differences in Preferences." *Journal of Economic Literature* 47 (2): 448–474.

Di Tommaso, Maria Laura, Dalit Contini, Dalila De Rosa, Francesca Ferrara, Daniela Piazzalunga, and Ornella Robutti. 2024. "Tackling the Gender Gap in Mathematics with Active Learning Methodologies." *Economics of Education Review* 100:102538.

Dohmen, Thomas, Armin Falk, David Huffman, Uwe Sunde, Jürgen Schupp, and Gert G. Wagner. 2011. "Individual Risk Attitudes: Measurement, Determinants, and Behavioral Consequences." *Journal of the European Economic Association* 9 (3): 522–550.

Einiö, Elias, Josh Feng, and Xavier Jaravel. 2025. *Social Push and the Direction of Innovation.* Working Paper 3383703.

Ersoy, Fulya, and Jamin D. Speer. 2025. "Opening the Black Box of College Major Choice: Evidence from an Information Intervention." *Journal of Economic Behavior & Organization* 231:106800.

Fischer, Stefanie. 2017. "The Downside of Good Peers: How Classroom Composition Differentially Affects Men's and Women's STEM Persistence." *Labour Economics* 46:211–226.

Folke, Olle, and Johanna Rickne. 2022. "Sexual Harassment and Gender Inequality in the Labor Market." *The Quarterly Journal of Economics* 137 (4): 2163–2212.

Funk, Patricia, Nagore Iriberri, and Giulia Savio. 2024. "Does Scarcity of Female Instructors Create Demand for Diversity among Students? Evidence from an M-Turk Experiment." *Labour Economics* 90:102606.

Gallen, Yana, and Melanie Wasserman. 2023. "Does Information Affect Homophily?" *Journal of Public Economics* 222:104876.

Gelbach, Jonah B. 2016. "When Do Covariates Matter? And Which Ones, and How Much?" *Journal of Labor Economics* 34 (2): 509–543.

Genda, Yuji, Ayako Kondo, and Souichi Ohta. 2010. "Long-Term Effects of a Recession at Labor Market Entry in Japan and the United States." *Journal of Human Resources* 45 (1): 157–196.

Giustinelli, Pamela. 2016. "Group Decision Making with Uncertain Outcomes: Unpacking Child–Parent Choice of the High School Track." *International Economic Review* 57 (2): 573–602.

Gong, Jie, Yi Lu, and Hong Song. 2021. "Gender Peer Effects on Students' Academic and Noncognitive Outcomes: Evidence and Mechanisms." *Journal of Human Resources* 56 (3): 686–710.

Gong, Yifan, Lance Lochner, Ralph Stinebrickner, and Todd R. Stinebrickner. 2021. *The Consumption Value of College.* Working Paper.

Goulas, Sofoklis, Silvia Griselda, and Rigissa Megalokonomou. 2024. "Comparative Advantage and Gender Gap in STEM." *Journal of Human Resources* 59 (6): 1937–1980.

Hainmueller, Jens, Dominik Hangartner, and Teppei Yamamoto. 2015. "Validating Vignette and Conjoint Survey Experiments against Real-World Behavior." *Proceedings of the National Academy of Sciences* 112 (8): 2395–2400.

Hampole, Menaka, Francesca Truffa, and Ashley Wong. 2024. *Peer Effects and the Gender Gap in Corporate Leadership: Evidence from MBA Students.* Working Paper.

Heckman, James J. 1981. "Heterogeneity and State Dependence." In *Studies in Labor Markets,* edited by Sherwin Rosen, 91–140. University of Chicago Press.

Högn, Celina, Lea Mayer, Johannes Rincke, and Erwin Winkler. 2025. *Preferences for Gender Diversity in High-Profile Jobs.* Working Paper.

Hoogendoorn, Sander, Hessel Oosterbeek, and Mirjam van Praag. 2013. "The Impact of Gender Diversity on the Performance of Business Teams: Evidence from a Field Experiment." *Management Science* 59 (7): 1514–1528.

Hsieh, Chang-Tai, Erik Hurst, Charles I. Jones, and Peter J. Klenow. 2019. "The Allocation of Talent and U.S. Economic Growth." *Econometrica* 87 (5): 1439–1474.

Hyslop, Dean R. 1999. "State Dependence, Serial Correlation and Heterogeneity in Intertemporal Labor Force Participation of Married Women." *Econometrica* 67 (6): 1255–1294.

Jacob, Brian, Brian McCall, and Kevin Stange. 2018. "College as Country Club: Do Colleges Cater to Students' Preferences for Consumption?" *Journal of Labor Economics* 36 (2): 309–348.

Karpowitz, Christopher F., Stephen D. O'Connell, Jessica Preece, and Olga Stoddard. 2024. "Strength in Numbers? Gender Composition, Leadership, and Women's Influence in Teams." *Journal of Political Economy* 132 (9): 3077–3114.

Kessler, Judd B., Corinne Low, and Colin D. Sullivan. 2019. "Incentivized Resume Rating: Eliciting Employer Preferences without Deception." *American Economic Review* 109 (11): 3713–3744.

Kirkebøen, Lars, Edwin Leuven, Magne Mogstad, and Jack Mountjoy. 2025. *College as a Marriage Market.* Working Paper 28688. National Bureau of Economic Research.

Koffi, Marlène. 2025. "Innovative Ideas and Gender (In)Equality." *American Economic Review* 115 (7): 2207–2236.

Koning, Rembrand, Sampsa Samila, and John-Paul Ferguson. 2021. "Who Do We Invent for? Patents by Women Focus More on Women's Health, but Few Women Get to Invent." *Science* 372 (6548): 1345–1348.

List, John A., and Jason F. Shogren. 1998. "Calibration of the Difference between Actual and Hypothetical Valuations in a Field Experiment." *Journal of Economic Behavior & Organization* 37 (2): 193–205.

List, John A., Paramita Sinha, and Michael H. Taylor. 2006. "Using Choice Experiments to Value Non-Market Goods and Services: Evidence from Field Experiments." *The B.E. Journal of Economic Analysis & Policy* 6 (2).

Long, Dede, and Yuki Takahashi. 2025. *Can Curricular Reform Close the STEM Gender Gap? Evidence from an Introductory Computer Science Course.* Working Paper.

Low, Corinne. 2024. "Pricing the Biological Clock: The Marriage Market Costs of Aging to Women." *Journal of Labor Economics* 42 (2): 395–426.

Macchi, Elisa. 2023. "Worth Your Weight: Experimental Evidence on the Benefits of Obesity in Low-Income Countries." *American Economic Review* 113 (9): 2287–2322.

Ministry of Education, Culture, Sports, Science and Technology. 2021. *On the Current State of High School Education [Original in Japanese].* Report.

Miserocchi, Francesca. 2024. *Discrimination through Biased Memory.* Working Paper.

Moriguchi, Chiaki. 2014. "Japanese-Style Human Resource Management and Its Historical Origins." *Japan Labor Review* 11 (3): 58–77.

Mouganie, Pierre, and Yaojing Wang. 2020. "High-Performing Peers and Female STEM Choices in School." *Journal of Labor Economics* 38 (3): 805–841.

Müller, Maximilian W. 2024. *Parental Pressure and Educational Choices.* Working Paper.

Nakajima, Koji. 2018. "Analysis on Deviation Value and Employment in Major Companies: Reconsideration of System for Recruiting New Graduates [Original in Japanese]." *Kansai University Journal of Higher Education* 9:57–68.

OECD. 2018. "PISA 2018 Database." https://www.oecd.org/en/data/datasets/pisa-2018-database.html.

OECD. 2019. *PISA 2018 Results: Where All Students Can Succeed.* Vol. 2. Paris, France: OECD Publishing.

Owen, Ann L., and Paul Hagstrom. 2021. "Broadening Perceptions of Economics in a New Introductory Economics Sequence." *The Journal of Economic Education* 52 (3): 175–191.

Perez, Caroline Criado. 2019. *Invisible Women: Data Bias in a World Designed for Men.* London, UK: Vintage Books.

Pestel, Nico. 2021. "Searching on Campus? The Marriage Market Effects of Changing Student Sex Ratios." *Review of Economics of the Household* 19 (4): 1175–1207.

Porter, Catherine, and Danila Serra. 2020. "Gender Differences in the Choice of Major: The Importance of Female Role Models." *American Economic Journal: Applied Economics* 12 (3): 226–254.

Riise, Julie, Barton Willage, and Alexander Willén. 2022. "Can Female Doctors Cure the Gender STEMM Gap? Evidence from Exogenously Assigned General Practitioners." *The Review of Economics and Statistics* 104 (4): 621–635.

Riley, Emma. 2024. "Role Models in Movies: The Impact of Queen of Katwe on Students' Educational Attainment." *The Review of Economics and Statistics* 106 (2): 334–351.

Schuh, Rachel. 2024. *Miss-Allocation: The Value of Workplace Gender Composition and Occupational Segregation.* Working Paper.

Shan, Xiaoyue. 2024. *Gender Diversity Improves Academic Performance.* Working Paper.

Tadjfar, Nagisa, and Kartik Vira. 2025. *Friends in Higher Places: Social Fit and University Choice.* Working Paper.

Train, Kenneth E. 2009. *Discrete Choice Methods with Simulation.* 2nd ed. Cambridge, UK: Cambridge University Press.

Truffa, Francesca, and Ashley Wong. 2025. "Undergraduate Gender Diversity and the Direction of Scientific Research." *American Economic Review* 115 (7): 2414–2448.

Valdebenito, Rocío. 2023. *Peer Influence and College Major Choices in Male-Dominated Fields.* Working Paper.

Wiswall, Matthew, and Basit Zafar. 2018. "Preference for the Workplace, Investment in Human Capital, and Gender." *The Quarterly Journal of Economics* 133 (1): 457–507.

———. 2021. "Human Capital Investments and Expectations about Career and Family." *Journal of Political Economy* 129 (5): 1361–1424.

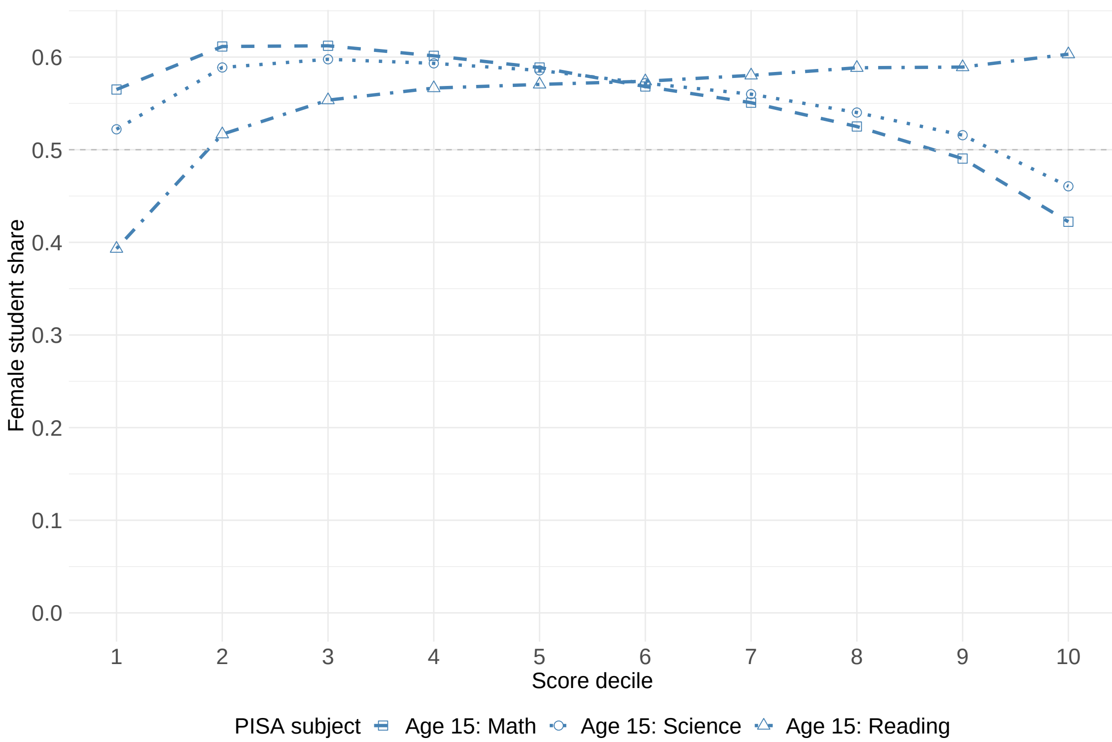
Zafar, Basit. 2013. "College Major Choice and the Gender Gap." *Journal of Human Resources* 48 (3): 545–595.

Zizzo, Daniel John. 2010. "Experimenter Demand Effects in Economic Experiments." *Experimental Economics* 13 (1): 75–98.

Online Appendix

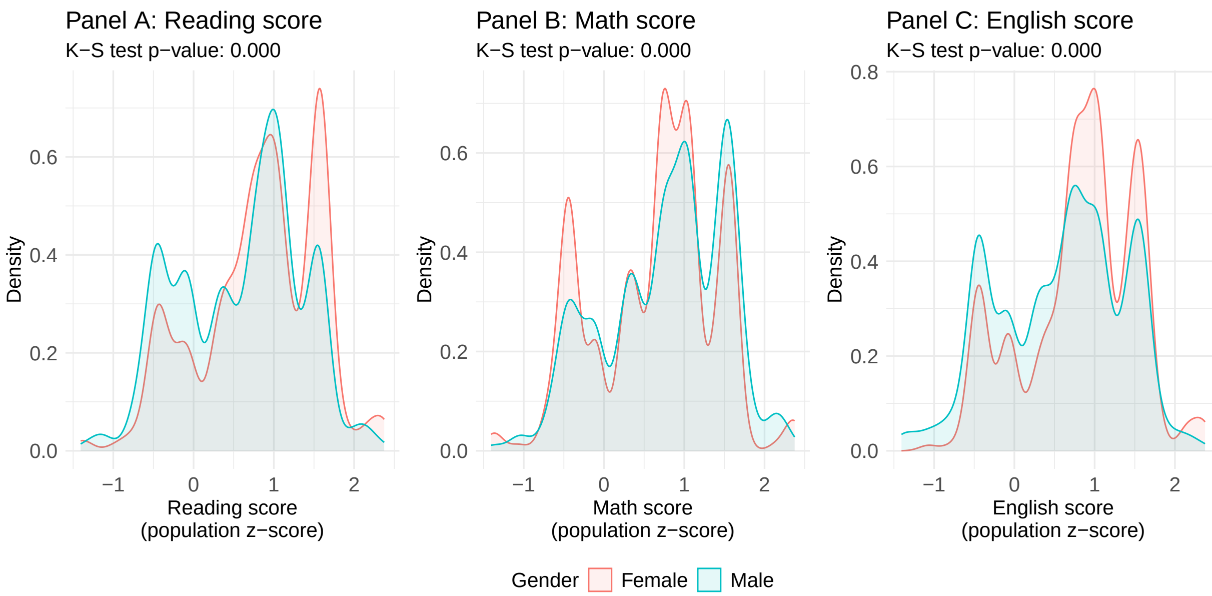
A Additional Figures and Tables

Figure A1: Talent Misallocation – OECD Average



Notes: This figure plots the OECD average share of female students in each of the 10 deciles of the PISA mathematics (dark dashed), science (dark dotted), and reading (dark dash-dotted) score distributions among 15 year-old students in 2018 who plan to pursue tertiary education. *Sources:* OECD (2018).

Figure A2: Distribution of Abilities



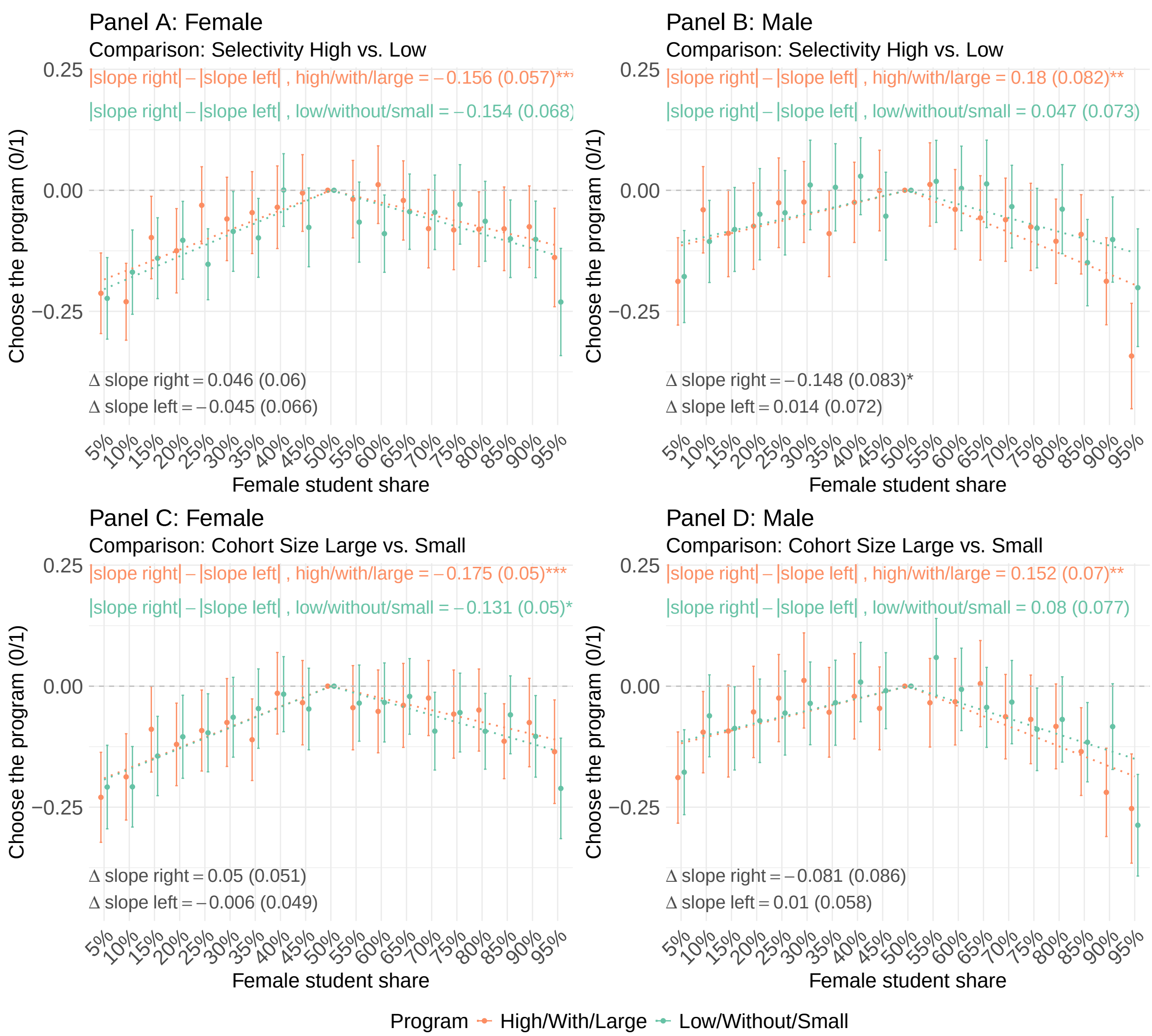
Notes: This figure presents the distribution of abilities of female and male students in reading (Panel A), mathematics (Panel B), and English (Panel C). K-S test p-value shows the Kolmogorov-Smirnov test p-value for differences in the distribution between female and male students.

Table A1: Attribute Values

General attributes	
College name:	AA, AB, AC, AD, AE, AF, AG, AH, AI, AJ, AK, AL, AM, AN, AO, AP, AQ, AR, AS, AT, AU, AV, AW, AX, AY, AZ, BA, BB, BC, BD
Department:	
Non-STEM:	Literature, Law, Business, Economics, Sociology, Foreign Language
STEM:	Physics, Chemistry, Biology, Engineering, Information Technology, Agriculture
Department attributes	
Selectivity index:	55, 57.5, 60, 62.5, 65, 67.5, 70, 72.5
Cohort size:	200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900
Female student share:	5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%
College attributes	
International exchange program:	Yes, Yes, Yes, Yes, No
Club participation rate:	40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%

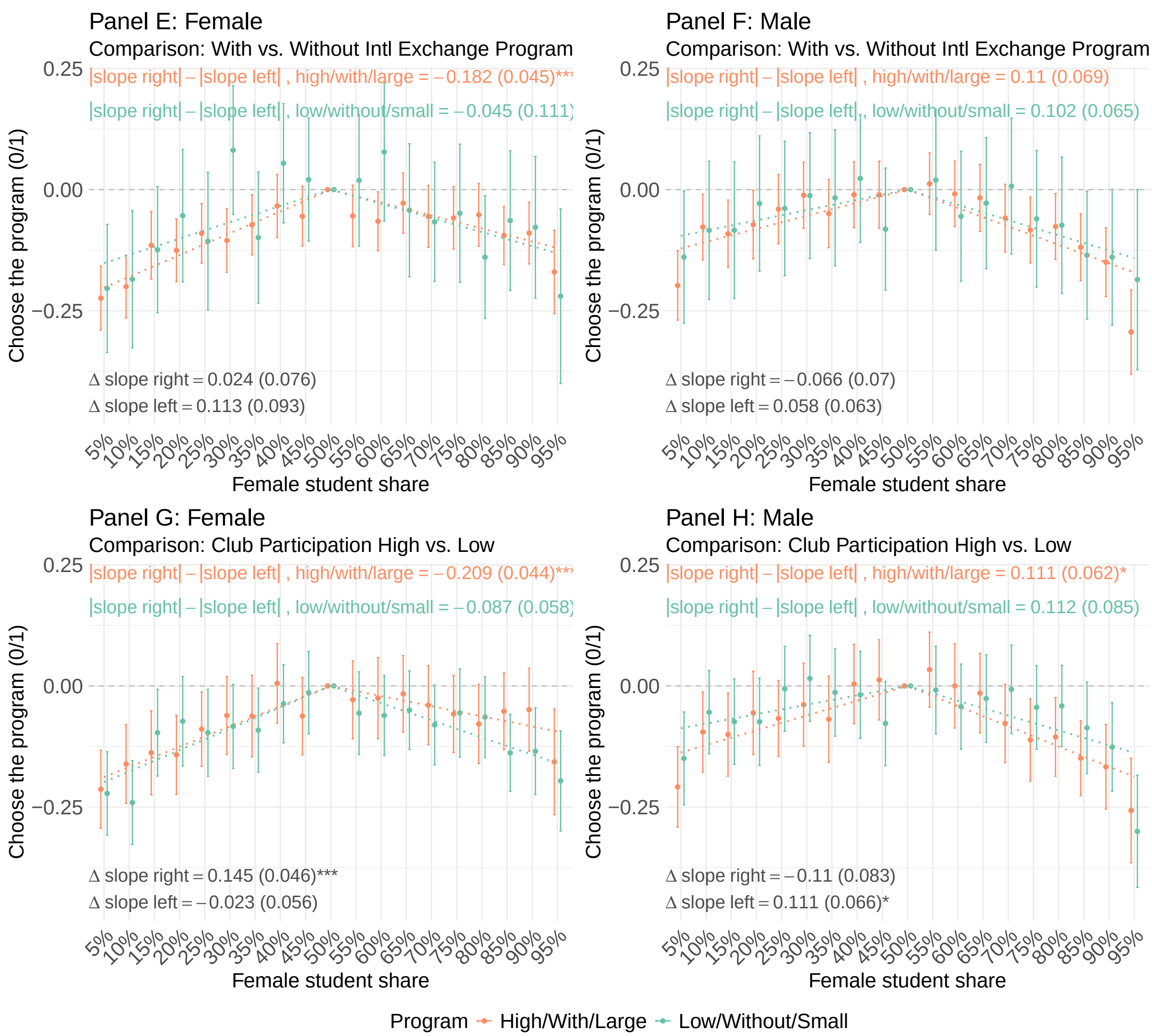
Notes: This table presents the values each attribute can take in a given program.

Figure A3: Heterogeneity of Preferences for the Gender Ratio by Selectivity and Cohort Size



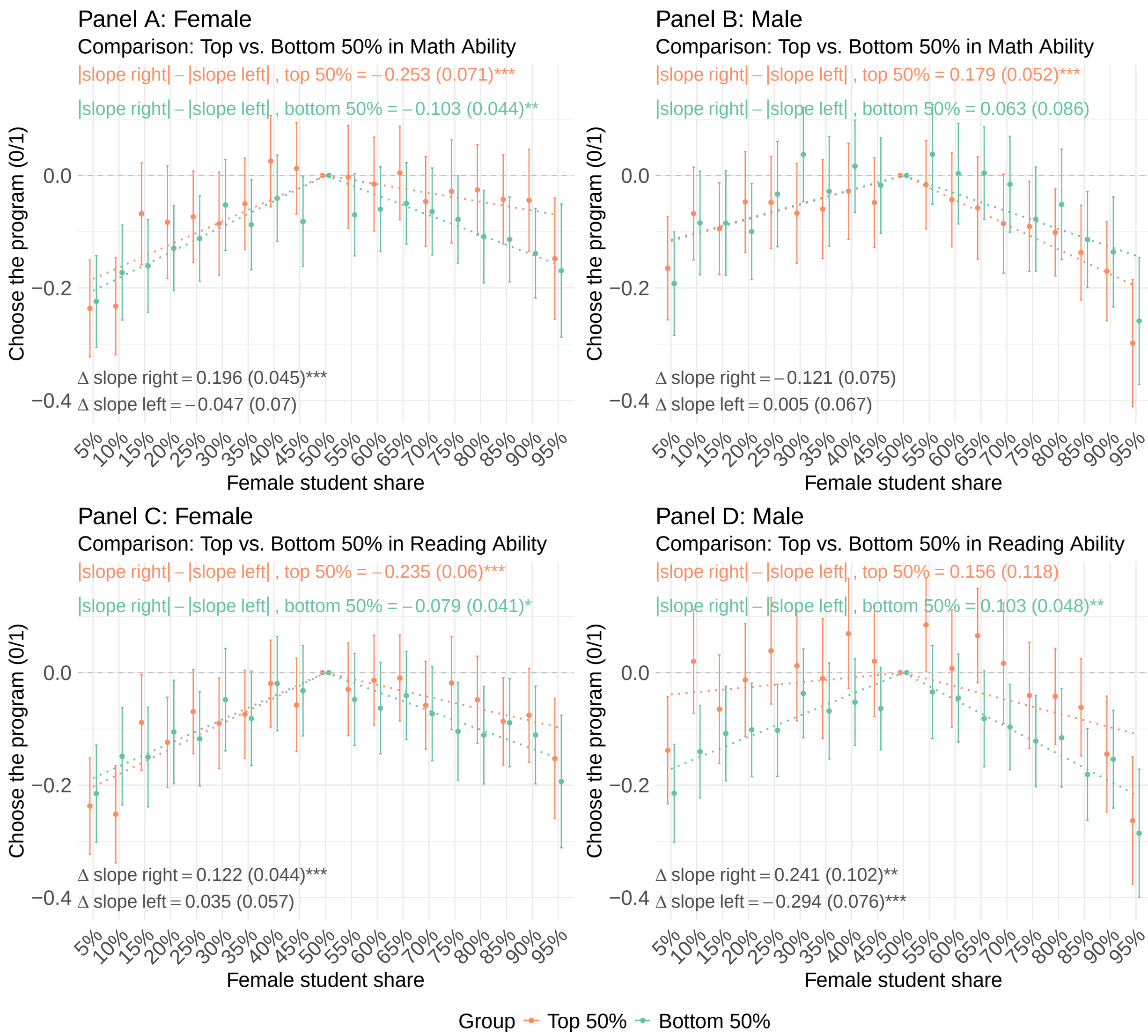
Notes: This figure plots coefficient estimates and the 95% confidence intervals for female and male students where we interact the attributes with an indicator variable for program attributes other than STEM vs. non-STEM. Panels A and B interact with an indicator variable for high selectivity program and Panels C and D interact with an indicator variable for large cohort size program. Standard errors are clustered at the student level. The dotted lines are weighted least squares linear fits of each point, with inverse variance weighting, for each side (below 50% and above 50%) for two groups. We imposed the constraint that the lines pass through the 50% point. The differences in the slopes between the two groups are calculated using the weighted least squares standard errors. Significance levels: * 10%, ** 5%, and *** 1%.

Figure A4: Heterogeneity of Preferences for the Gender Ratio by Intl Exchange and Club Participation



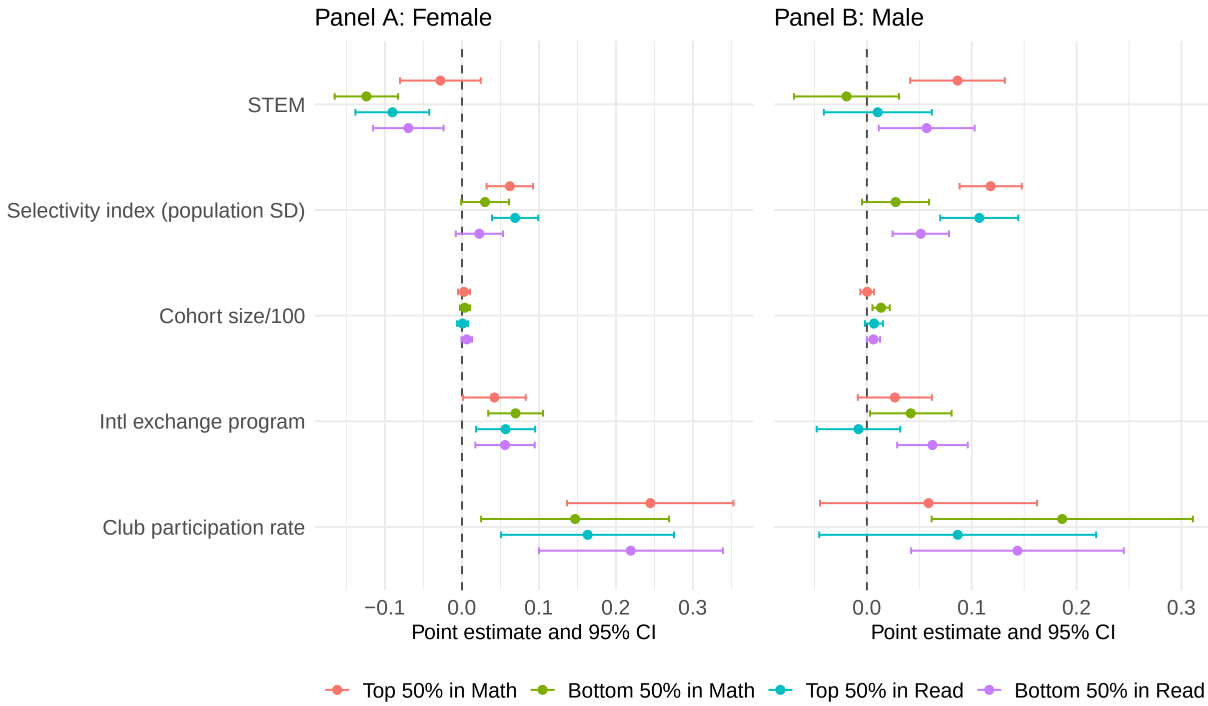
Notes: This figure plots coefficient estimates and the 95% confidence intervals for female and male students where we interact the attributes with an indicator variable for program attributes other than STEM vs. non-STEM. Panels A and B interact with an indicator variable for programs with international exchange program and Panels C and D interact with an indicator variable for high club participation rate. Standard errors are clustered at the student level. The dotted lines are weighted least squares linear fits of each point, with inverse variance weighting, for each side (below 50% and above 50%) for two groups. We imposed the constraint that the lines pass through the 50% point. The differences in the slopes between the two groups are calculated using the weighted least squares standard errors. Significance levels: * 10%, ** 5%, and *** 1%.

Figure A5: Heterogeneity of Preferences for the Gender Ratio by Abilities



Notes: This figure plots coefficient estimates and the 95% confidence intervals for female and male students where we interact the attributes with an indicator variable for academic abilities. Panels A and B interact with an indicator variable for top 50% mathematics abilities students and Panels C and D interact with an indicator variable for top 50% reading abilities students. Standard errors are clustered at the student level. The dotted lines are weighted least squares linear fits of each point, with inverse variance weighting, for each side (below 50% and above 50%) for two groups. We imposed the constraint that the lines pass through the 50% point. The differences in the slopes between the two groups are calculated using the weighted least squares standard errors. Significance levels: * 10%, ** 5%, and *** 1%.

Figure A6: Heterogeneity of Preferences for Other Attributes by Ability



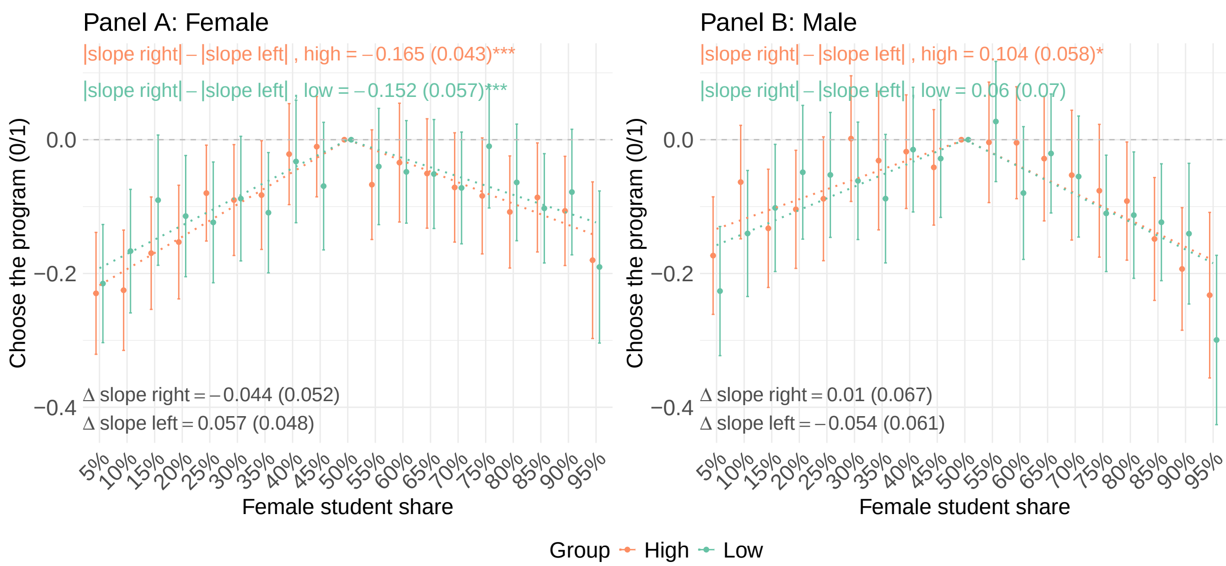
Notes: This figure plots coefficient estimates and the 95% confidence intervals of program attributes for female (Panel A) and male students (Panel B) with different academic abilities. Female student share is included in the estimation but omitted from this figure for brevity. Standard errors are clustered at the student level.

Table A2: Preferences for Program Attributes – Reasons as Dependent Variables

Sample:	Female	Male	Female	Male	Female	Male	Female	Male
Outcome:	Fit in		Role model		Studies		Career	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
STEM	-0.037***	0.007	-0.047***	0.011	-0.061***	0.041***	-0.062***	0.021
	(0.012)	(0.013)	(0.013)	(0.013)	(0.015)	(0.015)	(0.016)	(0.017)
Female student share	1.819***	1.099***	0.708***	0.605***	0.239***	0.453***	0.336***	0.407***
	(0.089)	(0.092)	(0.085)	(0.089)	(0.086)	(0.084)	(0.087)	(0.085)
Female student share squared	-1.477***	-1.384***	-0.621***	-0.685***	-0.225***	-0.503***	-0.320***	-0.450***
	(0.094)	(0.093)	(0.084)	(0.089)	(0.084)	(0.086)	(0.086)	(0.086)
Selectivity index (population SD)	-0.014	0.015	0.088***	0.080***	0.084***	0.094***	0.075***	0.080***
	(0.009)	(0.010)	(0.011)	(0.011)	(0.015)	(0.014)	(0.010)	(0.011)
Cohort size/100	0.002	0.005*	0.012***	0.013***	-0.002	0.000	0.004*	0.006**
	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Intl exchange program	0.021	0.020	0.049***	0.040***	0.034***	0.047***	0.053***	0.051***
	(0.014)	(0.013)	(0.014)	(0.012)	(0.013)	(0.013)	(0.014)	(0.015)
Club participation rate	0.282***	0.200***	0.651***	0.443***	0.086**	0.070*	0.165***	0.158***
	(0.042)	(0.043)	(0.054)	(0.050)	(0.039)	(0.041)	(0.040)	(0.043)
Constant	0.504***	0.479***	0.518***	0.490***	0.483***	0.480***	0.494***	0.495***
	(0.007)	(0.007)	(0.007)	(0.008)	(0.007)	(0.008)	(0.007)	(0.008)
Adj. R-squared	0.159	0.100	0.104	0.062	0.032	0.037	0.035	0.029
No. observations	4649	4451	4649	4451	4649	4451	4649	4451
No. students	310	297	310	297	310	297	310	297

Notes: This table presents the same specifications as Table 2 but with indicator variables for the four reasons in place of choice as the outcome variables. Significance levels: * 10%, ** 5%, and *** 1%.

Figure A7: Heterogeneity of the Preferences by Parental Education Levels



Notes: This figure plots coefficient estimates and the 95% confidence intervals for female and male students, just like Figure 6, but interact the attributes with an indicator variable for whether both parents have bachelor's degrees or above (High if 1, Low if 0). Standard errors are clustered at the student level. The dotted lines are weighted least squares linear fits of each point, with inverse variance weighting, for each side (below 50% and above 50%) for two groups. We imposed the constraint that the lines pass through the 50% point. The differences in the slopes between High and Low are calculated using the weighted least squares standard errors. Significance levels: * 10%, ** 5%, and *** 1%.

Table A3: Preferences for Program Attributes – Logit

Sample:	Female		Male		All	
Outcome:	Choose the program (0/1)					
	(1)	(2)	(3)	(4)	(5)	(6)
STEM	-0.078***	-0.078***	0.038**	0.038**	0.037**	0.038**
	(0.018)	(0.018)	(0.018)	(0.017)	(0.017)	(0.017)
Female student share	0.090***	0.819***	-0.036	0.815***	-0.036	0.811***
	(0.022)	(0.097)	(0.024)	(0.097)	(0.024)	(0.094)
Female student share squared		-0.742***		-0.872***		-0.868***
		(0.095)		(0.099)		(0.096)
Selectivity index (population SD)	0.046***	0.046***	0.074***	0.074***	0.074***	0.074***
	(0.011)	(0.011)	(0.012)	(0.012)	(0.012)	(0.012)
Cohort size/100	0.003	0.003	0.006**	0.006**	0.006**	0.006**
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Intl exchange program	0.059***	0.060***	0.033**	0.034**	0.032**	0.034**
	(0.014)	(0.014)	(0.013)	(0.013)	(0.013)	(0.013)
Club participation rate	0.193***	0.196***	0.122***	0.130***	0.121***	0.130***
	(0.042)	(0.043)	(0.041)	(0.041)	(0.041)	(0.041)
Female					0.016	0.015
					(0.011)	(0.011)
STEM x Female					-0.116***	-0.116***
					(0.025)	(0.025)
Female student share x Female					0.127***	0.011
					(0.033)	(0.131)
Female student share squared x Female						0.123
						(0.131)
Selectivity index (population SD) x Female					-0.028*	-0.027*
					(0.016)	(0.016)
Cohort size/100 x Female					-0.003	-0.003
					(0.004)	(0.004)
Intl exchange program x Female					0.027	0.026
					(0.019)	(0.019)
Club participation rate x Female					0.074	0.067
					(0.059)	(0.059)
Log Likelihood	-3133.13	-3090.73	-3032.28	-2977.35	-6165.41	-6068.08
No. observations	4649	4649	4451	4451	9100	9100
No. students	310	310	297	297	607	607

Notes: This table presents the logit coefficient estimates in average marginal effects on the program attributes with choice as the dependent variable. Columns 1 and 2 present estimates for female students, columns 3 and 4 present estimates for male students, and columns 5 and 6 present estimates for differences between female and male students. The average marginal effects for constant term are undefined and thus are not shown. Standard errors are clustered at the student level. Significance levels: * 10%, ** 5%, and *** 1%.

Table A4: Mixed Logit Parameter Estimates

Sample:	Female	Male
Outcome:	Choose the program (0/1)	
	(1)	(2)
STEM	-0.556***	0.240***
	(0.054)	(0.051)
Female student share	5.394***	4.916***
	(0.524)	(0.512)
Female student share squared	-4.823***	-5.313***
	(0.514)	(0.518)
Selectivity index (population SD)	0.290***	0.522***
	(0.052)	(0.056)
Cohort size/100	0.025*	0.028*
	(0.014)	(0.014)
Intl exchange program	0.392***	0.179**
	(0.075)	(0.075)
Club participation rate	1.311***	0.819***
	(0.231)	(0.237)
Intercept for right	-0.005	-0.049
	(0.040)	(0.040)
Log Likelihood	-2708.980	-2606.917
No. observations	9298	8902
No. students	310	297

Notes: This table presents mixed logit parameter estimates for female students (column 1) and male students (column 2). We assume the density for β_i^g is a normal distribution and allow arbitrary correlations among elements in β_i^g. We make 1000 Halton quasi-Monte Carlo draws to evaluate the integral instead of pure Monte Carlo draws for stability and faster convergence as suggested by (Train 2009, Section 9). Significance levels: * 10%, ** 5%, and *** 1%.

B Post-Experimental Questionnaire

Post-Experimental Questionnaire (English translation)

Questionnaire 1/4

Please tell us about yourself and your family.

- Your gender: [Male, Female, Non-binary or Other]
- Your father's academic background: [Below high school, High school, Vocational school, Associate degree, Bachelor's degree, Master's degree or above, I do not know]
- Your mother's academic background: [Below high school, High school, Vocational school, Associate degree, Bachelor's degree, Master's degree or above, I do not know]
- Extra schooling per week: [No extra schooling, one day a week, two days a week, three days a week, four days a week, five days a week or more]

Please recall the exam held on [Month Day]. What was your score in the following subjects?

- Reading: [Integer]
- Mathematics: [Integer]
- English: [Integer]

Please click "→" to proceed.

——— Page break ———

Questionnaire 2/4

Do you think **your scores in the exam held on [Month Day] accurately reflect your abilities?** Please answer for each of the subjects below.

	My abilities are lower than the score	My abilities are slightly lower than the score	It reflects my ability accurately	My abilities are slightly higher than the score	My abilities are higher than the score
Reading	○	○	○	○	○
Mathematics	○	○	○	○	○
English	○	○	○	○	○

- Do you consider yourself someone who is **willing to compete with others**, or someone who **avoids competing with others**? [Avoid competing with others, Slightly avoid competing with others, Neither avoid nor willing to compete with others, Slightly willing to compete with others, Willing to compete with others]

- Do you consider yourself someone who is generally **willing to take risks**, or someone who **avoids taking risks**? [Avoid taking risks, Slightly avoid taking risks, Neither avoid nor willing to take risks, Slightly willing to take risks, Willing to take risks]

Please click "→" to proceed.

——— Page break ———

Questionnaire 3/4

What do you think is **the average female-student ratio** in the following departments across the colleges in Japan?

	Below 10%	11-20%	21-30%	31-40%	41-50%	51-60%	61-70%	71-80%	81-90%	91% or above
Humanities Departments (Literature, history, philosophy, etc.)	○	○	○	○	○	○	○	○	○	○
Social Sciences Departments (Law, Economics, Sociology, etc.)	○	○	○	○	○	○	○	○	○	○
Sciences and Engineering Departments (Physics, Biology, Mechanical Engineering, Information Technology, etc.)	○	○	○	○	○	○	○	○	○	○
Medicine and Nursing Departments (Medicine, Dentistry, Pharmacy, Nursing, etc.)	○	○	○	○	○	○	○	○	○	○

Please click “→” to proceed.

——————————— Page break ———————————

Questionnaire 4/4

Please tell us your opinion about this survey.

- Was it easy to follow? [Difficult to follow, Slightly difficult to follow, Neither difficult nor easy to follow, Slightly easy to follow, Easy to follow]
- Which parts did you find it difficult to answer? [Text]
- What do you think is the purpose of this survey? [Text]
- Other comments? (optional) [Text]