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

Chihiro Inoue Asumi Saito Yuki Takahashi
Job Market Paper

August 11, 2025

Abstract

Although the gender gap in mathematics and sciences in OECD countries is negligible, female students are still less likely to major in STEM fields in college, which can lead to several social issues. One explanation that has received less attention in the literature is that STEM programs are predominantly male-dominated, which makes female students a minority. We study whether the gender ratio at colleges affects high school students' college choices and the extent to which it contributes to the gender gap in STEM programs. We begin by using administrative data to show that the gender ratio has become more balanced in both STEM and non-STEM programs over the last 16 years, especially in programs where students are supposed to have less trade-off among attributes. We then use an incentivized discrete choice experiment and show that the gender ratio at colleges does affect both female and male students' college choices: both female and male students prefer gender-balanced college programs over those with a male or female majority. Students avoid programs where they will be a minority mainly because they expect it to be difficult to fit into such environments. A counterfactual analysis suggests that the low female student share in STEM programs reduces the likelihood of female students choosing STEM programs by 6.0 percentage points or 15.7%, and they incur a utility cost equivalent to 0.58 standard deviations of program selectivity. Removing this constraint would lead to female students with high mathematics ability replacing male students with low mathematics but high reading ability in STEM. Thus, the gender ratio at colleges is an important factor for high school students' college choices, and making STEM programs more gender-balanced can help narrow the STEM gender gap and address social issues arising from it, and can improve the allocation of talent.

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: Department of Economics, Tilburg University, email: y.takahashi@uvt.nl.

We thank Sonia Bhalotra, Jochem de Bresser, Caterina Calsamiglia, Damian Clarke, Lucia Corno, Sota Ichiba, Clément Imbert, Shoya Ishimaru, Daiki Kishishita, Boon Han Koh, Annalisa Loviglio, Akira Matsushita, Daan van Soest, Takeshi Murooka, Johannes Rincke, Teodora Tsankova, Atsushi Yamagishi, Basit Zafar, and participants at EEA-ESEM Rotterdam, the ERATO Meeting, LEER Conference, the Trans Pacific Labor Seminar, 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. Keigo Furukawa, Mifuyu Kira, and Yoshino Kodama 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 pre-registered on the AEA Registry under AEARCTR-0012577.

1 Introduction

Although female and male students perform equally well in mathematics and sciences at age 15 in OECD countries (OECD 2019), female students remain less likely to major in Science, Technology, Engineering, and Mathematics (STEM) at colleges (OECD 2018). 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 not received much attention in the literature is that STEM programs are predominantly male-dominated and make female students a gender minority in the programs.[1] Indeed, several studies show disadvantages of being a gender minority in the workplace and schools (Bostwick and Weinberg 2022; Cullen and Perez-Truglia 2023; Folke and Rickne 2022; Hampole, Truffa, and Wong 2024). Thus, it is possible that female students anticipate these disadvantages and avoid programs that are male-dominated.

This paper studies whether the gender ratio at colleges affects high school students' college choices and the extent to which it contributes to the gender gap in STEM programs. We first establish empirical patterns about the gender ratios across college programs over the years using administrative data that covers 16 years of the universe of college programs in Japan. We then conduct an incentivized discrete choice experiment with students from selective academic high schools in Japan to elicit their preference over independently varied college program attributes, including the gender ratio, STEM or non-STEM, and program selectivity. In the experiment, students see 15 hypothetical college program pairs one by one with randomly assigned attributes and choose one of them that they want to attend. After that, they see four statements that many students care about when choosing college programs, and choose which program these four statements better apply to for each pair.[2] We incentivize these choices using the incentivized resume rating method (Kessler, Low, and Sullivan 2019) by providing career advice that is tailored based on their choices.

We first document using the administrative data that the gender ratio has become more balanced in both STEM and non-STEM programs from 2008 to 2023. However, the change is more pronounced in programs where students are supposed to have more choices for the attributes and thus the trade-off between gender ratios and other attributes are less binding. We then use the experimental data and show that the gender ratio at colleges does affect students' college choices of both female and male students. Specifically, both female and male students prefer gender-balanced programs over male- or female-majority programs. Decomposition of their choices shows that both female and male students avoid being a minority in a program mainly because they expect it to be difficult to fit into such an environment. On the other hand, students avoid being a majority for various

1. A notable exception is Ersoy and Speer (2025), who show that students consider non-job-related factors when choosing their college major, including student gender balance.

2. These four statements are selected from the open-ended questions in the pre-test.

reasons. We do not find that the preference for gender balance differs whether the program is STEM or non-STEM, whether a student excels in academic abilities or not, or whether students are from households with high or low socio-economic status (SES). Thus, the preference for balance is present in both STEM and non-STEM programs, regardless of students' academic abilities or SES backgrounds.

We then move to counterfactual analysis to quantify the extent to which preferences for the gender balance contribute to the gender gap in STEM programs where the female student share is low. Extracting individual-level student preference parameters using mixed logit (Train 2009), we show that while 38% of female students prefer STEM over non-STEM programs, the low female student share in STEM reduces their STEM choice probability by 6.0 percentage points or 15.7%, and they incur the utility cost equivalent to 0.58 standard deviation program selectivity. The gender gap in STEM is reduced by 18.5% if this constraint were removed. We also show that removing the constraint does not reduce applicant quality; on the contrary, it can improve the talent allocation: it would increase the share of female students with high mathematics abilities choosing STEM programs and reduces the share of male students with low mathematics abilities but high reading abilities choosing STEM programs.

Taken together, our findings suggest that the gender ratio at colleges is an important determinant for high school students' college choices, and making STEM programs more gender-balanced can help close the STEM gender gap and address social issues and can improve talent allocation.

Related Literature Our paper contributes to several strands of research. First, it enriches and extends the emerging literature on the preferences for gender balance in the workplace. Schuh (2024) finds that US workers, both women and men, have preferences for gender diversity in the workplace, and it reduces female employment in male-dominated occupations. Relatedly, Högn et al. (2025) find that German university students are willing to pay about 5% of their expected salary for a gender diverse future workplace, and that female students have a higher willingness to pay than male students. We add to these studies by showing that high school students also have preferences for gender balance at colleges and that the reason for avoiding being a gender minority is a concern for the difficulty of fitting into programs with a low own gender share.

Second, our paper contributes to the literature on the disadvantages of being a gender minority. Folke and Rickne (2022) find that women receive more sexual harassment in male-dominated jobs, contributing to gender differences in occupational sorting. Interestingly, they find that men also receive more sexual harassment in female-dominated jobs. Relatedly, Cullen and Perez-Truglia (2023) find that workers assigned to an opposite gender manager are less likely to get promotions due to less frequent social interactions. In educational settings, Bostwick and Weinberg (2022) find that female STEM PhD students are more likely to drop out of the program when they are assigned to a male-majority cohort, Hampole, Truffa, and Wong (2024) show that female MBA graduates are less likely to advance to a senior management positions than male counterparts when there are fewer female peers, and Karpowitz et al. (2024) show that female students in a male-majority work teams in a college course have less influence to the team than male students. Shan (2024), on the other hand, finds that female students in introductory economics courses achieve lower grades when placed in female-only study groups. We complement this literature by showing that women and men make college choices in anticipation of these disadvantages, which constrain their choices for male- or female-dominated fields.

Third, our paper contributes to the literature on the policies to close the gender gap in STEM. The most prominent policy so far is role model intervention where we expose students to successful female STEM professionals (Breda et al. 2023; Carrell, Page, and West 2010; Riise, Willage, and Willén 2022; Riley 2024).[3] Another policy is changing pedagogical practices where we either make the learning process more interactive (Di Tommaso et al. 2024) or teach the societal relevance of the discipline (Long and Takahashi 2025).[4] Our findings suggest that a college gender quota can be another effective policy, where we ensure a certain fraction of students will be women in STEM programs.

Finally, our study contributes to the literature on the determinants of college major choice. Previous studies find that women tend to prefer majors that lead to flexible and stable jobs, whereas men prefer majors that are associated with higher earnings (Wiswall and Zafar 2018). Women are also more likely than men to consider future family formation when selecting majors (Wiswall and Zafar 2021). In addition, Ersoy and Speer (2025) show that providing information about majors to students in the US, including student gender composition and other non-job-related factors, affects their major choices. We build on these studies, especially Ersoy and Speer, and show that the student gender ratio is indeed an important determinant for students' major choice, and the main reason for avoiding being in a major dominated by the opposite gender is expected difficulty in fitting in.

The remainder of the paper is structured as follows. Section 2 explains the Japanese high school and college application system, as well as the job market that comes afterwards. Section 3 details the experimental design. Section 4 describes summary statistics of the experimental data. Section 5 presents the main results. Section 6 provides results of the counterfactual analysis. Section 7 discusses policy implications of the results. Section 8 concludes.

2 Institutional Background

2.1 Japanese Education System

High School In the Japanese educational system, high school runs from grades 10 to 12, typically from age 15 to 18. Although it is not compulsory, nearly 99% of junior high school graduates attend it (Ministry of Education, Culture, Sports, Science and Technology 2021). College enrollment rate is also high, nearly 60%.[5] However, just like in the US, many colleges are similar to vocational

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

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

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

schools in Europe, and only a handful of top colleges are academically oriented. We call the former as non-selective and the latter as selective in this paper. In addition, as Science, Technology, Engineering, and Mathematics (STEM) requires good mathematical skills, most STEM programs are in selective colleges. As such, not all high schools can prepare students for selective colleges or STEM programs; only academic high schools can do so. Thus, the choice between STEM and non-STEM exists primarily in academic high schools. Students at academic high schools regularly take exams from grade 10 to prepare for college entrance exams.

Students at academic high schools choose a track at the end of grade 10, and the choice 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, on the other hand, students study reading (Japanese), English, sciences (biology, chemistry, and/or physics), and advanced mathematics. As discussed later, the track choice limits students' college majors.

A similar track system exists in other countries as well, such as Italy, where students choose their track between humanities and sciences at the end of grade 9 (see, for example, Carlana and Corno 2022), France, where students choose their track between humanities, social sciences, and sciences at the end of grade 10 (see, for example, Breda et al. 2023), and the Netherlands, where students in academic secondary schools (VMBO) choose their track between science, health, social sciences, and humanities at the end of grade 9 (see, for example, Buser, Niederle, and Oosterbeek 2014).

College Application Like in Europe but unlike in the US (Bordon and Fu 2015), students apply for specific college programs and cannot change their major later. Programs vary in their attributes, such as major, selectivity, tuition, public or private, and location, among others. As living alone can be expensive, and some parents do not want their daughters to live alone, many college students live with their parents. Among the various locations in Japan, the greater Tokyo area offers the widest variety of college attributes: nearly 29% of all colleges are in that region, where about 41% of all college students study (Obunsha 2024).

Most programs employ an exam-based, meritocratic admission system. Programs rank applicants based on their exam score and make offers from the top of the list. However, each program requires exams on different subjects. Humanities and social sciences programs typically require exams on advanced reading, English, history/social studies, and mathematics. On the other hand, science, engineering, and medicine programs typically require exams on advanced mathematics, English, sciences, and reading. Since students who choose the science track in high school do not study advanced readings and history/social studies, they cannot apply for humanities and social sciences programs. On the other hand, since students who choose the humanities track in high school do not study advanced mathematics and sciences, they cannot apply for science, engineering, and medicine programs. In this way, high school track choice determines students' college majors.

Job Search Colleges are very important signaling devices in the Japanese job market: each program is assigned a single number called the "selectivity index," which is 50 plus the standard deviation of a given program multiplied by 10. The index of a program a student graduates from significantly affects their job search, both at the extensive and intensive margins. At the extensive margins, it is highly associated with the quality of jobs a student can get (Nakajima 2018). At the intensive margins, it is associated with the promotions in the first few years after starting the job (Araki, Kawaguchi, and Onozuka 2016).

The Japanese labor market is characterized by its rigidity with very limited job mobility (Moriguchi 2014). On-the-job training is the norm, and workers accumulate non-transferable skills, working at the same company for their entire career. Indeed, the labor market has been gradually changing, and the job separation rate has been increasing, especially among young workers (Kambayashi and Kato 2017). However, sticking to the same company for the entire career is still the norm, especially for graduates from selective colleges that we consider in this study. Thus, after getting the first job, most students stay in the first job and rarely switch jobs. As a result, the quality of the first job significantly affects their career prospects. Hence, going to a college program with a higher selectivity index is very important for high school students as it determines their career prospects.

2.2 Descriptive Evidence: Changes in the Gender Ratio at Colleges over 16 Years

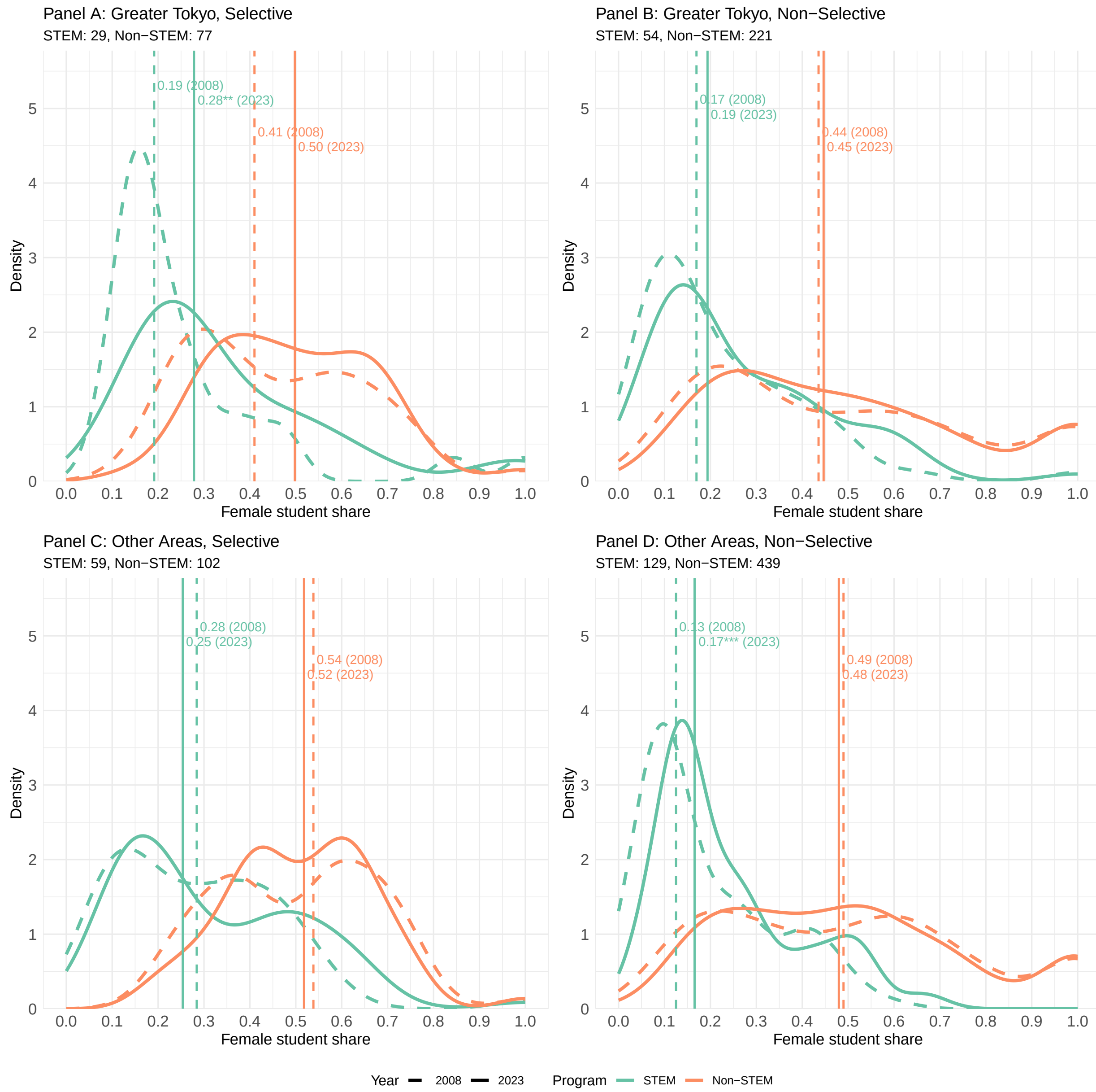
We provide some descriptive evidence that the gender ratio at college programs affect high school students' college choices using the School Basic Survey collected by the Ministry of Education, Culture, Sports, Science and Technology, which covers the universe of college programs in Japan from 2008 to 2023. Figure 1 plots female student share across all college programs in Japan in 2008 (dashed line) and 2023 (solid line) for STEM (green) and non-STEM (orange). Panel A shows the share for the selective programs (with the selectivity index of 55 or above, or 0.5 standard deviation above the mean or above) in the Greater Tokyo area, Panel B shows the share for the non-selective programs in the Greater Tokyo area, Panel C shows the share for the selective programs in the other areas, and Panel D shows the share for the non-selective programs in the other areas.[6,7] We drop programs that only exist in 2008 or 2023.

The figure shows that both in 2008 and 2023, male students are indeed the majority in STEM programs regardless of the areas or the selectivity. It also shows that, although non-STEM have roughly the same number of female and male students on average, male students are still over-represented in selective programs in the Greater Tokyo area.

6. The selectivity index of each program is defined as of 2023. The data was obtained from the list prepared by Kawaijuku, one of the most popular commercial college exam preparation companies in Japan: https://www.keinet.ne.jp/exam/ranking/index.html (accessed December 18, 2024). As public colleges require larger number of subjects in the entrance exam, we follow Araki, Kawaguchi, and Onozuka (2016) and add 5 to the index of public colleges.

7. Programs with a selectivity index of 55 or higher are classified as rank B or above in Goodman and Oka (2021)'s classification, which aligns with what Japanese people consider a selective program.

Figure 1: Female Student Share across College Programs in 2008 and 2023



Notes: This figure plots female student share across all college programs in Japan in 2008 (dashed line) and 2023 (solid line) for STEM (green) and non-STEM (orange). Panel A shows the share for the selective programs (with the selectivity index of 55 or above, or 0.5 standard deviation above the mean or above) in the Greater Tokyo area, Panel B shows the share for the non-selective programs in the Greater Tokyo area, Panel C shows the share for the selective programs in the other areas, and Panel D shows the share for the non-selective programs in the other areas. We drop programs that only exist in 2008 or 2023.
Source: School Basic Survey of the Ministry of Education, Culture, Sports, Science and Technology.

However, when we look at the changes in gender ratios from 2008 to 2023, programs in the Greater Tokyo area have become more gender balanced over the 16 years, especially among selective, academically-oriented programs, which have little emphasis on vocational training. This may be because students in the Greater Tokyo area have more choices in program attributes than students in other areas, making the trade-off between gender ratios and other attributes less binding.

We look at whether this shift towards gender balance is driven by high school students' preferences using discrete choice experiment.

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 conducted the 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 for three high schools and asynchronously online for one high school.

The experiment took an average of 40 minutes, including handing out the participation gifts. In total, 628 students participated, of which 619 provided valid responses (311 females, 298 males, 10 non-binaries). For the purpose of this study, which focuses on binary gender, we excluded non-binary students' responses from the analysis, leaving us with 609 responses with 15 observations each, yielding a total of 9135 observations.

3.1 Sample Selection

Schools We contacted teachers at academic high schools in the greater Tokyo area in Japan through our network and received their consent to conduct the experiment within 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 that the experimental content was relevant for students: the experiment was about college choices, and they must be planning to attend a college. Based on the schools' placement record, more than 95% of recent graduates attended college. Second, we wanted to have students whose mathematics abilities did not constrain their major choices: as discussed in Section 2, most STEM programs require good mathematics abilities, and only academic high schools can prepare students for STEM. Third, we did not want students take implicitly into account for potential location and financial constraints when making choices: as discussed in Section 2, the greater Tokyo area offers the widest variety of college programs, and these constraints are less likely to be in students' mind in this region. Fourth, we wanted the attributes in the hypothetical college programs to appear natural to students, and we wanted students to have a variety of program options available.

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 the 10th-grade students, prior to their track choice that 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. Instead, we explained that students would evaluate 15 hypothetical college programs and answer a short questionnaire, and receive a tailored career advice sheet based on the evaluations. We also explained that we would use the data from the experiment for academic research to improve the education policy. Almost 90% of the guardians and students provided consent and participated in the experiment.

3.2 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 that we would not share their responses or career advice sheet with anyone, including guardians, teachers, or peers, to minimize their influences as prior studies suggest they can influence students' study choices (Carlana 2019; Carlana and Corno 2022; Giustinelli 2016; Müller 2024). Students then saw 15 program pairs one by one and chose one of them that they wanted to attend; see Figure 2 for an example of a hypothetical program pair. In addition, they see four statements that many students care about when choosing college programs, selected from open-ended questions in the pre-test. For each pair, the student chose which program each of the statements better applies to.

After making choices for the 15 program pairs, students answered 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. After the questionnaire, students received a participation gift (a set of cute, functional pens) equivalent to 500 JPY (approx. 5.27 USD in 2022 PPP) for their participation.[8]

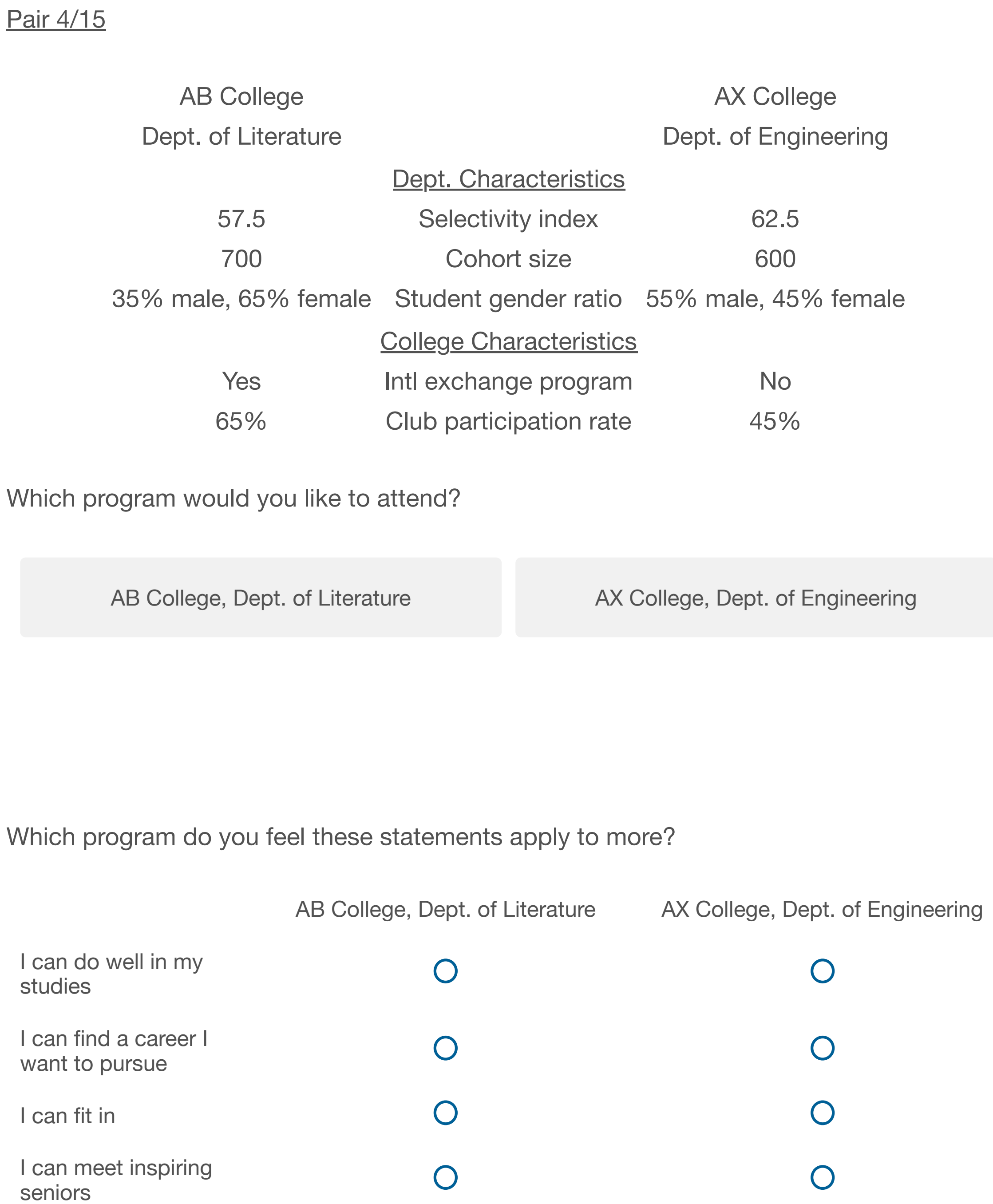
Several weeks later, each student received a career advice sheet we created. The career advice sheet consisted of a tailored part and a non-tailored part. The tailored part included the first three attributes students cared most about, as well as the top reason they care most about when choosing a college program, derived from the choices in the hypothetical programs. The non-tailored part contained non-individualized information valuable for most high school students, such as tips for choosing college programs, college admissions, financing, studying abroad, going to graduate schools, and getting jobs.

3.3 Attributes

We assigned each program randomly drawn attributes: 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. Among the attributes, our interests are (i) student gender ratio and (ii) department, which signals whether the program is STEM or non-STEM. We included other attributes to make the programs appear more natural to students and selected attribute value ranges that are realistic to students in our sample

8. 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).

Figure 2: Hypothetical Program Pair



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

to mitigate hypothetical bias (List and Shogren 1998; List, Sinha, and Taylor 2006). We asked them to assume that attributes not shown were the same between the programs.

College names consist of two alphabets and we draw them without replacement for each program in a given pair from a list, from AA to BD, that were unrelated to the actual college names. The Department was drawn from a list of 12 popular departments, where 6 departments were STEM and the other 6 departments were non-STEM. We first drew either STEM or non-STEM for one of the programs in a given pair. If we drew STEM, then we drew non-STEM with 75% probability and STEM with 25% probability for the other program in the pair to reduce the likelihood that both programs were STEM or non-STEM. We then drew a department within the list for STEM and non-STEM. STEM departments are Physics, Chemistry, Biology, Engineering, Information Technology, and Agriculture. Non-STEM departments are Literature, Law, Business, Economics, Sociology, and Foreign Language. We did not include the Medicine and Education departments. Although both are also popular, graduating from these programs leads to specialized occupations such as medical doctors, nurses, pharmacists, and school teachers, which are different from the majority of careers students would pursue, and the attributes we included may not be relevant for these programs.

Other attributes such as selectivity index that takes value from 55 to 72.5 with an increment of 2.5, cohort size that takes value from 200 to 900 with an increment of 50, student gender ratio that takes value of 5% to 95% for females and males but sum up to 100%, whether the college has an international exchange program that takes value of "Yes" with 80% of the time and "No" with 20% of the time, and club participation rate that take value from 40% to 85% with an increment of 5%.[9] These other attribute values were drawn with replacement for each program. Appendix Table A1 presents the values that each attribute can take.

3.4 Incentives

We used a hypothetical choice experiment as it allowed us to elicit students' preferences over independently varied attributes, which was essential for our study as there were very few STEM programs where male students were not the majority, and actual college names make students infer attributes not shown to them.

One concern with choice experiments is that students may not have incentives to state their true preferences without actual consequences. Although Hainmueller, Hangartner, and Yamamoto (2015) shows that choices in hypothetical vignettes and the actual behaviors are strongly correlated, we addressed this concern by incentivizing the choices using the incentivized resume rating method (Kessler, Low, and Sullivan 2019) by providing career advice based on their choices. Under the assumption that students considered we experimenters had new information that was potentially valuable for them, it is incentive compatible: the expected value of the advice is monotonically increasing in the truthfulness of their choices. In particular, because the students were from academic high schools interested in attending selective colleges, we assume most of them considered the experimenters had valuable academic and career information.[10]

Specifically, we provided the following information in the information letter and at the beginning

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

10. At the time of the experiment, Inoue obtained a PhD degree at one of Japan's most prestigious colleges, Saito obtained a master's degree in the US and has been active in industry, and Takahashi obtained a PhD degree from a university in Europe.

of the experimental instructions, which closely followed the information Kessler, Low, and Sullivan (2019) and Low (2024) provided in their experiments:

> *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.*

Several studies used the incentivized resume rating method to elicit preferences for attributes that are difficult to elicit from revealed preferences. For instance, 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 providing 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 providing booking options for doctors based on patients' choices.

4 Data

4.1 Variable Constructions

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 whole student pool. To do so, we use the latest placement records of graduates from each participating high schools, assign the selectivity index to each colleges in the records, rank the placements by the selectivity index, and assign the selectivity index to each students within a given high school ranked by their academic abilities in a given subject (reading, mathematics, English, and total). The selectivity index of each program was obtained from the list prepared by Kawaijuku, one of the most popular commercial college exam preparation companies in Japan.[11] As public colleges require larger number of subjects in the entrance exam, we follow Araki, Kawaguchi, and Onozuka (2016) and add 0.5 to the index of public colleges (5 in the raw index).

Behavioral Traits We elicited students' behavioral traits through post-experimental questionnaire: confidence in reading, mathematics, and English, competitiveness, and risk-taking in 5-point Likert scale with 3 being neutral.[12] For confidence questions, we asked the degree to which their recent exam scores they entered on the previous page represented their ability correctly. We convert these 5-point scales into [-1,1] intervals for ease of interpretation, with 0 being neutral.

11. The link to the list: https://www.keinet.ne.jp/exam/ranking/index.html (accessed December 18, 2024).

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

4.2 Summary Statistics

Table 1: Summary Statistics of Students in the Final Sample

	Female (N=311)		Male (N=298)		Difference (M – F)	
	Mean	SD	Mean	SD	Mean	SE
Panel A: Demographics						
Mother bachelor or above	0.59	0.49	0.56	0.50	-0.03	0.04
Father bachelor or above	0.80	0.40	0.77	0.42	-0.03	0.04
Both bachelor or above	0.50	0.50	0.48	0.50	-0.02	0.04
Extra schooling (no. days/week)	0.94	1.06	0.92	1.04	-0.02	0.09
Panel B: Academic abilities (population z-score)						
Reading score	0.82	0.74	0.59	0.75	-0.23***	0.06
Math score	0.65	0.75	0.77	0.75	0.12**	0.06
English score	0.82	0.69	0.58	0.78	-0.25***	0.06
Total score	0.71	0.74	0.61	0.76	-0.11*	0.06
Panel C: Behavioral traits						
Reading confidence [-1,1]	-0.01	0.39	0.05	0.49	0.06	0.04
Math confidence [-1,1]	-0.02	0.39	0.18	0.50	0.20***	0.04
English confidence [-1,1]	-0.00	0.40	0.10	0.45	0.10***	0.03
Competitiveness [-1,1]	0.03	0.67	0.11	0.68	0.08	0.05
Risk-taking [-1,1]	-0.36	0.61	-0.18	0.70	0.18***	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. Significance levels: * 10%, ** 5%, and *** 1%.

Table 1 presents summary statistics of 311 female and 298 male students in the final sample (609 students in total) as well as their differences. Panel A presents students' demographics and shows that there are no differences in parents' education level 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 are better at reading and English than male students, while male students are better at mathematics, and female students are slightly better in total score.[13] Yet, as expected, students in our sample perform about 0.6 to 0.7 standard deviations better than average Japanese high school students. Panel C presents students' behavioral traits and shows that male students are more confident in their mathematics and English abilities and less risk-averse than female students, consistent with the literature on gender differences in preferences (Croson and Gneezy 2009). Although statistically insignificant, the direction of the gender differences in competitiveness is also in line with the literature, where male students are more competitive than female students.

13. Appendix Figure A1 shows the distribution of abilities for female and male students.

5 Effect of the Gender Ratio at Colleges on High School 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 is represented with 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 the 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 via OLS for ease of interpretation and decomposition. Yet, we present the results with logit for a robustness check.

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

Denote the reduced form of equation 2 that we estimate via OLS as follows:

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

Where C^r_{ij} is an indicator variable equals 1 if student i chooses the right program in choice pair j, $FShare_{jd}$ is female student share of program d in pair j, $W_{jd} \equiv X_{jd} \setminus \{FShare_{jd}\}$ is a vector of attributes of program d in pair j other than the female student share, and α^g is intercept for the right program.

Now, replace C^r_{ij} with the four reasons:

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

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

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

$$C^r_{ij} = \phi^g + \eta^g(FRatio_{jr} - FRatio_{jl}) + (W_{jr} - W_{jl})'\kappa^g + \sum_{m=1}^{4} \theta^{m,g} M^m_{ij} + \upsilon_{ij} \tag{5}$$

Gelbach (2016) shows that:

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

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

When we discretize the female student share, as explained later, we perform this decomposition for each bin of the share.

5.2 Preferences for Gender Ratio

Table 2: Preferences for Program Attributes

Sample:	Female			Male		
Estimation:	OLS		Logit (AME)	OLS		Logit (AME)
Outcome:	Choose the program (0/1)					
	(1)	(2)	(3)	(4)	(5)	(6)
STEM	-0.079***	-0.079***	-0.078***	0.038**	0.038**	0.038**
	(0.017)	(0.017)	(0.018)	(0.017)	(0.017)	(0.017)
Female student share	0.090***	0.821***	0.819***	-0.036	0.818***	0.815***
	(0.022)	(0.088)	(0.097)	(0.024)	(0.089)	(0.097)
Female student share squared		-0.745***	-0.742***		-0.875***	-0.872***
		(0.088)	(0.095)		(0.090)	(0.099)
Selectivity index (population SD)	0.046***	0.047***	0.046***	0.075***	0.075***	0.074***
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.012)
Cohort size/100	0.003	0.003	0.003	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***	0.060***	0.033**	0.034**	0.034**
	(0.014)	(0.014)	(0.014)	(0.013)	(0.013)	(0.013)
Club participation rate	0.193***	0.196***	0.196***	0.122***	0.130***	0.130***
	(0.042)	(0.042)	(0.043)	(0.041)	(0.040)	(0.041)
Constant	0.503***	0.503***		0.488***	0.487***	
	(0.007)	(0.007)		(0.008)	(0.008)	
Adj. R-squared	0.036	0.054		0.021	0.045	
No. observations	4649	4649	4649	4451	4451	4451
No. students	310	310	310	297	297	297

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

Table 2 presents the coefficient estimates on the program attributes with choice as the dependent variable.[14] Columns 1 to 3 present estimates for female students, and columns 4 to 6 present estimates for male students. 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 choose them, consistent with the literature. Second, both female and male students prefer programs with a higher selectivity

14. Appendix Table A3 presents the same specifications but with indicator variables for the four reasons in place of choice as the outcome variables.

index, but male students prefer it more: a 1 standard deviation increase in the selectivity index increases female students' choice probability of the program by 4.6 to 4.7 percentage points, while it increases male students' choice probability by 7.5 percentage points.

Third, however, students also prefer the social aspects of the programs. Both female and male students prefer programs with an international exchange program, but female students prefer it more: having an international exchange program increases female students' choice probability by 5.9 to 6.0 percentage points, while it increases male students' choice probability by 3.3 to 3.4 percentage points. Club participation rates also affect students' choice probability: a 10% increase in club participation rate increases female students' choice probability by 19.3 to 19.6 percentage points, and male students' choice probability by 12.2 to 13.0 percentage points. In addition, male students slightly prefer a larger program while female students do not: an increase of a program size by 100 students increases male students' choice probability by 0.6 percentage points. Logit (average marginal effects or AME) estimates in columns 3 and 6 essentially present the same results as OLS.

However, our main variable of interest, female student share, is highly non-linear for both female and male students: the coefficient estimates on the squared female student share are statistically and quantitatively highly significant for both female and male students, as shown in columns 2 and 4. Hence, we discretize the female student share into 19 equally-spaced bins with an increment of 5 percentage points and re-estimate equation 2.

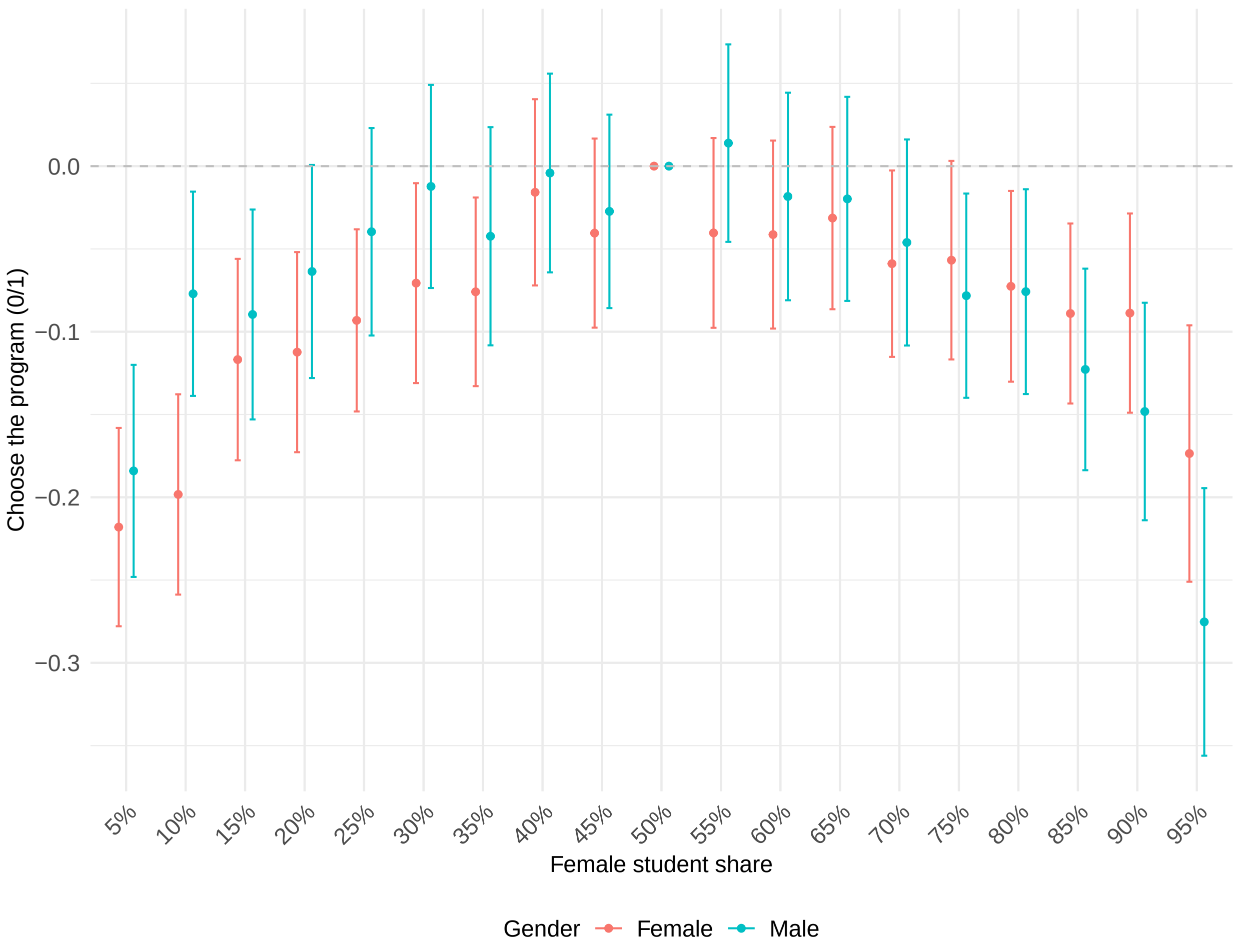
Figure 3 plots the coefficient estimates and the 95% confidence intervals of female student share discretized into 19 equally spaced bins, with 50% as the baseline, separately for female students (red) and male students (blue). The figure shows that the gender ratio does affect both female and male students' college choices. Specifically, both female and male students prefer programs where gender ratio is balanced over where there are fewer students of own gender: compared to a program where 50% of students are female, female students are about 22-20 percentage points less likely to choose a program where only 5-10% of students are female, and male students are about 27-15 percentage points less likely to choose a program where only 5-10% of students are male (95-90% of students are female).

Interestingly, both female and male students also prefer a program where the gender ratio is balanced over a program where there are too many students of own gender: compared to a program where 50% of students are female, female students are about 17 to 9 percentage points less likely to choose a program where 90-95% of students are female, and male students are about 18 to 7.5 percentage points less likely to choose a program where 90-95% of students are male (10-5% of students are female). However, if anything, both female and male students prefer a program where most students are their own gender over a program where only a few students are their own gender.

5.3 Reasons for the Preferences

To investigate why the gender ratio affects students' college choices and why they avoid being a minority and a majority, Figure 4 plots $\hat{\theta}^{m,g}\hat{\xi}^{m,g}$ $(m = 1, 2, 3, 4)$ and $\hat{\eta}^g$ from equation 6 estimated for each of the 19 bins separately for female (Panel A) and male (Panel B) students. The red area

Figure 3: Preferences for 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.

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 starkly differs when most students in the program are male or female. Specifically, when most students in the program are male, the gender ratio affects female students' program choices primarily due to the expected difficulty in fitting in. On the other hand, when most students in the program are female, no single reason explains why the gender ratio affects female students' program choices: all the four reasons – concerns for fitting in, finding a role model, doing well in studies, and finding a career to pursue – as well as unexplained factors explain why.

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

Figure 4: Decomposition of Preferences for Gender Ratio

Notes: This figure plots $\hat{\theta}^{m,g}\hat{\xi}^{m,g}$ $(m = 1, 2, 3, 4)$ and $\hat{\eta}^g$ 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.

expected difficulty in fitting in. When most students are male, on the other hand, all four reasons, as well as reasons other than these four, explain why.

5.4 Heterogeneity of the Preferences

STEM vs. Non-STEM Programs Appendix Figure A2 plots coefficient estimates and their 95% confidence intervals on discretized female student shares separately for STEM program (orange) and non-STEM program (green) for female (Panel A) and male (Panel B) students. The F statistic on the top left of each panel shows the differences in the estimates between STEM and non-STEM programs. The figure shows no significant heterogeneity between STEM and non-STEM programs for female or male students. This figure suggests that the gender ratio affects students' program choices regardless of whether it is STEM or non-STEM.

Top 50% vs. Bottom 50% Academic Abilities Appendix Figure A3 plots coefficient estimates and their 95% confidence intervals on discretized female student shares separately for students in the top 50% (orange) and the bottom 50% (orange) of mathematics abilities. The F statistic on the top left of each panel shows the differences in the estimates between the two groups of students. Panel A plots the estimates for female students, and Panel B plots the estimates for male students. The figure shows no significant heterogeneity between the two groups of students, both female and male.

Figure A4 plots the equivalent estimates but for reading abilities and shows a marginally statistically significant difference (10%) between the top 50% and the bottom 50% reading ability male students, but not female students. Specifically, male students in the top 50% of reading abilities are more likely to choose male-majority programs than male students in the bottom 50% of reading abilities. Still, male students with top 50% reading abilities are less likely to choose programs where most students are female. In summary, these figures still suggest that the gender ratio affects students' program choices regardless of their academic abilities.

High vs. Low SES Households Appendix Figure A5 plots coefficient estimates and their 95% confidence intervals on discretized female student shares separately for students from households with high socioeconomic status (SES) (orange) and with low socioeconomic status (green). The F statistic on the top left of each panel shows the differences in the estimates between the two groups of students. We proxy the SES by parents' education level: we classify students whose mother and father have at least bachelor's degree as high SES and students whose mother or father does not have bachelor's degree as low SES.[15] The figure shows no significant heterogeneity by household SES, suggesting that the gender ratio affects students' program choices regardless of students' household SES.

To summarize, there is surprisingly little heterogeneity for preferences for the gender ratio in terms of program type, students' academic ability level, and students' household SES status. Preferences for gender balance exist among STEM and non-STEM programs and students with different backgrounds.

6 Contribution to the Gender Gap in STEM Programs

To estimate the extent to which the preferences for the gender ratio and the low female student share in STEM programs contribute to the gender gap in these programs, we conduct a counterfactual analysis. We first estimate individual-specific preferences for each program attribute. We then use these individual parameters to predict STEM choice probabilities in the actual scenario where the gender ratio in STEM programs is set to students' median belief level and in the counterfactual scenario where the gender ratio in STEM programs is balanced. The gender ratio in non-STEM programs is set to the students' median belief level between the two scenarios. We then use the STEM choice probabilities under these two scenarios to conduct the analysis.

A caveat with this counterfactual exercise is that we do not take into account for other constraints, and assume that students are free to choose their programs solely based on their preferences. In reality, students' actual choices can be constrained by other factors such as financial considerations, parents' and teachers' suggestions, and other factors that we do not include in the program attributes in the experiments. Nevertheless, the comparison between the actual and the counterfactual scenarios can still be valid as those factors are kept constant between the two scenarios.

15. There are 266 students classified as those from household with high SES and 278 students classified as those from household with low SES. 65 students are unclassified because they answered they did not know either their mother's or father's education (or both).

6.1 Econometric Framework

Estimation of Individual-Level Preferences To estimate individual students' preferences for program attributes, we use mixed logit. Rewrite equation 2 as follows:

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

where P^g_{ijd} is student i of gender g's choice probability of program $d \in \{r, l\}$ with attributes X in pair j.

Now, assume β^g_i is a random variable with density $f(\beta^g|\theta^g)$. Then the choice probability can be written in a mixed logit form:

$$P^g_{ijd} = \int \frac{\exp(X'_{jd}\beta^g_i)}{\sum_{k\in\{r,l\}}\exp(X'_{jk}\beta^g_i)} f(\beta^g|\theta^g)d\beta^g \tag{8}$$

where we allow arbitrary correlations among elements in β^g_i.

We draw individual-specific parameters β^g_i from the estimates as a weighted average of the conditional distribution of β^g_i, $h(\beta^g|y, X, g, \theta^g)$, following Train (2009), where y is a vector of choices across the 15 pairs. We assume f to be a triangular distribution to prevent outliers from affecting the estimates following Kremer et al. (2011) and León and Miguel (2017). However, a normal distribution gives essentially the same results. We use a quadratic functional form for female student share to keep the number of model parameters adequate for the sample size.[16]

Prediction of Actual and Counterfactual STEM Choice Probabilities We predict the actual STEM choice probability with logit using individual-specific parameters $\hat{\beta}_i$:

$$\hat{P}^{STEM,Actual}_i = \frac{\exp(\bar{X}'_{STEM}\hat{\beta}_i)}{\sum_{k\in\{STEM,non\text{-}STEM\}}\exp(\bar{X}'_k\hat{\beta}_i)} \tag{9}$$

Where $\bar{X}_d$ is a vector of the average values of characteristics a program can take, except the female student share, which is set to students' median belief (25% for STEM, 55% for non-STEM; see Appendix Table A2).

Now, we predict the STEM choice probability with $\tilde{X}_d$, where we replace the female student share in STEM and non-STEM in $\bar{X}_d$ with 50%, so that i's choice is not constrained by the female student share:

$$\hat{P}^{STEM,CF}_i = \frac{\exp(\tilde{X}'_{STEM}\hat{\beta}_i)}{\sum_{k\in\{STEM,non\text{-}STEM\}}\exp(\tilde{X}'_k\hat{\beta}_i)} \tag{10}$$

Calculating Actual and Counterfactual Gender Balances in STEM We calculate the share of female and male students in STEM programs in the actual and counterfactual scenarios

16. The tipping point of female student share is 0.55 for female students and 0.47 for male students with quadratic female student share, which are not very different from the tipping points in Figure 3.

by taking the average of the STEM choice probabilities under each scenario. We apply a capacity constraint in calculating the student shares because, in reality, not everyone can be admitted to the program they want. The capacity constraint is applied to both actual and counterfactual scenarios by adding a shadow price term to equations 9 and 10. We define the shadow price as follows and subtract it from the indirect utility of choosing STEM but ignoring the error term, $X'_{STEM}\beta_i$:

$$\text{Shadow price}_i = \lambda(\overline{Selectivity} - \text{math score}_i) \tag{11}$$

where math score$_i$ is student i's math score, and $\overline{Selectivity}$ is the average selectivity index a hypothetical program can take (63.75). Thus, assuming λ is positive, this formulation makes it more costly for students with low mathematics ability and less costly for students with high mathematics ability to enter a STEM program. We can think of this shadow price as the effort cost to prepare for the entrance exam, the effort cost to catch up to the class once entered, or the risk of not getting 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 1/3. We choose 1/3 for convention as there are roughly three program categories – Humanities, Social Sciences, and STEM. This is equivalent to adjusting the STEM-alternative constant term (see, for example, Train 2009, Section 2.8). Unlike the alternative-specific constant, however, we vary it with individual mathematics ability.

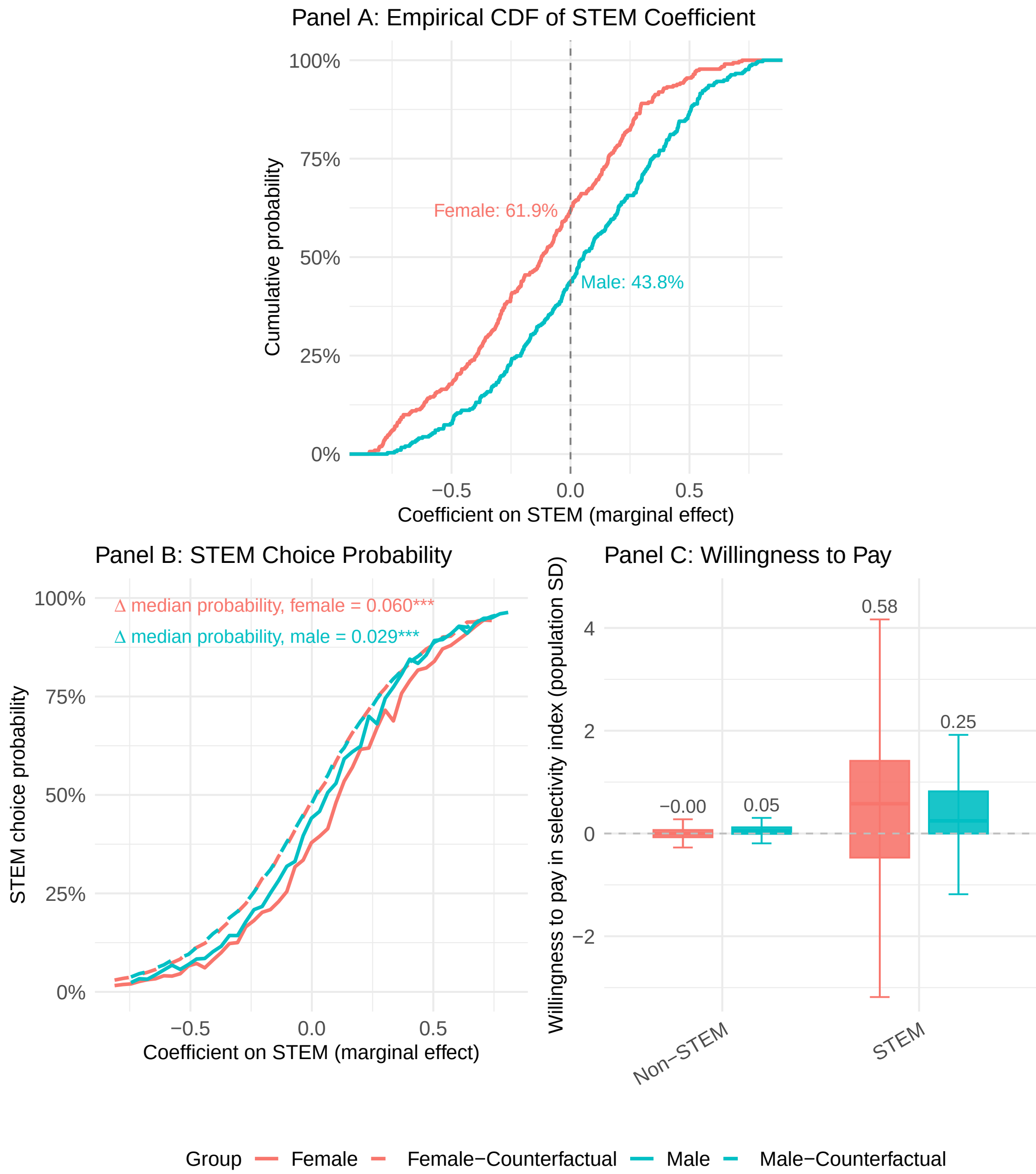
6.2 STEM Choice Probabilities

Panel A of Figure 5 plots empirical cumulative distribution functions (CDFs) of STEM coefficient for female (red) and male (blue) students. The figure shows that 38% of female students and 56% of male students have a positive STEM coefficient. This means that in the absence of any constraints, 38% of female students and 56% of male students prefer STEM over non-STEM programs.

However, the low female student share in STEM programs constrains their choices, especially female students' choices. Panel B of Figure 5 plots the actual (solid line) and counterfactual (dashed line) STEM choice probabilities for female (red) and male (blue) students. The texts on the top left show the median differences in the choice probabilities under the two scenarios. The figure shows that in the counterfactual scenario where the student gender ratio is balanced, students whose coefficient estimate on STEM is 0 are indifferent between STEM and non-STEM programs, meaning there is little constraint on their choice, as expected. When we look at the actual scenario, however, female students are less likely to choose STEM by 6 percentage points or by 15.7%. The low female student share is also pulling down male students' choice probability by 2.9 percentage points or by 5.2%.

To quantify the utility cost of low female student share in STEM programs, Panel C of Figure 5 plots the willingness to pay for moving from actual to the counterfactual scenarios measured by the selectivity index (in population standard deviation). The whisker indicates the range between the 5th and the 95th percentiles, and the number on the top indicates the median. The willingness to pay is obtained based on the indirect utility function of equation 1. Specifically, it is the additional

Figure 5: STEM Choice Probabilities Under Actual and Counterfactual Scenarios



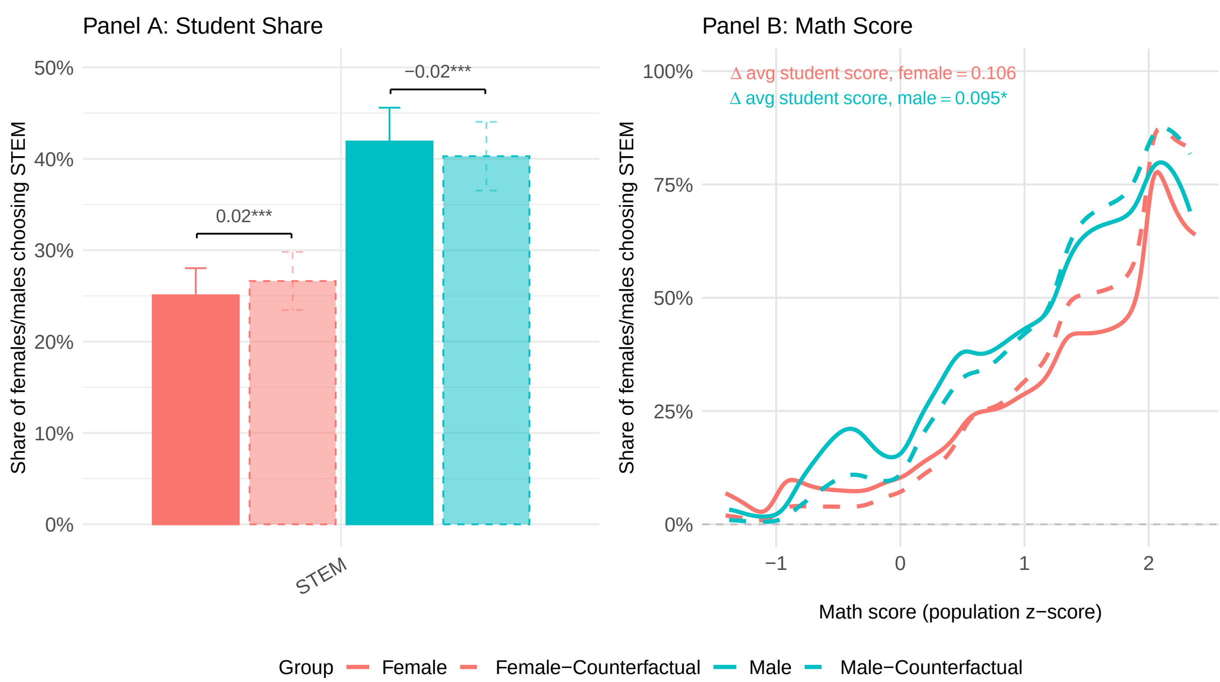
Notes: Panel A plots empirical cumulative distribution functions (CDFs) of STEM coefficient for female (red) and male (blue) students. Panel B plots the actual (solid line) and counterfactual (dashed line) STEM choice probabilities for female (red) and male (blue) students. The texts on the top left show the median differences in the choice probabilities under the two scenarios. The significance level is based on the Wilcoxon rank-sum test. Panel C plots the willingness to pay for moving from actual to the counterfactual scenarios measured by the selectivity index (in population standard deviation). The whisker indicates the range between the 5th and the 95th percentiles, and the number on the top indicates the median. Significance levels: * 10%, ** 5%, and *** 1%.

selectivity index that equates individual i's indirect utilities of female student shares of 25% and 50%. While the selectivity index is not a monetary measure, it is associated with the career success that many students care about.[17]

While it is a noisy measure, neither female nor male students are willing to sacrifice the selectivity index in non-STEM programs by a significant margin to achieve gender balance as expected, because the female student share in non-STEM programs is already above parity. However, in STEM programs where female student share is still low, female students are willing to sacrifice the selectivity index by 0.58 standard deviations. Male students are also willing to sacrifice the selectivity index, but less than female students are: they are willing to sacrifice by 0.25 standard deviation. Therefore, the current low female student share in STEM programs imposes utility costs on students, but particularly so for female students.

6.3 Student Shares in STEM

Figure 6: Student Share and Abilities in STEM Under Actual and Counterfactual Scenarios



Notes: Panel A plots share of female (red) and male (blue) students who choose STEM programs under actual (solid) and counterfactual (dashed) scenarios, and the significance level for gender share differences is based on a t-test. Panel B plots the mathematics ability of female (red) and male (blue) students who choose STEM under the actual (solid line) and the counterfactual (dashed line) scenarios, and the texts on the top left show the mean differences in the average student quality under the two scenarios where the statistical significance is obtained with the Horvitz-Thompson estimator using STEM choice probability as the selection probabilities (Horvitz and Thompson 1952). Significance levels: * 10%, ** 5%, and *** 1%.

Panel A of Figure 6 plots the share of female (red) and male (blue) students who choose STEM programs under actual (solid) and counterfactual (dashed) scenarios. It shows that under the

17. We are not the first to use a non-monetary measure to estimate willingness to pay. For example, Gallen and Wasserman (2023) measures willingness to pay for having a same gender mentor by the probability one wants to be mentored by a person in their preferred occupation.

actual scenario, 25.1% of female students and 41.9% of male students choose STEM programs over non-STEM programs, with the gender gap of 16.8 percentage points. However, under the scenario where the gender ratio is balanced, the share of female students increases by 1.5 percentage points and the share of male students decreases by 1.6 percentage points. Thus, when the gender ratio is balanced, we achieve higher gender balance: 26.6% of female students and 40.3% of male students choose STEM programs over non-STEM programs, and the gender gap shrinks to 13.7 percentage points or narrows by 18.5% relative to the actual scenario. Although female students replace male students in STEM, we show in the next subsection that this substitution can actually improve the average student quality.

Note that the realized gender ratio does not reach 50:50 even under the counterfactual scenario. It is because the underlying preference for STEM is not 50:50 as shown in Panel B of Figure 5. This suggests that other factors are at play: the lack of STEM role models and the gender stereotypes associated with STEM, as discussed in the introduction.[18] Thus, low female student share in STEM is one of the factors that prevents female students from choosing STEM programs. Still, it is a factor that has a sizable effect on their choices.[19]

6.4 Student Abilities in STEM

One concern policymakers would have is that increasing the female student share may attract students less prepared for STEM, which may deteriorate the quality of students. To address this concern, Panel B plots the mathematics ability of female (red) and male (blue) students who choose STEM under the actual (solid line) and the counterfactual (dashed line) scenarios. The texts on the top left show the mean differences in the average student quality under the two scenarios.[20]

The Panel B of Figure 6 shows that about 7% of female students whose mathematics ability is 1 standard deviation below the population mean choose STEM programs under the actual scenario. The share increases roughly with their mathematics ability, and around 75% of female students whose mathematics ability is 2 standard deviations above the population mean choose STEM programs. Under the counterfactual scenario, female students who choose STEM programs drops further when their mathematics ability is 1 standard deviation below the population mean and increases slightly when their mathematics ability is 2 standard deviations above the population mean. More generally, the likelihood of choosing STEM programs has higher association with the mathematics abilities under the counterfactual scenario than under the actual scenario.

We see a similar pattern for male students who choose STEM programs between actual and counterfactual scenarios; however, the drop in the share of students with low mathematics abilities is more pronounced for male students. As a result, the change in the average student quality is

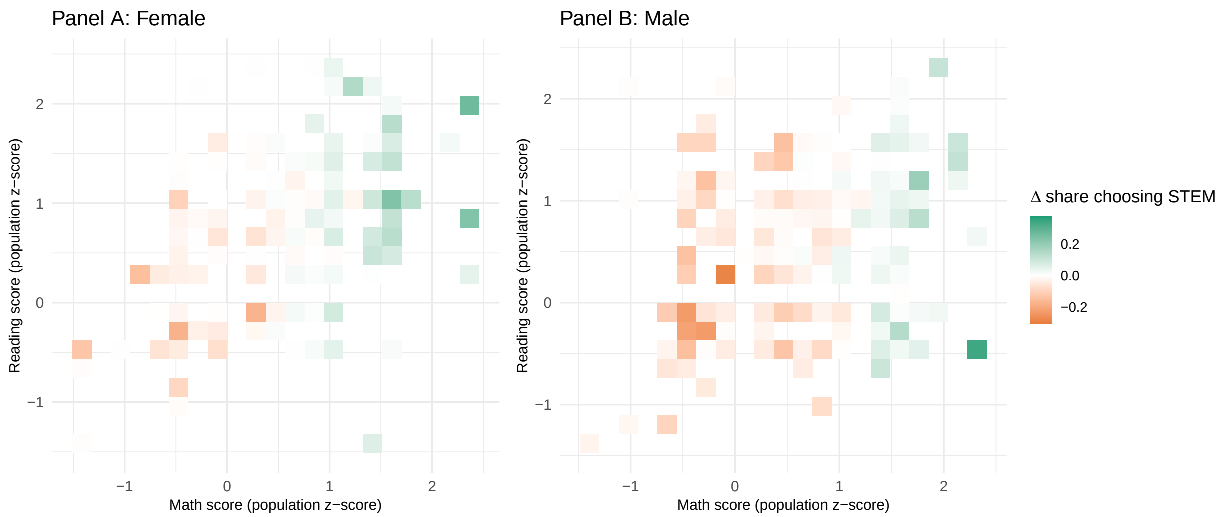
18. It is also possible that female students have comparative advantage in readings and thus more likely to choose non-STEM (Dekhtyar et al. 2018). However, it is likely due to misconception of STEM programs as verbal skills are necessary both in STEM and non-STEM.

19. Appendix Figure A6 presents the same plots as Figure 6 but without capacity constraint.

20. The statistical significance is obtained with the Horvitz-Thompson estimator using STEM choice probability as the selection probabilities (Horvitz and Thompson 1952).

marginally positive and statistically significant at 10% level.

Figure 7: Changes in Students' Abilities in Counterfactual Scenario



Notes: This figure plots a cross tabulation of mathematics (x-axis) and reading abilities (y-axis) of students who switch to STEM programs under the counterfactual scenario for females (Panel A) and males (Panel B).

To elaborate this point, Figure 7 plots a cross tabulation of mathematics (x-axis) and reading abilities (y-axis) of students who switch to STEM programs under the counterfactual scenario for females (Panel A) and males (Panel B).[21] The figure shows that under the counterfactual scenario, male students who excel in reading but do not do well in mathematics are more likely to choose non-STEM than under the actual scenario, leading to an improved talent allocation. Thus, on the contrary to the concern policymakers may have, the current low female student share may actually attract male students less prepared for STEM, and making STEM programs more balanced can attract students more prepared for STEM. On the other hand, we do not see this pattern for female students. Therefore, what happens behind the substitution in Panel A of Figure 6 is that female students with high mathematics abilities replacing male students with low mathematics but high reading abilities.

To summarize, the current low female student share in STEM programs constrains students' STEM choices, especially female students' choices. Removing this constraint can make STEM programs more gender balanced and can improve the allocation of talent.

7 Policy Implications

Although our results are based on a small fraction of high school students in Japan, and Japan differs from other countries in several ways, we can still draw some policy implications for affirmative action for women in STEM in higher education.

21. We see more changes along the 45 degree line because students who excel in mathematics also tend to excel in readings, especially female students. The correlation between mathematics and reading score is 0.481 for female students and 0.270 for male students.

First, a typical criticism of affirmative action relates to individual merit: it does not help underrepresented students. For example, Arcidiacono, Aucejo, and Hotz (2016) find that in the University of California campuses, an affirmative action for underrepresented minority (URM) students that essentially lowers the threshold for the admission criteria can actually have adverse effects on those students in STEM programs: they are more likely to dropout and took longer to graduate because they are less academically prepared.[22] Elaborating on this result, Arcidiacono and Lovenheim (2016) argue that while affirmative action can increase the quality of education that URM students can receive, it can lower the fit between those students' academic preparedness and the colleges' academic requirements. Our results suggest that this concern does not apply to policy targeting female students, as female students who would be affected by these policies are as academically prepared for STEM as currently enrolled male students.[23]

Another criticism of affirmative action relates to the negative spillover effects at the aggregate level: it can be discriminatory against non-minority groups. This is likely the reason behind the ban of affirmative action at colleges in the US Supreme Court in 2023 (see, for example, Mangan 2023) and sentiments of some US high school students (The Learning Network 2023). Although our results show a slight improvement in the quality of STEM students when we balance the gender ratio, we remain cautious about this point as some male students are indeed negatively affected.

Nevertheless, our findings indicate that some form of affirmative action may be justified due to the market failure: students are unlikely to take into account the social issues associated with women's underrepresentation in STEM, as discussed in the introduction, when choosing their college programs. Specifically, although banned in the US, implementing a gender quota that ensures a certain percentage of incoming students are female would probably outweigh the costs. Such a quota can generate short-run and medium-run effects. First, because the gender quota directly affects X in students' indirect utility in equation 1, it can reduce a barrier for female students pursuing STEM in the short term. Second, increasing the number of female students in STEM programs can shift societal perceptions in the medium term, which in turn impacts β in students' indirect utility in equation 1.

However, given the second criticism, the quota must be carefully designed to minimize negative spillover effects. Additionally, since students admitted through quotas may face stigma as being less capable than those admitted outside the quota (Yokoyama et al. 2024), and since the primary reason female students avoid being in the minority is the difficulty fitting in the environment, each program must implement measures to reduce such stigma.

8 Conclusion

Female students are less likely to pursue STEM fields at colleges, despite negligible gender differences in mathematics and sciences. Having fewer women in STEM programs at colleges causes several

22. URM groups in this context refer to African Americans, Hispanics, and Native Americans.

23. There is another evidence that an affirmative action – a gender quota for candidates in a Swedish political party – improved the competence of politicians (Besley et al. 2017).

social issues. One potential reason that has received less attention in the literature is that STEM programs are predominantly male-dominated, which may discourage women from pursuing STEM. In this paper, we examine whether the gender ratio affects students' college choices and to what extent it contributes to the gender gap in STEM programs.

We first show using administrative data that covers the universe of college programs in Japan that colleges have become more gender balanced both in STEM and non-STEM programs, especially in programs where students are supposed to face weaker trade-off between the gender ratios and other attributes. We then use an incentivized discrete choice experiment and show that the gender ratio at colleges does affect female but also male students' college choices. In particular, both female and male students dislike being a minority and being a majority, and the main reason they dislike being a minority is the expected difficulty of fitting into the program.

A counterfactual analysis shows that low female student share in STEM programs reduces the probability that a female student chooses a STEM program by 6 percentage points or 15.7%, and they incur the utility cost equivalent to 0.58 standard deviations of program selectivity. Removing this constraint would increase the gender balance in STEM programs and replace male students with low mathematics but high reading abilities with female students with high mathematics abilities. Taken together, the gender ratio at colleges is indeed an important determinant for students' college choices, and making STEM programs more gender balanced can help close the gender gap in STEM and address social issues that come from it, and can improve the allocation of talent.

A natural next question would be how to make STEM programs more gender balanced. As discussed in Section 7, a gender quota would probably bring net benefits, and exploring the effects of the quota or its optimal design would be promising. While answering these questions would be difficult with our data, future work could implement some interventions or exploit the introduction of gender quotas in certain college programs. While we should be cautious not to overgeneralize our findings, our paper highlights the importance of student gender ratios at colleges in closing the gender gap in STEM.

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.

Arcidiacono, Peter, Esteban M. Aucejo, and V. Joseph Hotz. 2016. "University Differences in the Graduation of Minorities in STEM Fields: Evidence from California." *American Economic Review* 106 (3): 525–562.

Arcidiacono, Peter, and Michael Lovenheim. 2016. "Affirmative Action and the Quality-Fit Trade-Off." *Journal of Economic Literature* 54 (1): 3–51.

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.

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.

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.

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.

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

Carlana, Michela, and Lucia Corno. 2022. *Parents and Peers: Gender Stereotypes in the Field of Study.* 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.

Cullen, Zoë, and Ricardo Perez-Truglia. 2023. "The Old Boys' Club: Schmoozing and the Gender Gap." *American Economic Review* 113 (7): 1703–1740.

Dekhtyar, Serhiy, Daniela Weber, Jonas Helgertz, and Agneta Herlitz. 2018. "Sex Differences in Academic Strengths Contribute to Gender Segregation in Education and Occupation: A Longitudinal Examination of 167,776 Individuals." *Intelligence* 67:84–92.

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.

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

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.

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.

Goodman, Roger, and Chinami Oka. 2021. "The Hensachi: Its Dominant Role in University Rankings." In *Handbook of Higher Education in Japan,* edited by Paul Snowden, 133–142. Amsterdam, The Netherlands: Amsterdam University Press.

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.

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

Horvitz, Daniel G., and Donovan J. Thompson. 1952. "A Generalization of Sampling Without Replacement from a Finite Universe." *Journal of the American Statistical Association* 47 (260): 663–685.

Kambayashi, Ryo, and Takao Kato. 2017. "Long-Term Employment and Job Security over the Past 25 Years: A Comparative Study of Japan and the United States." *ILR Review* 70 (2): 359–394.

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.

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.

Kremer, Michael, Jessica Leino, Edward Miguel, and Alix Peterson Zwane. 2011. "Spring Cleaning: Rural Water Impacts, Valuation, and Property Rights Institutions." *The Quarterly Journal of Economics* 126 (1): 145–205.

León, Gianmarco, and Edward Miguel. 2017. "Risky Transportation Choices and the Value of a Statistical Life." *American Economic Journal: Applied Economics* 9 (1): 202–228.

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. *Closing the Gender Gap in STEM: Evidence from a Curriculum Reform in Computer Science.* 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.

Mangan, Dan. 2023. "Supreme Court Rejects Affirmative Action at Colleges as Unconstitutional." *CNBC: Politics.*

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

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

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.

Obunsha. 2024. "Colleges Are on the Rise, Junior Colleges Are Declining; Undergraduate Enrollment Has Fallen for the First Time in Ten Years [Original in Japanese]." *Educational Information Center Article [Original in Japanese].*

OECD. 2018. "Indicator B5 Who Is Expected to Graduate from Tertiary Education?" In *Education at a Glance 2018: OECD Indicators,* 206–216. Paris, France: OECD Publishing.

———. 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.

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.

The Learning Network. 2023. "What Students Are Saying About the End of Race-Based Affirmative Action in College Admissions." *The New York Times: Current Events Conversation.*

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.

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.

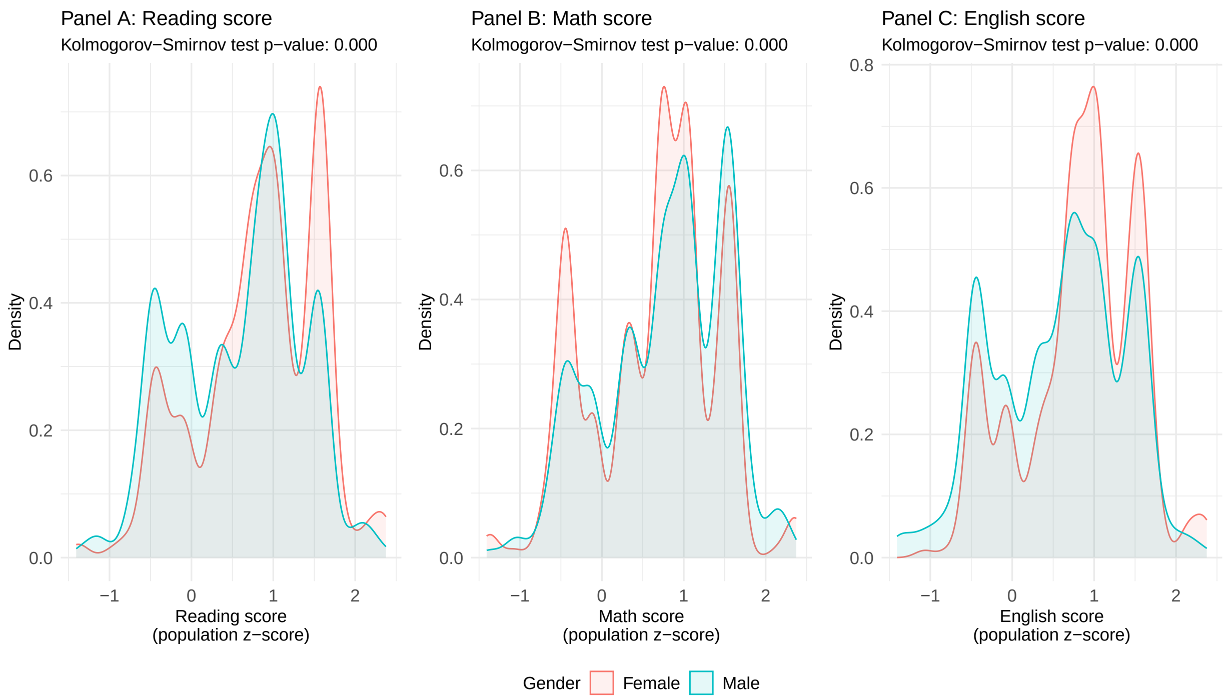
———. 2021. "Human Capital Investments and Expectations about Career and Family." *Journal of Political Economy* 129 (5): 1361–1424.

Yokoyama, Hiromi M., Yuko Ikkatai, Euan McKay, Atsushi Inoue, Azusa Minamizaki, and Kei Kano. 2024. "Can Affirmative Action Overcome STEM Gender Inequality in Japan? Expectations and Concerns." *Asia Pacific Business Review* 30 (3): 543–559.

Online Appendix

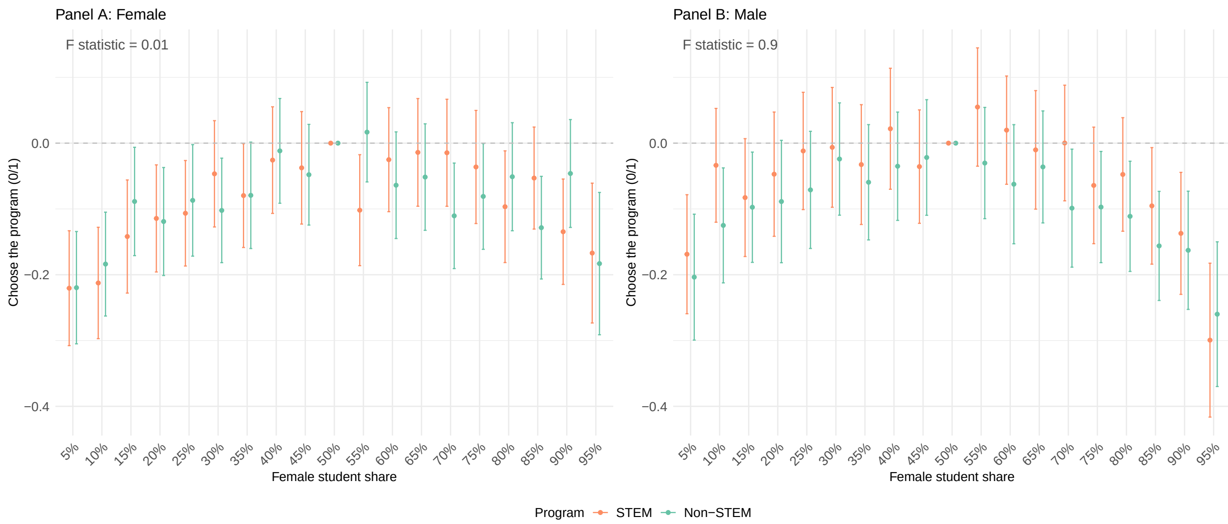
A Additional Figures and Tables

Figure A1: Distribution of Abilities



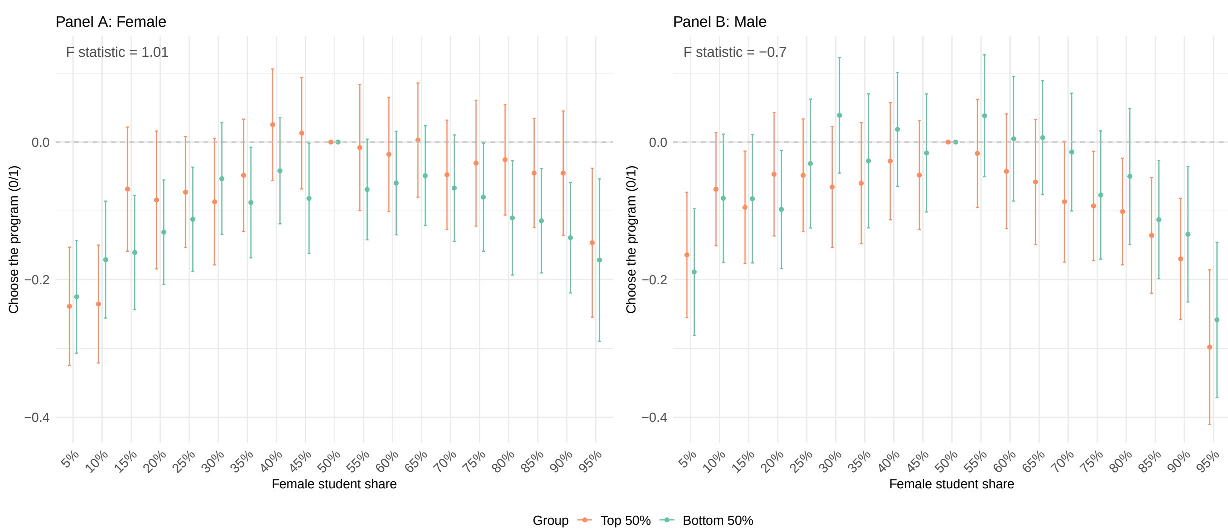
Notes: This figure presents the distribution of abilities of female and male students in reading (Panel A), mathematics (Panel B), and English (Panel C).

Figure A2: Preferences for Gender Ratio by STEM vs. Non-STEM Programs



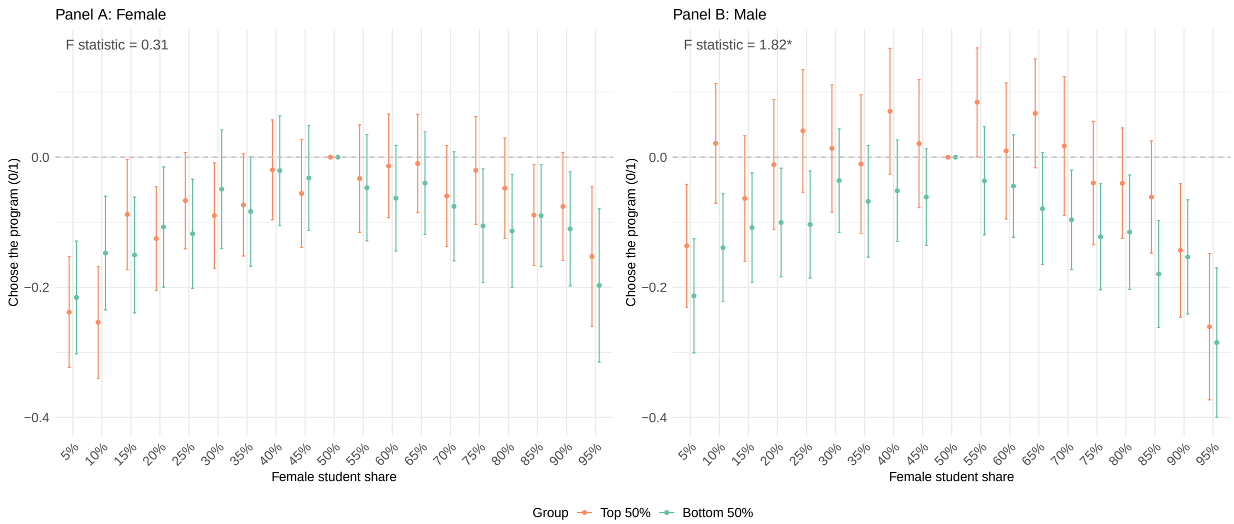
Notes: This figure plots coefficient estimates and their 95% confidence intervals on discretized female student shares separately for STEM program (orange) and non-STEM program (green) for female (Panel A) and male (Panel B) students. The F statistic on the top left of each panel shows the differences in the estimates between STEM and non-STEM programs. Significance levels: * 10%, ** 5%, and *** 1%.

Figure A3: Preferences for Gender Ratio by Top 50% vs. Bottom 50% Mathematics Abilities



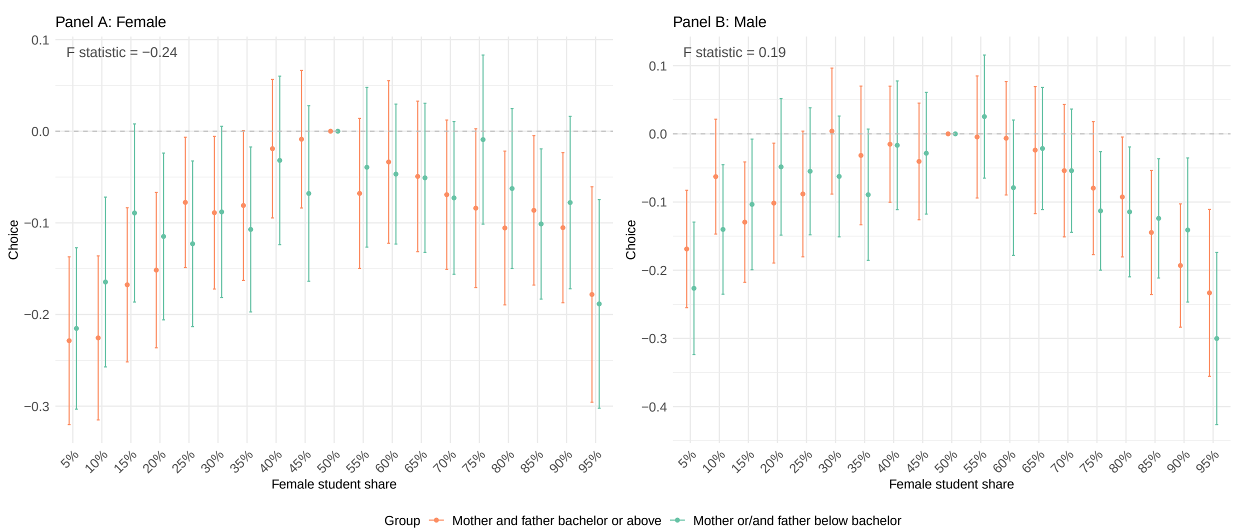
Notes: This figure plots coefficient estimates and their 95% confidence intervals on discretized female student shares separately for students in the top 50% (orange) and the bottom 50% (orange) of mathematics abilities. The F statistic on the top left of each panel shows the differences in the estimates between two groups of students. Significance levels: * 10%, ** 5%, and *** 1%.

Figure A4: Preferences for Gender Ratio by Top 50% vs. Bottom 50% Reading Abilities



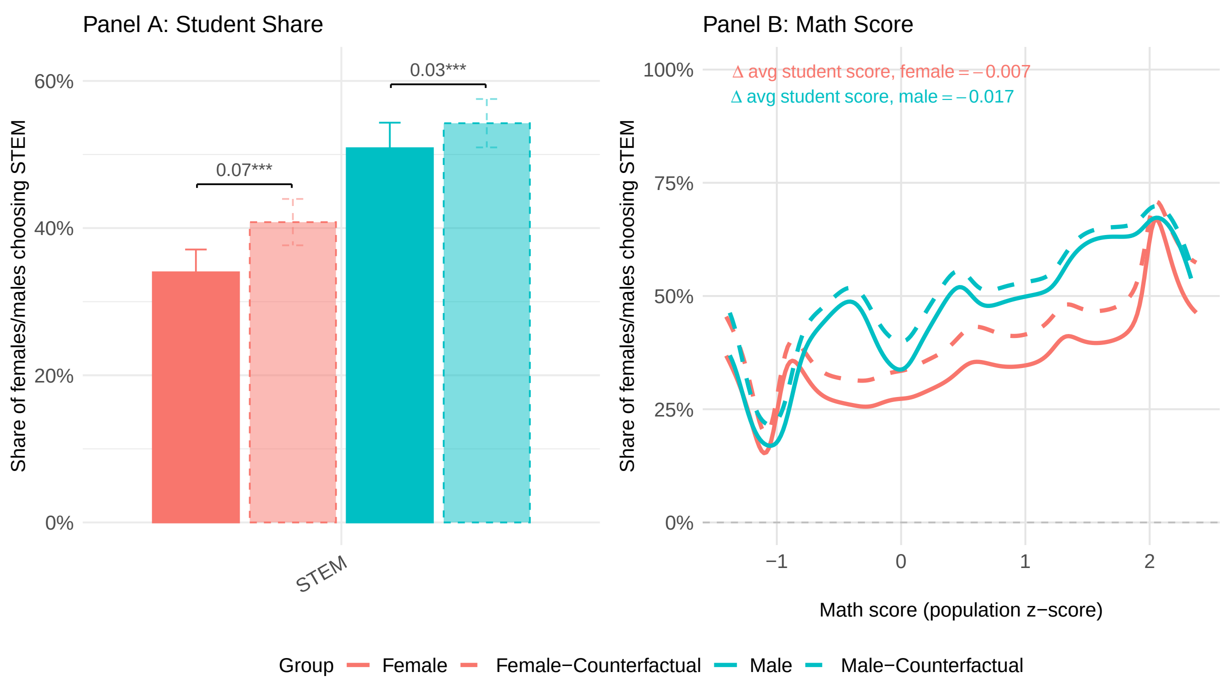
Notes: This figure plots coefficient estimates and their 95% confidence intervals on discretized female student shares separately for students in the top 50% (orange) and the bottom 50% (orange) of reading abilities. The F statistic on the top left of each panel shows the differences in the estimates between two groups of students. Significance levels: * 10%, ** 5%, and *** 1%.

Figure A5: Preferences for Gender Ratio by Household Socioeconomic Status



Notes: This figure plots coefficient estimates and their 95% confidence intervals on discretized female student shares separately for students from households with high socioeconomic status (SES) (orange) and with low socioeconomic status (green). The F statistic on the top left of each panel shows the differences in the estimates between two groups of students. Significance levels: * 10%, ** 5%, and *** 1%.

Figure A6: Student Share and Abilities in STEM Under Actual and Counterfactual Scenarios – No Capacity Constraint



Notes: Panel A plots share of female (red) and male (blue) students who choose STEM programs under actual (solid) and counterfactual (dashed) scenarios without capacity constraint, and the significance level for gender share differences is based on a t-test. Panel B plots the mathematics ability of female (red) and male (blue) students who choose STEM under the actual (solid line) and the counterfactual (dashed line) scenarios without capacity constraint, and the texts on the top left show the mean differences in the average student quality under the two scenarios where the statistical significance is obtained with the Horvitz-Thompson estimator using STEM choice probability as the selection probabilities (Horvitz and Thompson 1952). Significance levels: * 10%, ** 5%, and *** 1%.

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.

Table A2: Median Belief about Female Student Share in Each Program

	Females' belief (N=311)	Males' belief (N=298)	Actual
Non-STEM	55.0%	50.0%	45.2%
STEM	25.0%	25.0%	22.8%

Notes: This table presents female and male students' median beliefs about female student share in each program.

Table A3: 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%.

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]