Gender Gap in High-Stakes Exams: What Role for Exam Preparation?

Cécile Bonneau* Léa Dousset†

October 29, 2024

Abstract: Admission to elite higher education institutions often relies on competitive entrance examinations, which are presumed to objectively measure students' abilities. However, emerging evidence suggests a gender gap in those examinations, with women frequently underperforming in high-stakes and competitive environments. This paper corroborates these findings within a highly selected group of elite STEM program undergraduates in France. We extend our investigation beyond the effect of the high-stakes exam itself to consider the role of the learning environment for such high-stake exams on this gender performance gap. Our analysis reveals that heightened competitiveness in exam preparation exacerbates the gender performance gap in admission to the most selective programs. These insights enhance our understanding of the gender gap in access to selective STEM higher education, which is particularly important in explaining the persistent gender pay gap at the top of the income distribution.

JEL codes: I23, J16, J24.

Keywords: Higher Education, Gender, STEM Field, Selective Programs, High-Stakes Exams.

*Aalto University. Email: *cecile.bonneau@aalto.fi.* Professional address: Ekonominaukio 1, 02150 Espoo, Finland.

†Paris School of Economics (PSE) and Université Paris 1 Panthéon-Sorbonne. Email: *lea.dousset@psemail.eu.* Professional address: 48 Boulevard Jourdan, 75014 Paris, France.

1 Introduction

In numerous countries, admission to elite higher education institutions heavily relies on competitive entrance examinations (OECD, 2019). Rigorous preparation is typically required for taking those exams, which are perceived as cornerstones of meritocratic admissions policies, providing a fair and equitable assessment of students' abilities.

Yet, economic experimental literature has identified sources of gender gaps in these competitive entrance examinations, highlighting gender differences in attitudes towards competition. Women's performances and their willingness to compete have been shown to be lower than those of men (Gneezy et al., 2003; Niederle and Vesterlund, 2007), particularly in tasks stereotypically associated with males (Niederle and Vesterlund, 2011). These findings are corroborated in real-life settings, where female students underperform in high-stakes examinations compared to continuous assessment (Azmat et al., 2016; Montalbán and Sevilla, 2023). This phenomenon, to which we refer to as the *D-day exam effect*, detrimentally impacts female students' higher educational outcomes (Arenas and Calsamiglia, 2022). While much of the literature has focused on the impact of high-stakes exams on the gender performance gap, little is known about the effect of the learning environment for these high-stakes exams on such gender disparities.

In this paper, we confirm that female students underperform compared to their male counterparts in high-stakes exams. We contribute to the existing literature by showing that this trend persists even among a highly selected group of high achievers who opt for selective STEM studies and are presumably well-prepared for such exams. We also provide compelling evidence that an intensified competitive environment during exam preparation exacerbates the gender performance gap, in addition to the impact of the *D-day exam effect*.

Our study takes advantage of the specific context of Science, Technology, Engineering, and Mathematics (STEM) elite higher education programs in France. Admission to French elite STEM graduate schools (*Grandes Écoles scientifiques*) requires high school graduates to first enroll in STEM preparatory (prep) programs (*Classes Préparatoires aux Grandes Écoles*). The admission in these prep programs is already highly selective, based on high school academic performance.[1] Prep programs are highly intensive academic courses that span two to three years, meticulously designed to prepare students for the competitive national entrance examinations for admission into elite graduate schools. To explore how a higher level of competition in the learning environment impacts the performance of

[1]Only 2.5% of a given birth cohort pursue this higher educational path and in 2021, 55 percent of the students in these programs graduated from high school with highest honors (i.e., with a GPA of 16/20 or higher), compared to 9 percent of all students graduating from high school in France (DEPP, 2022).

female and male students differently, we leverage a unique characteristic of these programs which creates exogenous variation in the competitiveness of the learning environment that students face: tracking by ability during the second year of these programs. In many STEM prep programs, second-year students are divided between standard classes and "star" classes (*classes étoiles*), with the latter comprising high-performing students preparing for the most prestigious competitive exams. This effectively creates a tracking system for high-achievers, allowing teachers to tailor instruction to the appropriate level. Due to frequent peer-based rankings, star classes foster a more competitive and intellectually stimulating environment. This feature of elite undergraduate STEM programs generates varying competitive study environments within the same program — without changing the stakes of the exams themselves[2] —, which enables us to assess the gender performance difference in more competitive exam preparation settings.

Our study rely on novel comprehensive administrative data from the centralized admission process to elite STEM graduate schools (*Service des concours écoles d'ingénieurs*, SCEI), which includes applicant demographics, prep program enrollments, exam applications, results, students' preferences and admission outcomes for 165,450 applicants from 2015 to 2023. This is enriched with administrative data on students' prior academic achievements.[3] These administrative datasets provide information on achievement at the high school graduation exam and at the end of the prep program, during the admission process to STEM graduate schools. To open the black box of performance within prep programs, we conducted a large-scale data collection on grades obtained by students in 18 preparatory programs, covering 21,801 students. This within-prep-program data is crucial for observing the evolution of the gender gap in performance over time and for controlling for ability just before high-stakes exams. This data is also essential for our empirical strategy, which involves a regression discontinuity at the margin of admission to the top track.

With this extensive dataset, we first observe a puzzling fact: while female students self-select more into undergraduate STEM prep programs in terms of academic achievement,[4] they end up underrepresented in the most selective STEM graduate schools. The underrepresentation of women in the most selective STEM programs may arise from (i) differences in preferences, (ii) a gender performance gap during high-stakes entrance examinations, and/or (iii) a gender performance gap

[2]Students in both types of classes are eligible to take the same competitive entrance exams. In practice, students in star classes tend to choose more selective exams, while those in standard classes opt for somewhat less selective exams. Nevertheless, there is significant overlap in the exams chosen by students in both types of classes: 89 percent of students in standard classes and 98 percent of students in star classes take at least one highly selective entrance exams.

[3]*Direction de l'évaluation, de la prospective et de la performance*, DEPP.

[4]In our sample of STEM prep program students, 59% of female students achieved their high school graduation exam with the highest honors, compared to 47% of male students.

in exam preparation. Our detailed data analysis allows us to decompose the contribution of each of these factors. The results of this decomposition discount preferences as the primary cause and reveal that the gender gap in access to top programs is partly due to performance disparities on exam day (*D-Day exam effect*) but is primarily driven by a widening gender gap in performance during exam preparation. This finding leads us to further explore the role of the learning environment during exam preparation.

To analyze the gender differential impact of preparing exams in a more competitive environment, we compare gender gaps in performance in classes of varying competition level — created by the tracking system between *star* and *standard* classes —, while controlling for detailed previous academic achievement and program fixed effects.[5] Our results indicate that preparing exams in a more selective and competitive environment is significantly less advantageous for female students than for male students, contributing to the underrepresentation of women in the most prestigious STEM graduate schools and subsequently to the gender pay gap among alumni of these institutions. On average, the gender gap in the probability of gaining admission to the top 10% most selective STEM graduate schools is 3.3 percentage points larger in star classes than in standard classes, representing a gender gap that is more than twice as large. This finding holds *on average* when comparing the relative gender gap in performance in more or less selective classes within a program. We use the smaller subsample of students with within-prep program grades to show that this result remains true *at the margin* of star class admission through a regression discontinuity analysis: admission to a star class is shown to significantly enhance the performance of male students but has no impact on the performance of female students. These results are consistent across various definitions of school selectivity, whether based on the prior academic achievements of admitted students or applicants' revealed preferences (Avery et al., 2013). Regarding expected earnings one year post-graduation, the gender pay gap is between 50% and 80% (depending on the inclusion of bonuses) larger for star class alumni compared to standard class alumni.[6]

Related literature and contributions. This paper relates to several strand in the literature and make several contributions. Firstly, we add to the literature on the success of female students who decide to pursue STEM studies. There exists a substantial body of literature aimed at explaining the underrepresentation of women in STEM fields, synthetized in Kahn and Ginther (2018). However, less attention has been given to women who choose to pursue STEM studies and careers. A few papers

[5]Crucial for our identification, we check that admission to star classes is the same for men and women. This is indeed the case as conditional on performance, there is no gender gap in access to star classes.

[6]These figures are computed using aggregated earnings data at the program $\times$ cohort $\times$ gender.

delve into the determinants of female students persistence within STEM programs (Griffith, 2010; Canaan and Mouganie, 2023). This literature has shown that women are more likely to leave STEM studies (Ellis et al., 2016; Landaud and Maurin, 2020; Kugler et al., 2021) and even when they pursue STEM studies, there is evidence that female STEM graduates are also less likely to remain in STEM occupations (Beede et al., 2011; Delaney and Devereux, 2022). Our article enhances this literature by examining the determinants of women underepresentation in the most selective STEM programs, as opposed to average ones. This is crucial for several reasons. First, this underrepresentation likely contributes to the substantial gender pay gap observed at the top of the income distribution and among STEM professionals: in our sample, we observe a gender pay gap of €1,600 among STEM workers one year post-graduation. Adjusting for STEM program selectivity, the gap narrows to €750, indicating that roughly half of the gender pay disparity is due to unequal access to selective STEM programs.[7] Secondly, as the representation of women in these programs is very low, it offers significant room for increasing female representation. Lastly, while the general underrepresentation of women in STEM may be attributed to differences in preferences — challenging to address — the specific underrepresentation in top-tier programs might not solely be preference-driven. The highly detailed nature of our administrative data, including rarely available student preferences, allows us to examine the role of these preferences in detail.

Our article more directly relates to the literature on gender differences in competitive environments. Over the past twenty years, there has been growing literature on the under-performance of women in competitive environments (Gneezy et al., 2003; Croson and Gneezy, 2009; Gneezy et al., 2009; Niederle and Vesterlund, 2007, 2011). More recent papers have shown that the gender gap in willingness to compete is particularly large for young people (Flory et al., 2018) and among the most competitive individuals (Saccardo et al., 2018). However, most of these studies are based on laboratory experiments. Buser et al. (2014) show that a laboratory measure of competitiveness is associated with real-life study choices, and can explain a substantial part of the gender gap in these choices, making it all the more important to study the effect of competitiveness in real-life educational settings. Some recent studies observe gender differences in performance in real-life setting, using variations in the stake of the exams (Azmat et al., 2016; Montolio and Taberner, 2021), in the level of competitiveness (Morin, 2015; Iriberri and Rey-Biel, 2019), in the selectivity of high schools (Landaud et al., 2020), in the test-taking environment (Montalbán and Sevilla, 2023), or in chess field tournaments (De Sousa and Hollard, 2023). Most of these studies are based on primary or secondary

[7]There is an annual wage gap of around €3,300 between the top 10% most selective graduate schools and the next top-tier institutions in terms of selectivity. These figures are computed using aggregated earnings data at the program $\times$ cohort $\times$ gender.

education settings. There is an emerging literature on the under-performance of female students in college entrance exams. These studies highlight in different contexts that female students underperform in high-stakes exams compared to their performance in low-stakes or mock exams (Cai et al., 2019; Schlosser et al., 2019; Arenas and Calsamiglia, 2022), underperform more in entrance exams at the most selective universities (Jurajda and Münich, 2011), and are less strategic than men in their answering of multiple-choice questions (Pekkarinen, 2015). Our research focuses on a more selective context than the majority of existing studies, and on an examination that takes place after two or three years of higher education, rather than on entry into higher education. Closely related to us, Ors et al. (2013) and Parodi et al. (2022) examine the entrance examinations of the most selective French elite schools of management and administration, respectively. Their findings highlight an observable discrepancy in female student performance, noting that these students generally score lower on high-stakes entrance exams compared to lower-stakes tests such as end-of-high-school examinations. We complement this literature in three ways. First, our focus is on STEM schools where the underrepresentation of female students is markedly more pronounced than in management or administration schools. Second, thanks to the comprehensiveness of our administrative data, we can extend the scope of our analysis beyond the most selective institution in a field to include a broader range of graduate schools. More importantly, most studies use variation in the degree of competitiveness of the high-stakes exams, rather than that of the learning environment for preparing such exams. Our study adds to this literature by showing that the test preparation context is important in itself as well. This is made possible thanks to the rarely available, within-program data on grades, which enables the observation of the gender gap during exam preparation, not just during high-stakes exams.

Lastly, our research engages with the existing literature on tracking in education, reviewed in Betts (2011). This literature has shown that students in the higher track benefit from interaction with higher-performing peers (Card and Giuliano, 2016), but that low-performing students can also benefit from teachers who target their level (Duflo et al., 2011). Closely related to us, Landaud and Maurin (2022) shows that in the context of French preparatory programs, tracking can increase social gaps in access to elite schools. We complement this literature by studying the gender-heterogeneous effect of tracking.

The article proceeds as follows: Section 2 details the institutional context of the French higher education system, with a focus on undergraduate STEM preparatory programs. Section 3 outlines the data, analysis samples, and descriptive statistics. Section 4 presents a decomposition of the gender gap in admission to the best STEM graduate schools into several explaining factors. In Section 5, we examine the role of the selectivity and competitiveness of the learning environment on these gender disparities. Section 6 explores the role of peers. Section 7 concludes.

2 Institutional Background

French Higher Education System. The French higher education system is characterized by a significant degree of academic hierarchy. Upon successfully completing the high school graduation exam, known as the *Baccalauréat*, students who pursue higher education have three main tracks from which to choose: (i) a technical and vocational track, which accounted for approximately 30 percent of first-year students in the 2021-2022 academic year, (ii) a non-selective academic track represented by public universities,[8] which enrolled around 50 percent of first-year students, and (iii) a selective academic track composed of preparatory programs (*Classes Préparatoires aux Grandes Écoles (CPGE)*) and elite graduate schools (*Grandes Écoles*), which admitted 7 percent of first-year students. The coexistence of these two academic tracks is a distinctive feature of the French higher education system. The remaining 13 percent of first-year students are enrolled in other specialized programs, such as paramedical training or specialized schools. Diagram A1 displays a simplified version of the different tracks available to students in France. The STEM elite academic track, which is the focus of our paper, is highlighted in blue.

Preparatory Programs and Elite Graduate Schools. Given their limited number of available seats, access to preparatory (prep) programs is highly selective. Admission is largely based on academic performance, as assessed by grades and teachers' assessments from the students' junior and senior years of high school. Prep programs consistently attract the highest-achieving students. For instance, in the academic year 2016-2017, 43 percent of the students enrolled in prep programs passed the high school graduation exam with highest honors. In comparison, only 3 percent of students in the technical and vocational track and 8 percent of university students achieved this level of academic distinction (Bonneau et al., 2021).

Prep programs are highly intensive academic courses that span two to three years, meticulously designed to prepare students for the competitive national entrance examinations required for admission to elite graduate schools. These higher education programs are predominantly hosted by

[8]Until 2018, public universities were non-selective, and access was formally granted to anyone holding the high school graduation exam. In the case of oversubscribed universities and programs, random lotteries were drawn to select students. Since 2018, universities have been allowed to select their students based on their own criteria. However, most university programs remain undersubscribed and, therefore, not selective in practice: Bechichi et al. (2021) estimated that in 2018 and 2019, 84% of university programs are non-selective, in the sense that they refuse less than 5 percent of applicants.

renowned high schools for their academic excellence. The academic requirements of prep programs are notably demanding. For a two-year program, the national curriculum is closely aligned with the content that would be covered in a bachelor's degree in several subject: mathematics, physics, chemistry, engineering, and/or computer science. Students enrolled in these prep programs engage weekly in four to six hours of written mock exams, often scheduled on Saturdays, on top of two oral assessments per week. Students are also subjected to frequent peer-based ranking with their classmates based on these written and oral mock exams.

The elite graduate schools play a pivotal role in shaping the country's elite in France. Historically, these renowned institutions were established following the French Revolution with the explicit aim of training a cadre of individuals to serve in leadership positions across various sectors, including politics, business, science, the military, and academia. The selection criteria for admission to these institutions have predominantly been *meritocratic*, in contrast to the *aristocratic* selection practices of the past. In more recent cohorts, approximately 6 percent of a birth cohort graduates from one of these elite schools, with 3 percent graduating from a STEM-focused institution. Notably, a hierarchical structure prevails among these elite graduate schools, with some being exceptionally selective, while others maintain a less rigorous admission process.[9]

STEM Preparatory Programs and Star Classes. This study focuses on Science, Technology, Engineering, and Mathematics (STEM) preparatory programs, which admit approximately 25,000 students annually. This represents about 2.5 percent of a given birth cohort, and roughly two-thirds of the total enrollment in preparatory programs. These STEM prep programs encompass five major tracks (see Diagram A3), each emphasizing distinct disciplines : biology, engineering science, mathematics, physics-chemistry, and/or computer science. In all the tracks but the biology one, students face a tracking system during their second-year classes, between more selective classes which are referred to as "star" classes (*classe étoile*), or standard classes (as displayed in Diagram A1). Although the core curriculum is theoretically consistent between star and standard classes, students in star classes often engage more deeply with academic content beyond the official curriculum and receive enhanced preparation for the most competitive entrance examinations. Their weekly written assessments are structured to mirror the format of the most selective entrance exams. Due to frequent peer-based rankings, star classes foster a more competitive and intellectually stimulating environment. Star classes typically have fewer students and are taught by more experienced teach-

[9]The average percentile rank in the high school graduation exam for students admitted to the top 10% of the most selective elite graduate schools is 91, while it is 38 for those admitted to the bottom 10% of less selective schools (Bonneau et al., 2021).

ers compared to standard classes. Not all preparatory programs offer tracking between star and standard classes in the second year; this feature is more commonly found in the bigger and most prestigious preparatory programs. We exclude the biology track from our analyses because it does not have this tracking system during the second year.

Competitive Entrance Examinations. At the end of their second year in prep programs, students take competitive entrance examinations for admission to elite graduate schools, of which there are more than 200. To manage the volume of exams and prevent scheduling conflicts, several schools have collaborated to share the same entrance examinations, forming five large groups of competitive entrance exams. Candidates can then choose a selection of entrance exams that align with their preferences and objectives.

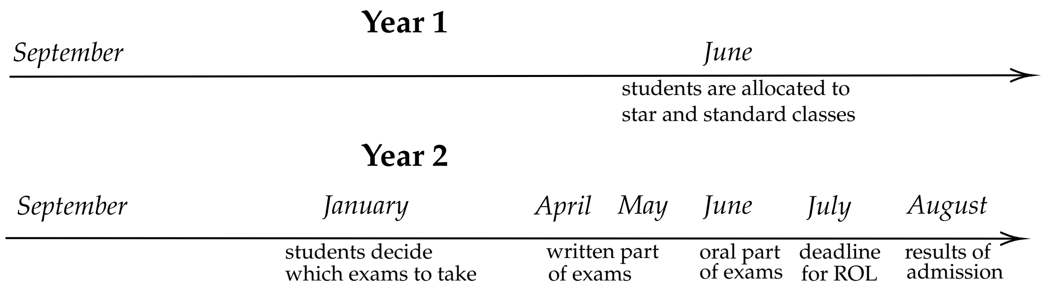
The admission process to elite STEM graduate schools is centralized and overseen by the *Service de Concours Ecoles d'Ingénieurs* (SCEI), based on the college-proposing Gale-Shapley Deferred Acceptance mechanism (Gale and Shapley, 1962), with no limits on the number of programs students can rank in their list of STEM graduate schools.

The admission timeline is detailed in Figure 1. Initially, between December and January, students select which competitive exams to take. The written part of the exams are conducted in April and May and typically include two mathematics tests, two physics tests, a foreign language test, a literature test, and, depending on the track[10] and optional subjects chosen, additional exams in chemistry, engineering sciences, and/or computer science.

Candidates achieving satisfactory results in the written portion are then eligible for oral exams held in June and July. Following these, they must submit their rankings of preferred schools to the clearinghouse by the end of July, with initial offers from the centralized admission process being made shortly thereafter. There are five rounds of offer and acceptance in the admission process, spanning from the end of July to the beginning of September, due to approximately 20 percent of students repeating the second year of the prep program to improve their entrance exam results, thereby exiting the admission procedure and freeing up their spots for others.

[10]There are four main tracks, with slightly differing specialties (more math-intensive, physics-intensive, engineering science-intensive). The organization of these tracks is displayed in Diagram A3.

Figure 1: Schedule of Prep Program and Competitive Entrance Exams



Notes: This diagram shows the preparatory programs calendar and STEM graduate schools admission timeline over the two years of prep programs. Some students repeat the second year to try to get into a better STEM graduate school. ROL stands for rank-ordered list of schools submitted by students.

3 Data

3.1 Data Sources

We exploit four distinct data sources to carry out our analyses.

Administrative data from the centralized admission process to STEM graduate schools (SCEI). We use novel administrative data from the *Service de Concours Écoles d'Ingénieur* (SCEI) from 2015 to 2023. The data contains students' demographic information (age, gender, social background, geographical origin, need-based scholarship status, disability status), students' academic information (prep program where the student is registered, track, star or standard class, number of time taking the competitive exam), the list of competitive exams that each applicant decided to take — which allows us to recover the set of graduate schools to which they have applied, their results at the competitive exams, each student rank-ordered list (ROL) of schools, and admission offers for each round of the admission process.

Administrative data from the Ministry of Education (DEPP, 2010-2020). In order to gain detailed information on prior academic achievement before entering the prep programs, we match through an encrypted identifier the SCEI data with the administrative data from the *Direction de l'évaluation, de la prospective et de la performance* (DEPP), which is in charge of statistics for the Ministry of Education in France. We control for previous academic achievement through high school graduation exam (*Baccalauréat*) and middle school graduation exam (*DNB*) results (both GPA and detailed subject grades).

Data collected in 18 STEM prep programs. Between February 2022 and July 2022, and February and April 2023, we collected data in 18 STEM prep programs who accepted to take part in our study.

Figure A2(a) displays the location of prep programs surveyed. We gathered information about the universe of their recent previous cohorts students and current students, depending on the prep programs' records keeping. We collected students' demographics (gender, date of birth) and academic information (track, class, grades in each subject, admission outcome). We statistically matched this surveyed data to our administrative data on the centralized admission process (SCEI dataset).[11] We thus have detailed information on grades during the prep programs for around 17% of our study sample. Given that the tracking system between standard and star classes is implemented between the first and second years of the prep program, this data allows us to conduct a regression discontinuity design at the margin of star class admission. More generally, this data enables us to open the black box of within-program performance, which is often an overlooked aspect of the literature on the gender gap in STEM programs.

Aggregated earnings data (CTI). We web-scrapped data on the median earnings of alumni from STEM elite graduate schools, by school, graduation cohort, and gender. This information was sourced from the website of the *Commission des Titres d'Ingénieurs* (CTI), for individuals who graduated between 2015 and 2023. The CTI provides, for each STEM graduate school, data on the median earnings of alumni from the previous cohort.[12]

3.2 Samples of Analyses

We restrict our sample to STEM prep programs in France[13] because we have more information about their students, especially about their middle and high school graduation exam results. We also restrict our sample to the combinations of prep programs and tracks that are present throughout the nine-year period covered by our data and have more than 10 students enrolled per year. We exclude the biology track for two reasons: (i) we want to concentrate on programs in which female students are in minority; and (ii) unlike the other tracks, the biology one does not have a tracking system between *star* and *standard* classes in its second year, which is the main feature used in our analysis.

We use different samples in our analyses: the full sample of applicants (N=165,450), a restricted study sample — which includes only applicants from prep programs and tracks that have both a

[11]One prep program opposed this statistical matching. They are thus not matched, but are included in the analyses which only use surveyed prep programs data.

[12]In cases where data for a particular year at a school was unavailable, we made inferences based on information from adjacent cohorts. For the most recent cohorts that have not yet graduated within our study sample, we estimated earnings by referencing the previous cohorts and factoring in the average annual increase in earnings specific to that school. Overall, earnings data is unavailable for 3.8% of individuals of our study sample.

[13]There are also some prep programs abroad included in the administrative data, mainly in Morocco.

standard class and a star class in the second year (N=89,079),[14] and a survey sample, consisting of data collected from prep programs in 2022 and 2023 about their students' grades, which we matched with our administrative data (N=14,746), including those used for the regression discontinuity analysis (N=6,782). The intersection of these different samples is detailed in Figure B4.

3.3 Descriptive Statistics

The descriptive statistics of our sample are displayed in Table 1. We have the information of the students for nine years of entrance examinations, from 2015 to 2023. There are 200 STEM prep programs in mainland France, with more than a quarter of them located in the Parisian area. In Column (4), we compare these descriptive statistics with those of all freshmen and sophomores enrolled in higher education in France during the academic year 2016-2017, with figures sourced from Bonneau et al. (2021).[15]

Female students, while representing 54% of all students enrolled in the first two years of higher education in France, account for only a quarter of all prep program students. The student population of prep programs is predominantly composed of high-achieving individuals with high socioeconomic status (SES). Only 28% of them are need-based scholarship holders (a proxy for low-income students), compared to 38% of all students in France. Furthermore, 62% of prep program students have a father, and 48% have a mother, coming from a high socio-economic status, in stark contrast to 32% (resp. 21%) of all students.[16] Additionally, 50% of them passed the high school graduation exam with the highest honors, while this achievement was attained by only 10% of all freshmen and sophomores. The proportion of prep program students coming from Paris is twice as large as that of all students coming from Paris (8% versus 4%). Regarding enrollment, 18% of prep program students are enrolled in a program in Paris, compared to 10% of all students. 36% of all prep program students are enrolled in star classes.

When we divide our sample between students in star and standard classes, we observe that star classes have fewer female students, and a greater proportion of high SES and high-ability students. Only 40 percent of preparatory programs (82 out of 200) implement this tracking system in the second year, even though these programs accommodate 54 percent of all students (Table B3). Moreover,

[14]This data restriction is mainly because we aim to compare students in star and standard classes within the same programs.

[15]We do not possess exhaustive information on all freshmen and sophomores for the exact same cohort as our prep program students (2015-2023), but the 2016-2017 academic year is likely representative of these cohorts.

[16]For socio-economic status, we rely on the Department of Education's statistical service (DEPP), which classifies occupations into four groups by socio-economic status (SES): high SES includes white-collar professionals, managers, CEOs, teachers, and artists.

these programs have a higher concentration in Paris compared to the overall distribution of preparatory programs.

Table B2 in the Appendix presents the descriptive statistics for male and female students separately. Overall, male and female students exhibit relatively similar observable characteristics with two notable exceptions: women tend to have higher prior academic achievement (59% of them obtained the high school graduation exam with highest honors compared to 47% of men) and are less likely to repeat in the prep program (16% versus 20%).

Table B3 in the Appendix compares the students across our three study samples. The three samples exhibit relatively similar demographic characteristics. However, the study sample, which exclusively comprises programs and tracks with both star and standard classes, includes a higher proportion of students from high socioeconomic status (SES) backgrounds and high achievers. In our surveyed sample, there is a slight overrepresentation of programs from the Parisian area and more recent years, due to the schools' record-keeping practices. Apart from these distinctions, the surveyed sample is quite representative of the study sample in terms of demographic characteristics.

3.4 Outcome Variables

Our primary outcome is the graduate school where applicants receive a final offer in the centralized admission process, regardless of whether the student ultimately accepts the offer. In our analysis, we focus on admission to the most selective graduate schools and we use two distinct definitions of school selectivity.

Selectivity. The first measure assesses a more *objective* selectivity. It relies on the average percentile rank of admitted students at their high school graduation exams. In the the rest of the paper, this definition will be referred to as *selectivity*.

Desirability. The second definition of selectivity is grounded in students' revealed preferences and assesses a more *subjective* selectivity. Following Avery et al. (2013), we estimate the following rank-order logit model:

$$U_{i,j} = \theta_j + \epsilon_{i,j} \tag{1}$$

where $U_{i,j}$ is the rank of school j for student i in its rank-ordered list of schools and θ_j is school j fixed-effect. We use the estimates $\hat{\theta}_j$ (standardized to have mean zero and a standard deviation of one) as a proxy for program desirability. Schools that are more desirable receive a higher volume of

Table 1: Descriptive Statistics (2015-2023)

	Prep. program students (2015-2023)			All students (2016-2017)
	All (1)	Star classes (2)	Standard classes (3)	(4)
A. Students				
Female	0.25	0.23	0.26	0.54
Age	19.6	19.5	19.6	19.2
	(0.8)	(0.8)	(0.8)	(1.4)
Need-based scholarship holder	0.28	0.23	0.31	0.38
Father is high SES	0.62	0.71	0.58	0.32
Mother is high SES	0.48	0.56	0.44	0.21
High School Graduation Exam				
Highest honors	0.50	0.69	0.40	0.10
High honors	0.32	0.24	0.37	0.18
Honors	0.14	0.06	0.19	0.31
Without Honors	0.03	0.01	0.05	0.42
Percentile rank	0.78	0.86	0.74	0.51
	(0.18)	(0.14)	(0.19)	(0.29)
Percentile rank re-weighted (exams. coeffs.)	0.87	0.93	0.84	–
	(0.16)	(0.10)	(0.17)	
From Paris (in high school)	0.08	0.12	0.06	0.04
From Parisian area - outside Paris (in high school)	0.18	0.19	0.17	0.15
Enrolled in Paris	0.18	0.26	0.14	0.10
Enrolled in Parisian area - outside Paris	0.13	0.13	0.13	0.11
Star class	0.36	1.00	0.00	–
Repeater in prep program	0.19	0.19	0.19	–
B. Prep Programs				
Number of prep programs	200	82	197	–
in Paris	23	15	22	–
in Parisian area (outside Paris)	31	12	29	–
Number of classes	546	185	361	–
Number of students	165,450	58,775	106,675	1,090,356

Notes: This table displays the descriptive statistics for all the students in our sample, and then divide our sample between students in star classes and students in standard classes. The sample is constructed from the SCEI administrative datasets, which cover the entire universe of applicants to elite STEM graduate schools from 2015 to 2023, excluding the biology track. Only applicants enrolled in prep programs located in France are included. We retain only those prep programs that were present throughout the nine-year period covered by our data and had more than 10 students per year. Age is the age of the candidates when taking the competitive entrance exam. SES is defined using the classification of the Ministry of National Education. In Column (4), we compare these descriptive statistics with those of all freshmen and sophomores enrolled in higher education in France in the academic year 2016-2017, with figures sourced from Bonneau et al. (2021). Table B2 in the Appendix presents these descriptive statistics for male and female students separately. Table B3 compares the students in our three samples: all applicants, our restricted study sample — which includes only applicants from prep programs with both a standard class and a star class —, and the sample of prep programs that we surveyed in 2022 and 2023 to collect students' grade data during the prep programs.

applications and are more frequently ranked higher on students' rank-ordered lists, indicating their heightened desirability.[17] We also perform a robustness check by defining desirability based on the revealed preferences of female students exclusively. We thus compute distinct school fixed effects for female ($\hat{\theta_{jF}}$) and male students ($\hat{\theta_{jM}}$). This allows us to observe which school characteristics are valued by both genders and to ascertain the role of differences in preferences between male and female students in the underrepresentation of females in the most selective graduate schools. In the rest of the paper, these definitions will be referred to as *desirability* and *desirability for female students*.

Top 10% of STEM graduate schools. Regardless of the specific definition of selectivity, we focus on the top 10 percent of schools, identified as the top-tier STEM graduate schools. The primary variable of interest in our analysis is a binary variable, taking the value 1 if a student receives an offer from a STEM graduate school in the top 10% of most selective schools, and 0 otherwise. Selectivity is determined separately for each of the four tracks considered in our study, as not all tracks lead to the exact same set of graduate schools. Our rankings of graduate school selectivity closely align with popular online rankings accessible to students, such as those found on the *L'Étudiant* website.[18] The various definitions of selectivity exhibit a substantial degree of concordance (for a comprehensive list of the top 10 percent most selective schools under each definition and for each track, please refer to Table B4 in the Appendix).

Motivation of the variables of interest. Securing admission to a top-tier elite STEM graduate school carries profound significance, often leading to significantly enhanced labor market outcomes. Using aggregate data at the school, graduation cohort, and gender level, we identify a pronounced earnings advantage for graduates of more selective schools just one year after graduation. Notably, the earnings gap between alumni from schools in the 9th and 10th selectivity deciles is particularly striking, with an annual difference of approximately €3,300 (for more detailed information, refer to Table B5 in the Appendix). In our sample of engineering graduates, the raw gender pay gap stands at €1,600, while the adjusted gap, accounting for STEM school fixed effects, narrows to €750 — around half of the initial disparity. This suggests that the substantial underrepresentation of women in the

[17] These methodologies have been criticized, partly because even in strategy-proof mechanisms, students may misreport their true preferences (Fack et al., 2019). In our context, misreporting is less likely than in most settings, since there is no limit to the number of schools students can include in their rank-ordered list of preferred graduate schools. Nonetheless, we observe some evidence of "skipping the impossible" strategies. Therefore, we are currently attempting to redefine school desirability using the methodology proposed in Fack et al. (2019), estimating preferences through a conditional logit model on feasible graduate schools.

[18] L'Étudiant website: https://www.letudiant.fr/classements/classement-des-ecoles-d-ingenieurs.html, consulted on May 21, 2023.

most selective schools plays a critical role in contributing to the gender wage gap among STEM workers.

Moreover, elite STEM graduate schools significantly propel their alumni into leadership roles in major corporations. As of October 2022, among the forty CEOs of CAC40 companies, twenty possess a science background. Nineteen of these twenty CEOs were educated at the top 10 % most selective STEM graduate schools.

Finally, another method for evaluating the selectivity of these top schools is their acceptance rate, defined as the percentage of applicants admitted to the school — those ranked above the school's admission threshold — relative to the total number of individuals who took the entrance exam. Figure B7 in the Appendix shows that the most selective graduate schools have the lowest admission rates (12%). This admission rate is comparable to that of the most selective U.S. universities, like Ivy League schools (5.2% in 2022), particularly given the high level of selectivity within the candidate pool, which represents only 2.5% of the highest-performing students in a birth cohort.

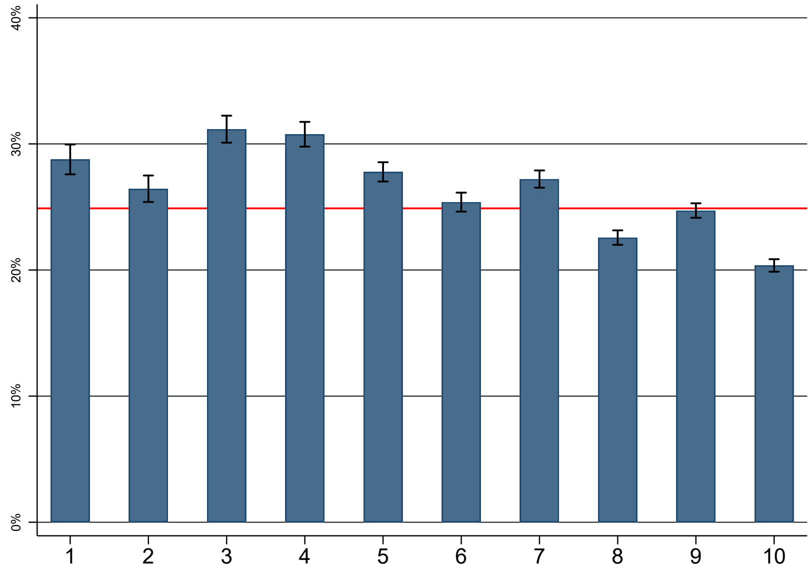
4 Stylized Facts

Figure 2 displays the proportion of women by decile of selectivity of STEM graduate schools. This proportion decreases as the selectivity of the STEM graduate schools increases, reaching its lowest in the top 10 percent most selective graduate schools. In this section, we provide a descriptive analysis to explain this gender gap in access to the most prestigious STEM programs. This discrepancy could arise from (i) a gender gap in achievement before prep programs; (ii) a gender gap in achievement during prep programs; (iii) a gender gap in high-stakes exam performance; or (iv) gender differences in preferences. We aim to quantify the contribution of each of these factors through a decomposition. We can study this thoroughly thanks to our unique dataset, which covers the entire process of elite STEM graduate school recruitment, enriching the burgeoning literature on gender disparities across selection stages (Farré and Ortega, 2023).

For this analysis, we mostly rely on the smaller subsample of programs for which we have grades within the prep program, as we want to observe performance throughout the preparation for high-stakes entrance exams. Table 2 and Figure 4 present the main results of the decomposition. Overall, we observe the evolution of the raw gender gap in the probability of reaching the top 10 percent most selective graduate schools when adding various controls for academic achievement and students' preferences. This raw gender gap, is equal to 6.2 percentage points (Table 2, column 1).[19]

[19]To compute the initial gender gap, we control for demographic characteristics, which are relatively balanced by

Figure 2: Proportion of Female Students, by Decile of Selectivity of STEM Graduate Schools



Notes: This figure shows the proportion of female students in STEM graduate schools, by decile of selectivity of STEM graduate schools. The average number of female students in these schools is represented with the red line. Selectivity of schools is measured using the average percentile rank of admitted students at the high school graduation exam over the period 2015-2023. Figure B8 in the Appendix presents the same statistics using an alternative measure of school selectivity, based on students' preferences (Avery et al., 2013). Figure B9 in the Appendix presents the same statistics for each of the specific tracks (i.e., differing specialties) in prep programs. It shows that the underrepresentation of women in the most selective STEM graduate schools is present across all five tracks. This includes the biology track, which is excluded from most analyses in the paper.

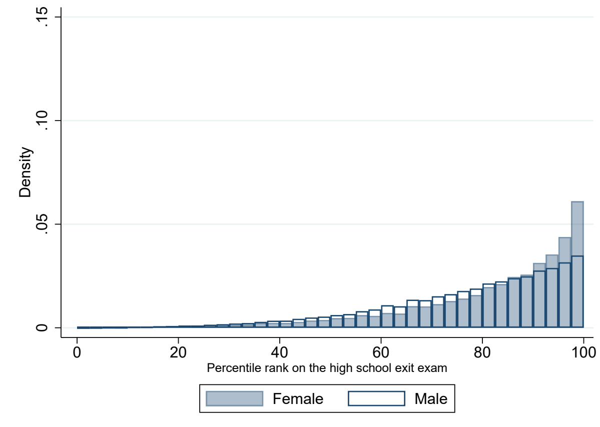
Performance before prep program. Women entering prep programs typically perform better on the national high school graduation exam (Figure 3).[20] This holds true both when using high school graduation exam GPA (Figure 3(a)) and when reweighting the GPA with coefficients from the most competitive entrance exams (see Table A1 for coefficients used for this reweighting), which puts much more emphasis on STEM subjects (Figure 3(b)), although the differences are less pronounced with this reweighting. Therefore, it is not surprising that previous academic achievement does not help explain the gender gap in access to the most selective programs. Based on prior academic achievement alone, we would expect women to enter the more selective programs more frequently than their male counterparts. Controlling for previous academic achievement through detailed grades in each subject obtained at the national high school exam actually increases the gender gap by approximately 20 percent (Table 2, column 2).

Performance during prep program. Subsequent columns in Table 2 control for achievement during the prep program (both decile of average GPA and quintile in each subject) at the beginning of the

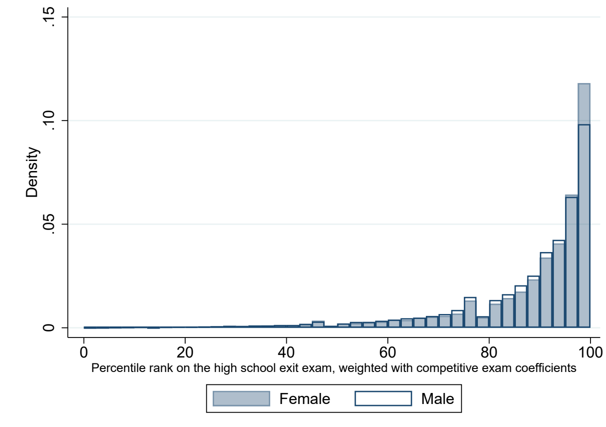
gender (see Table B2), as well as for program, track, and year fixed effects, which are less balanced. Our focus is on the decisions students make after choosing to attend a STEM prep program and selecting which program to attend. The raw gender gap is 8 percentage points (Figure 4), indicating that differences in observable characteristics and program and track choices account for about 20 percent of the raw gender gap.

[20]This finding is consistent with Bechichi and Bluntz (2019), who show that female students self-select more when applying to STEM prep programs.

Figure 3: Percentile Rank at the High School Graduation Exam, by Gender



(a) Raw Percentile Ranks



(b) Percentile Ranks with Re-Weighted GPA

Notes: These histograms illustrate the distribution of percentile ranks of prep program students on the high school graduation exam, by gender. Percentiles are computed among all high school students. In Figure (a), the percentile ranks are derived from the average GPA achieved on the high school graduation exam. In Figure (b), scores are re-weighted with the coefficients of the most selective competitive exams. The coefficients used for this re-weighting can be found in Table A1.

prep program (1st year, 1st semester) and by the end, just before taking the high-stakes entrance exams (2nd year, 2nd semester). Overall, the gender gap forming and increasing in the prep program appears to be a major explanatory factor for the gender gap observed in access to the most selective graduate programs. 52 percent (44 percent when considering only positive factors summing to 100) is explained by performance by the end of the first semester in the prep program. 35 percent (30 percent when considering only positive factors summing to 100) is explained by the increasing gender gap in performance between the first semester and the last semester in the prep programs.[21]

Figure B10 in the Appendix details the evolution of the gender gap in performance (defined as the probability of being in the top of the class) over time in the prep program for each semester. The figure reveals a slightly U-shaped pattern in the gender performance gap.[22]

The prep program period sees a reversal and widening of the gender gap in performance. These dynamics explain around three-quarters of the gender gap in access to the most selective STEM graduate schools.

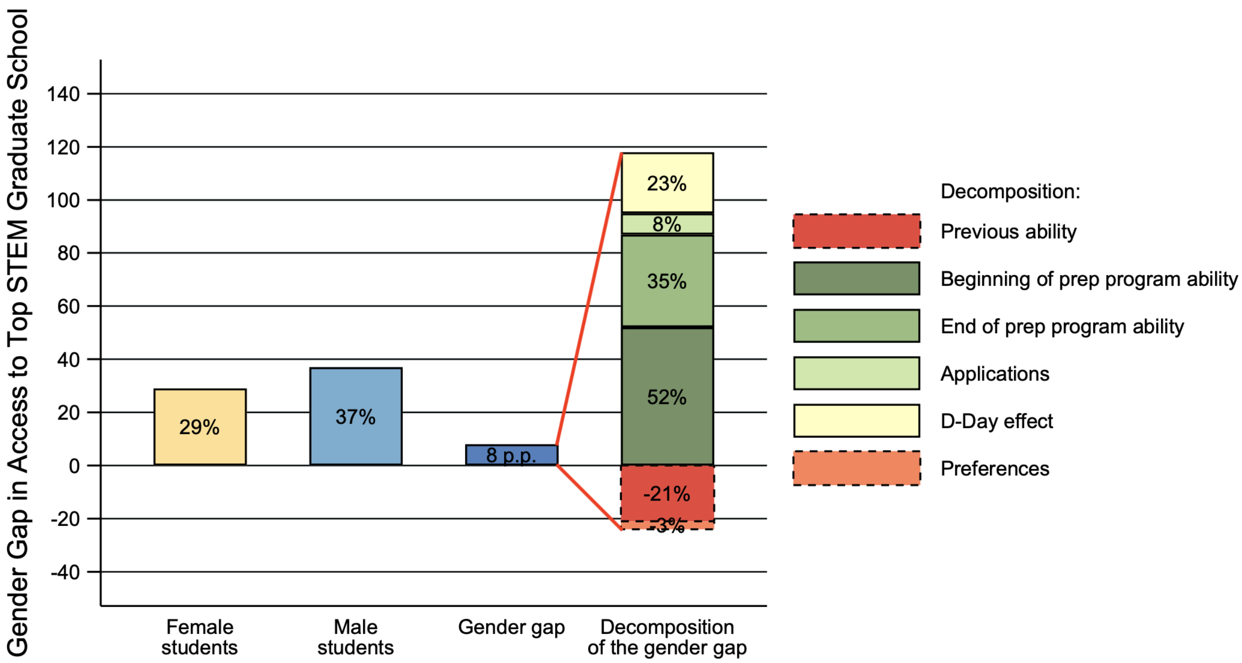
Choice of competitive exams. Overall, differences in the choice of competitive exams to take[23] explain 8 percent (7 percent when considering only positive factors summing to 100 percent) of the gen-

[21]For these last controls, we also include star fixed effects since decile of GPA and quintile of grades are defined at the class level and thus not comparable for students enrolled in the top track (star classes) or non-top track (standard track).

[22]It is crucial to acknowledge that third-year students repeating the prep program represent a self-selected subgroup, given that only 20 percent of students opt to repeat their second year to enhance their admission prospects, with 21 percent of men and only 16 percent of women making this choice.

[23]Applications to top graduate schools are controlled by dummies for application to each of the exams leading to top graduate schools, defined track by track.

Figure 4: Decomposition of the Gender Gap in Access to Top 10% Most Selective Graduate Schools



Notes: This graph displays the decomposition of the gender gap in access to the top 10% most selective graduate schools. We use the subsample of students for whom we have collected grades during the prep program to be able to control precisely for those grades. The decomposition is computed by observing how much the gender gap in access to top STEM graduate schools is reduced when an additional control is added, compared to the raw gender gap of 8 percentage points. Selectivity of schools is defined by the average percentile rank of admitted students at the high school graduation exam. Demographic controls include geographic origin (Paris or the Parisian area), low-income status, the socio-economic status of each parent (categorized into four groups), French nationality, repeater status, disability status, whether the student was in a science academic track during high school, the student's option (either engineering science or computer science), and the gender composition in the *track* $\times$ *program*. We include year, track, and program fixed effects. Previous ability is controlled by (i) decile rank dummy variables from both high school and middle school graduation exams' GPAs, and (ii) quintile rank dummy variables for each grade at the high and middle school graduation exams in the subjects studied in prep programs — mathematics, physics and chemistry, engineering science, French (both written and oral), and foreign languages. Prep programs' ability controls include (i) dummy variables for each decile of the weighted GPA, and (ii) dummy variables for each quintile in grades for mathematics, physics, chemistry, engineering science, computer science, French, and foreign languages, during the first semester (1st year, 1st semester) and last semester (2nd year, 2nd semester) in the prep program. For end-of-prep program ability, we also control for star class status, as performance is defined within each class. Applications to top graduate schools are controlled by dummies for application to each of the exams leading to top graduate schools, defined track by track. D-Day effect refers to performance during competitive entrance exams and is controlled for by (i) dummies for whether or not the applicant is ranked at competitive exams leading to top graduate schools and (ii) the percentile rank of their rank in those exams, controlled linearly for each entrance exam. Preferences are controlled for by adding dummy variables for whether or not each of the top graduate schools is ranked among students' ranked-ordered lists of schools. The order of these variables is chosen to follow the chronological order of decisions made by students.

der gap in access to the top graduate schools (Table 2, Column 5). This suggests that even at the same performance level by the end of the prep programs, male and female students do not opt equally for the most selective exams. This result aligns with previous literature indicating that women tend to shy away from competitive environments (Niederle and Vesterlund, 2007). It's worth noting, though, that the effect of this factor is relatively small compared to the other factors considered here.

Performance on the day of the high-stakes exams. Column 6 of Table 2 indicates that the gender gap in performance on the day of the high-stakes exams explains 23 percent (20 percent when considering only positive factors summing to 100 percent) of the gender gap in access to the top 10 percent most selective STEM graduate schools.[24] This corresponds to the *D-Day Effect* of the exams: for a similar level of achievement by the end of the prep program, women slightly underperform at the most selective high-stakes exams, partly explaining their underrepresentation in the most selective STEM elite graduate schools. These results confirm previous findings from the literature (Azmat et al., 2016; Arenas and Calsamiglia, 2022) on a very selective student population of high achievers who have opted for STEM studies and a competitive selection process. This is striking because these students are supposed to be very well prepared for such high-stakes exams, taking written and oral mock exams that mirror the exact format of the high-stakes entrance exams every week.

Differences in preferences. Differences in preferences are often cited to explain women's underrepresentation in STEM fields (Kahn and Ginther, 2018). To what extent is their underrepresentation in the most selective STEM programs also attributable to gender differences in preferences over STEM graduate schools? Column 7 of Table 2 reveals that the role of preferences is minimal, barely changing the coefficient on the gender gap in access to the most selective STEM graduate schools.[25] If anything, the main differences in preferences are observed regarding the choice of which competitive exams to take, as mentioned earlier, but not on whether to include top STEM graduate schools in their rank-ordered-list of graduate schools. For individuals with the same performance in prep programs, the same set of competitive exams, and the same results at those exams, preferences for top STEM schools are strictly similar.

To corroborate these results, we leverage the very detailed data on students' preferences (i) to observe which characteristics of the graduate schools are valued by the students and (ii) to run coun-

[24]Performance during competitive entrance exams is controlled for by (i) dummies for whether or not the applicant is ranked at competitive exams leading to top graduate schools and (ii) the percentile rank of their rank in those exams, controlled linearly for each entrance exam.

[25]Preferences over graduate schools are controlled for by adding dummy variables for whether or not each of the top graduate schools is ranked among students' ranked-ordered lists of schools.

Table 2: Decomposition Gender Gap in Access to Top 10% Most Selective Schools

	Receive an offer from top 10% most selective graduate schools						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Female student	-0.062***	-0.075***	-0.043***	-0.021**	-0.016*	-0.002	-0.004
	(0.009)	(0.009)	(0.009)	(0.008)	(0.008)	(0.007)	(0.007)
Decomposition		-21%	52%	35%	8%	23%	-3%
% explained by additionnal control							
<u>**Controls:**</u>							
Demographics and Fixed-effects	✓	✓	✓	✓	✓	✓	✓
Demographic characteristics							
Program, track, and year FE							
Previous ability		✓	✓	✓	✓	✓	✓
HS & MS grade FE							
Beginning of prep program ability			✓	✓	✓	✓	✓
1st year - 1st semester grade FE							
End of prep program ability				✓	✓	✓	✓
2nd year - 2nd semester grade FE							
Star class							
Applications					✓	✓	✓
Apply to top school exams							
D-Day Effect						✓	✓
Percentile rank at top school exams							
Preferences							✓
Top schools in ROLs							
N	8,779	8,778	8,778	8,778	8,778	8,778	8,778
Adj-R^2	0.358	0.416	0.516	0.577	0.590	0.724	0.730

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: This table shows the evolution of the gender gap in access to the top 10% most selective graduate schools when adding various controls. We use the subsample of students for whom we have collected grades during the prep program to be able to control precisely for those grades. Table B6 in the Appendix displays the same decomposition on the full sample of students, but without controls for prep program ability. The decomposition raw is computed by observing how much the gender gap is reduced when an additional control is added, compared to the raw gender gap observed in the first column. Selectivity of schools is defined by the average percentile rank of admitted students at the high school graduation exam. Demographic controls include geographic origin (Paris or the Parisian area), low-income status, the socio-economic status of each parent (categorized into four groups), French nationality, repeater status, disability status, whether the student was in a science academic track during high school, the student's option (either engineering science or computer science), and the gender composition in the *track* $\times$ *program*. We include year, track, and program fixed effects. Previous ability is controlled by (i) decile rank dummy variables from both high school and middle school graduation exams' GPAs, and (ii) quintile rank dummy variables for each grade at the high and middle school graduation exams in the subjects studied in prep programs — mathematics, physics and chemistry, engineering science, French (both written and oral), and foreign languages. Prep programs' ability controls include (i) dummy variables for each decile of the weighted GPA, and (ii) dummy variables for each quintile in grades for mathematics, physics, chemistry, engineering science, computer science, French, and foreign languages, during the first semester (1st year, 1st semester) and last semester (2nd year, 2nd semester) in the prep program. For end-of-prep program ability, we also control for star class status, as performance is defined within each class. Applications to top graduate schools are controlled by dummies for application to each of the exams leading to top graduate schools, defined track by track. Performance during competitive entrance exams is controlled for by (i) dummies for whether or not the applicant is ranked at competitive exams leading to top graduate schools and (ii) the percentile rank of their rank in those exams, controlled linearly for each entrance exam. Preferences are controlled for by adding dummy variables for whether or not each of the top graduate schools is ranked among students' ranked-ordered lists of schools. The order of these variables is chosen to follow the chronological order of decisions made by students. Standard errors are clustered at the *track* $\times$ *program* $\times$ *cohort* level.

terfactual simulations of the school-student matching algorithm, estimating the preferences of female students based on those of male students. These results are more precise as they rely on the full structure of students' preferences rather than a binary indicator of whether a graduate school is ranked in students' rank-ordered lists (ROLs) (Table 2, Column 7).

Overall, male and female students tend to value the same graduate school characteristics (Table 3). We regress specific school characteristics and their interactions with gender against school fixed effects from revealed preferences (Equation 1). Comparing school fixed effects from female ($\hat{\theta_{jF}}$) and male ($\hat{\theta_{jM}}$) preferences shows no significant differences. Both genders value top 10% most selective schools, overall selectivity, expected earnings, and military schools similarly. Some coefficients for female students are slightly negative, indicating a lower valuation for these characteristics, but the differences are minor and statistically insignificant. This consistency holds even when all characteristics are considered jointly (Column 6, Table 3). These results are consistent with the fact that graduate schools in the top decile of desirability for both men and women are very similar (Table B4).

Table 3: STEM Graduate School Characteristics Valued by Students, by Gender

	(1)	(2)	(3)	(4)	(5)	(6)
	School FE	School FE	School FE	School FE	School FE	School FE
Top Decile of Selectivity	2.07***					0.75***
	(0.076)					(0.064)
Top Decile of Selectivity × Female students	-0.082					-0.023
	(0.11)					(0.091)
Decile of Selectivity		0.32***				0.20***
		(0.011)				(0.0081)
Decile of Selectivity × Female students		-0.020				-0.019
		(0.015)				(0.011)
Expected Earnings			0.15***			0.065***
			(0.0073)			(0.0048)
Expected Earnings × Female students			-0.0033			0.00078
			(0.010)			(0.0068)
All Military School				1.59***		0.46***
				(0.18)		(0.074)
All Military School × Female students				0.042		0.097
				(0.25)		(0.10)
Strictly Military School					0.45*	
					(0.25)	
Strictly Military School × Female students					-0.19	
					(0.35)	
Female students	-0.041	0.078	0.077	-0.059	-0.049	0.041
	(0.046)	(0.11)	(0.41)	(0.062)	(0.066)	(0.26)
Track fixed effects	✓	✓	✓	✓	✓	✓
Analytical Weights (number of admitted students)	✓	✓	✓	✓	✓	✓
N	976	976	920	976	976	920
R^2	0.605	0.647	0.499	0.170	0.031	0.869

Standard errors in parentheses
$^{*}\ p < 0.1$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$

Notes: This table presents the regression coefficients of various school characteristics — as well as their interactions with a gender dummy variable — on school fixed effects computed separately for male and female students ($\hat{\theta_{jF}}$ and $\hat{\theta_{jM}}$), allowing an analysis of the attributes valued by students in general and female students in particular. The school fixed effects are calculated using a rank-ordered logit model based on students' ranked lists of preferred schools (Avery et al., 2013), and these effects also serve to quantify school desirability. In this analysis, we incorporate weights equal to the number of students admitted to each school, to avoid placing too much weight on small schools. The school fixed effects are normalized to have a mean of 0 and a standard deviation of 1, and are computed separately for each track. The regression thus also includes track fixed effects."Expected earnings" refers to the median gross annual salary, inclusive of bonuses, of the most recent alumni cohort from each school, with data retrieved from the (CTI website). The category "Military schools" encompasses all such institutions (*Ecole de l'Air*, *Ecole Navale*, *ENSTA Bretagne*, *ENSTA Paris*, *ESM Saint-Cyr*, *Polytechnique*), while "Strictly military schools" includes the same list minus *Polytechnique*, a prestigious and multifaceted school that, despite its military affiliation, does not exclusively lead to military careers.

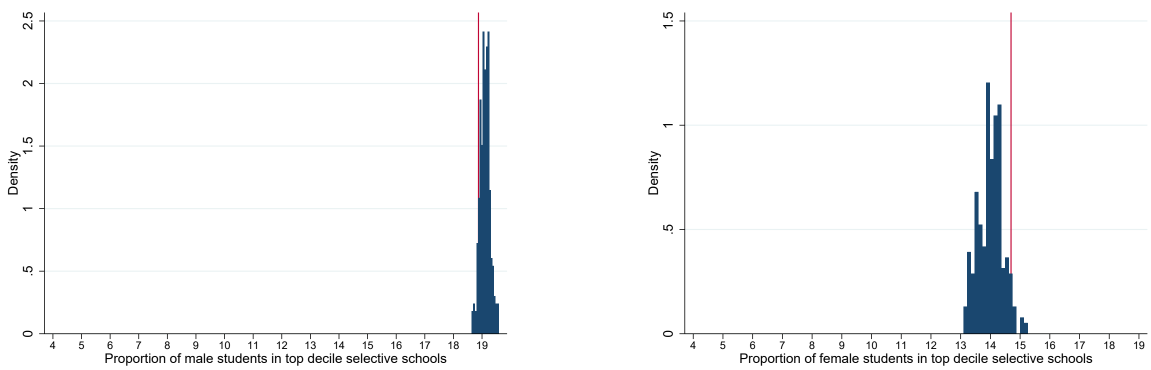
To gauge gender differences in preferences more precisely, we estimate male students' preferences using a rank-order logit model (Equation 1) and predict valuations for both genders. We then run 300 simulations of the matching algorithm using different error vectors (with $\epsilon_{i,j}$ being independent and identically distributed (i.i.d.) according to a Type I extreme value (Gumbel) distribution), observing the gender distribution in the top 10% most selective schools.[26] Results suggest that if women had the same preferences as men, their representation in the top 10% most selective graduate schools might

[26]These methods face criticism as students may misreport preferences even in strategy-proof mechanisms (Fack et al., 2019). Misreporting is less likely here since there is no limit to the number of schools in the rank-ordered list. However, some 'skipping the impossible' strategies are noted. We are refining preference estimations using methods proposed by Fack et al. (2019).

actually slightly decrease (Figure 5). This could be due to intensified competition for the same STEM graduate schools, considering that women tend to underperform in high-stakes exams. Current minimal gender differences in preferences might benefit women's representation in top schools.

Taken together, these results suggest that women's underrepresentation in the most selective STEM graduate schools is not primarily due to differing preferences.

Figure 5: Access to the 10% Most Selective Graduate Schools When Estimating Student Preferences From Male Ones



(a) Male students

(b) Female students

Notes: These figures depict results from 300 counterfactual simulations of the school-student matching algorithm, using preferences of male students to predict the preferences of all students. The red bars display the current proportion of male and female students accessing the top 10% most selective STEM graduate schools; the blue bars display this proportion under the counterfactual simulation scenario. Specifically, we estimate a rank-ordered logit model based on schools ranked in the ROLs by male students and predict the school valuation for all students. We then generate 300 different error vectors, each distributed according to a Type I extreme value (Gumbel) distribution, and re-run the school-student matching algorithm 300 times. These results are preliminary since we are currently attempting to redefine the estimation of students' preferences using the methodology proposed in Fack et al. (2019), estimating preferences through a conditional logit model on feasible STEM graduate schools.

Table B6 in the Appendix presents the results of the same decomposition as Table 2, but using the full sample of prep program students. The advantage of this analysis is the increased sample size (N=165,418 compared to N=8,778), while the drawback is the lack of detailed information on within-prep program performance, aside from whether the student was enrolled in a star class, which is a crude measure of prep program performance. According to this decomposition, previous achievement does not explain the gender gap in access to top programs (-13%). Achievement during the prep program, measured roughly by star class status, explains 31%. The choice of competitive exams accounts for 24%, performance on the exam day (D-Day) accounts for 56%, and preferences for graduate schools do not explain this gap at all.

Overall, we are able to observe which factors drive the gender gap in admission to the top 10% most selective graduate schools. We show that this gap is primarily explained by an increasing gender gap in performance between the entry point and the end point of the preparatory program, and to a lesser extent, by differences in performance on the exam day (D-Day). This suggests that studies focusing only on performance on the day of the exam miss a significant aspect: the widening gender gap in performance during exam preparation. It also highlights the need to explore the impact of the learning environment on the performance of both genders.

In the following section, we leverage the variation in the learning environment induced by the tracking between *star* and *standard* classes in prep programs to observe how changes in the learning environment for exam preparation affect each gender differently.

5 Results

In this section, we aim to test the hypothesis that the competitive *learning environment* within preparatory programs disproportionately benefits male students' performance over female students', thereby widening the gender gap in performance over time in the prep programs. Addressing this question require exogenous variation in the competitiveness of the learning environment experienced by students. We therefore leverage a specific feature of these programs: the tracking in the second year between *star* and *standard* classes. The star classes are comprised of the higher-achieving students, which effectively creates a more selective and competitive learning environment. Star classes feature increased selection, heightened competition, enhanced intellectual stimulation, and a more rigorous preparation for the STEM graduate schools' entrance examinations. We thus examine whether preparing exams in a star class, as opposed to a standard class, has a different impact on the success of male and female students at the competitive entrance examinations using two different complementary methodologies.

First, we discuss the average relative performance of male and female students in star and standard classes, while controlling for detailed previous academic achievement. Secondly, we study what is happening for students at the margin of star class admission, using a regression discontinuity design. The first methodology uses the full sample of prep program students enrolled in a program that offers both standard and star classes (N=89,065) to estimate an Average Treatment Effect (ATE): it does not specifically focus on students who are marginally admitted to star classes. However, it does not allow us to control for unobservable factors that may impact differently male and female students after prep program entry, though it does control for detailed achievement just before entry. The second methodology applies to a smaller subsample (N=6,782 in the RDD sample) and estimates a Local Average Treatment Effect (LATE) valid only for students at the margin of star class admission. The strength of this approach is its robustness to unobservable factors, as it compares students with similar observable and unobservable characteristics who differ only by star class status.

In the last part of the section, we discuss the mechanisms that could explain the increased gender gap in star classes compared to standard classes.

5.1 Comparison of Star and Standard Classes

We aim to uncover the gender differential impact of preparing exams in a more selective and competitive environments on students' academic success. This analysis constitutes a "double" difference, comparing both gender (*male students* versus *female students*) and class status (*star* versus *standard*). In this analysis, we restrict the full sample of students to the combinations of tracks and prep programs that have both a standard and a star class, as we aim to compare similar students who differ only in their second-year class status. This reduces the sample to 54 percent of the original sample (see Table B3 in the Appendix).

5.1.1 Empirical Strategy

We compare the average relative performance of male and female students preparing for exams in star and standard classes, while we control for detailed previous academic achievement. We thus perform the following reduced-form estimation:

$$y_{ikpc} = \alpha_0 + \alpha_1 \times F_i \times S_i + \alpha_2 \times F_i + \alpha_3 \times S_i + \gamma \boldsymbol{X_i} + \lambda_k + \lambda_p + \lambda_c + \epsilon_{ikpc} \quad (2)$$

where y_{ikpc} is a dummy variable that takes value 1 if student i in track k in program p in cohort c is admitted in a STEM graduate school in the top 10% of selectivity or desirability, F_i is an indicator of individual i being a female student, S_i is an indicator of a individual i being enrolled in a star class, $\boldsymbol{X_i}$ is a vector of characteristics of student i, including detailed measures of previous academic achievement.[27] We include track (λ_k), program (λ_p) and cohort (λ_c) fixed effects. We estimate Equation 2 with a linear probability model, and test the robustness of our results with a probit model (Table C23 in the Appendix). Our coefficient of interest is α_1.

To ensure that our results regarding the interaction between star class and gender are not influenced by the gender differential impact of other observable characteristics, we also perform the

[27] Demographic controls include geographic origin (Paris or Parisian area), low-income status, the socioeconomic status of each parent (into four categories), French nationality, repeater status, disability status, whether the student was in a science academic track during high school, the student's option (either engineering science or computer science), and the gender composition in the *track* $\times$ *program*. Previous ability is controlled by (i) decile rank dummy variables from both high school and middle school graduation exams GPA, and (ii) quintile rank dummy variables for each grades at the high and middle school graduation exams at the subjects studied in prep programs — mathematics, physics and chemistry, engineering science, French (both written and oral), and foreign languages.

following reduced-form equation:

$$y_{ikpc} = \alpha_0 + \alpha_1 \times F_i \times S_i + \alpha_2 \times S_i + \gamma_M \boldsymbol{X_i} + \gamma_F \boldsymbol{X_i} \times F_i + (1 + F_i) \times (\lambda_k + \lambda_p + \lambda_c) + \epsilon_{ikpc} \quad (3)$$

This equation closely resembles Equation 2, with the introduction of interactions between all observable characteristics and fixed effects with gender.

For the double difference between gender and star class status to capture the differential impact of star classes on male and female achievement, we must make three assumptions:

Assumption 1 *Assignment to a star class is determined solely by ability.*

Assumption 2 *The variance in the achievement distribution of male and female students at entering prep program is similar.*

Assumption 3 *There is an overlap between the competitive exams taken by standard and star class students.*

We test these assumptions using data on previous academic achievements, including high school exam results, detailed program grade data, and responses from a teacher survey.

Assumption 1. Assumption 1 posits that top-performing first-year students, regardless of gender, are placed in star classes for their second year, with no affirmative action for female students. If women had easier access to star classes, our coefficient regarding the gender gap in performance specific to those classes would be biased. This assumption aligns with prep programs' goals of enhancing success rates in entrance exams, as they are ranked based on this metric in popular student rankings. However, the significant underrepresentation of female students in STEM, especially in star classes, raises the possibility of lower-achieving females being admitted to maintain gender balance or support their STEM engagement.

Our data from 18 prep programs, covering 14,746 students, allows us to examine this assumption (Table C8). Results show that women have less access to star classes, but the gender gap disappears when adjusting for detailed subject grades,[28] indicating no gender gains easier access to star classes, given equivalent academic performance. Similarly, the size of the first stage of the regression discontinuity is very similar for men and women, showing that when they cross the eligibility threshold,

[28]We control for (i) decile rank dummy variables for the GPA, and (ii) quintile rank dummy variables for grades in mathematics, physics and chemistry, engineering science, computer science, French, and foreign languages.

men and women have the same probability of gaining access to star classes, thus supporting Assumption 1 (Table 5).

Furthermore, a 2021 survey of prep program teachers from 14 high schools, with a 70% response rate, corroborates this finding. All respondents confirmed that star class admissions are based solely on first-year academic performance, with no gender preference for students with the same achievement level.

Assumption 2. Assumption 2 posits equal variance in male and female students' achievements upon entering prep programs. If male students' abilities are more spread out while female students cluster around the mean, our method would isolate true ability differences rather than the competitive environment's effect on students. The argument of higher variability in male students' performance is often cited as an explanatory factor in the literature on gender achievement gaps (Hedges and Nowell, 1995; Machin and Pekkarinen, 2008; Johnson et al., 2009); however, few explanations have been found for this observation despite its occurrence in various contexts.

Our teacher survey in prep classes shows 64% observe no difference in gender achievements at the end of the first year.[29] Furthermore, analysis of high school graduation scores, adjusted by exam coefficients, confirms similar variance between genders, supporting Assumption 2. The standard deviation of the weighted percentile rank on the high school graduation exam scores is 28.1 for male students compared to 28.0 for female students. We cannot reject at the 10% level the test of equality of variance between male and female distributions of the weighted high school graduation exam GPA.

We also present evidence in Table C10 and Figure 6 showing that there were as many students at the top of the distribution among future star class enrollees before entering the star class, strongly supporting Assumption 2.

Overall, it does not appear that our results are driven by more variance in the ability of male students than female ones prior to entering these programs.

Assumption 3. We examined the overlap in exams between students in standard and star classes to confirm that the star class effect is on the learning environment rather than the exam-taking context. To differentiate, we defined exams leading to only top-tier schools as *very selective* and those leading to at least one top-tier school as *selective*. Findings show 98% of star class students take at least one selective exam (versus 89% in standard classes), and 84% take a very selective exam (versus 57% in

[29] Among the others, 18% state that male students are higher-achieving (which does not pose a threat to our identification strategy), 9% say female students are higher-achieving, 3% perceive male students are more centered around the mean with females in the tails, and 3% view the opposite.

standard classes), with star class students attempting more exams on average. Despite significant overlap in the exams taken, some difference persists, so we include a robustness test of our main results controlling for the number and type of exams taken (see Table C27).

5.1.2 Results

In this section, we present our results on the impact of studying in a more selective and competitive environment on female students compared to their male counterparts, specifically on the access to top 10% STEM graduate schools. We then proceed with a heterogeneity analysis and conduct several robustness checks.

Impact on access to top 10% STEM graduate schools. We present the results of a double difference between gender and star class status in Table 4, while controlling for detailed previous achievement. The initial two columns refer to graduate school selectivity, measured by the average percentile rank of admitted students at the high school graduation exam. In contrast, the last two columns refer to school desirability, based on the revealed preferences of applicants (Avery et al., 2013). Columns (1) and (3) exhibit the outcomes derived from Equation 2, while Columns (2) and (4) present the results from Equation 3, with all control variables and fixed effects interacted with gender.

All specifications show that female students benefit less from the selective and competitive educational environment of the star classes than their male counterparts. The gender gap in access to top-tier schools is wider in star classes compared to the same gender gap in standard classes. Our preferred estimates in Column (2) quantifies that, relative to the gender gap observed in standard classes, female students in star classes have a 3.3 percentage points lower probability of admission to the top-tier STEM graduate schools than male students in star classes. This 3.3 percentage points reduction corresponds to a 15 percent decrease in the probability of admission from a baseline admission probability of 22 percent. Similarly, female students have a lower probability of admission to top 10% desirable STEM graduate schools of 4.6 percentage points (Column 4), which represent a 20 percent reduction from baseline.[30] The gender gap in accessing top-tier STEM schools significantly differs between star and standard class students, being more than twice as large for star class students, at 6.0 (2.7 + 3.3) percentage points, compared to 2.7 percentage points for standard class

[30]Table C11 in the Appendix displays all the coefficients of the control variables included in the regression of Equation 2. Generally, students from high socioeconomic status have a higher probability to get an offer from a top-tier STEM graduate school, as do repeaters and students from Paris or, to a lesser extent, the Parisian area. Conversely, need-based scholarship recipients tend to perform less well. The proportion of female students in the *track* $\times$ *program* has no impact in our setting on student performance, and we verified that this holds true for both genders separately.

Table 4: Access to STEM Graduate Schools in the Top 10 Percent of Selectivity

	(1) Top 10% grad. schools (Selectivity)	(2) Top 10% grad. schools (Selectivity)	(3) Top 10% grad. schools (Desirability)	(4) Top 10% grad. schools (Desirability)
Baseline proba. of access	0.22	0.22	0.23	0.23
Female student × Star class	-0.042***	-0.033***	-0.058***	-0.046***
	(0.0061)	(0.0064)	(0.0060)	(0.0064)
Star class	0.27***	0.27***	0.29***	0.28***
	(0.0061)	(0.0062)	(0.0061)	(0.0062)
Female student	-0.027***		-0.028***	
	(0.0031)		(0.0030)	
Demographic Controls	✓	✓	✓	✓
MS & HS Exam Score	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
Track FE	✓	✓	✓	✓
Program FE	✓	✓	✓	✓
Demographic Controls × Female student		✓		✓
MS & HS Exam Score × Female student		✓		✓
Year Fixed-Effects × Female student		✓		✓
Track Fixed-Effects × Female student		✓		✓
Program Fixed-Effects × Female student		✓		✓
N	89,065	89,065	89,065	89,065

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: The table illustrates the change in the probability of admission to top-tier STEM graduate schools for female students in star classes over the period 2015-2023. Selectivity of schools is measured in two ways: Columns (1) and (2) use the average percentile rank of admitted students at the high school graduation exam, while Columns (3) and (4) are based on the revealed preferences of applicants (Avery et al., 2013). In Columns (2) and (4), we include interactions of all controls and fixed-effects with a gender dummy variable, allowing observable characteristics to have gender-specific performance impacts. Demographic controls include geographic origin (Paris or Parisian area), low-income status, the socio-economic status of each parent (into four categories), French nationality, repeater status, disability status, whether the student was in a science academic track during high school, the student's option (either engineering science or computer science), and the gender composition in the *track* × *program*. Previous ability is controlled by (i) decile rank dummy variables from both high school and middle school graduation exams GPA, and (ii) quintile rank dummy variables for each grades at the high and middle school graduation exams at the subjects studied in prep programs — mathematics, physics and chemistry, engineering science, French (both written and oral), and foreign languages. We include year, track and program fixed effects. Standard errors are clustered at the *track* × *program* × *cohort* level. For a comprehensive list of coefficients on the demographic control variables, refer to Table C11 in the Appendix.

students.

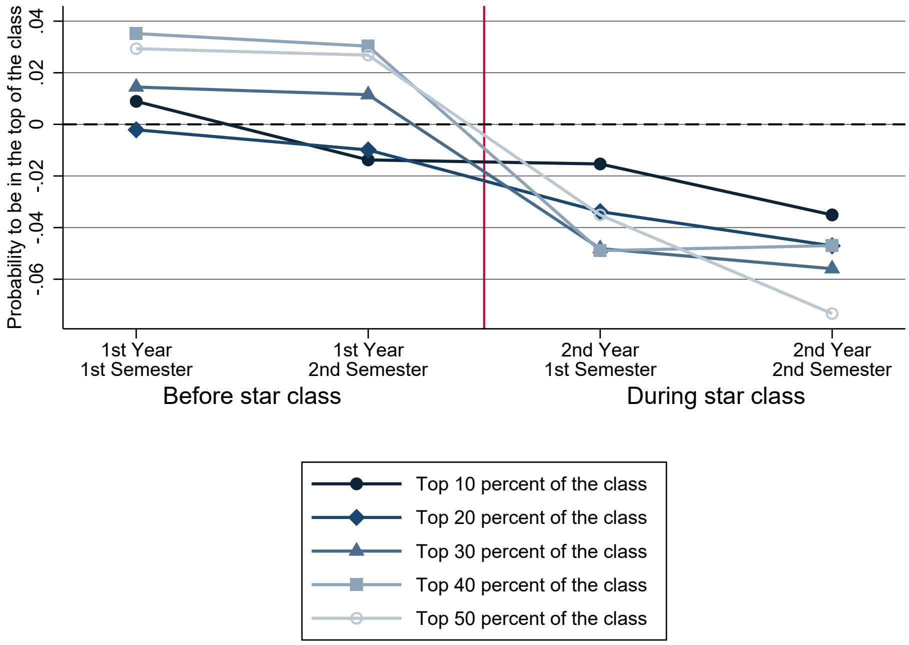
Interestingly, for both definitions of school selectivity, the coefficient with the interaction of all control variables and fixed effects with gender is about 80% of the coefficient without these interactions ($\frac{3.3}{4.2}$ and $\frac{4.6}{5.8}$). This suggests that the differential impact of observable characteristics at baseline on male and female students, particularly the differential impact of previous ability on future ability, could explain at most 20% of the increased gender gap observed in star classes.

During the Prep Programs. We use the subsample of students for whom we collected grades during the prep program to observe the double difference (Equation 2) in the probability of being among the high achievers in the class, both before and after entering the star classes. The results are presented in Figure 6 and Table C10 in the Appendix. Overall, the results indicate that a gender gap in the probability of being among the top achievers emerges after star class entry. Table C10 shows no significant gender gap in the likelihood of being among the high achievers during the first year, but a significant gap emerges after star class entry: women are 4.6 percentage points less likely to be among the top 20% of the class than their male counterparts in star classes, with minimal differences in standard classes. Consistent with these results, Figure B10(b) illustrates that the gender gap in the probability of being among high achievers is more pronounced in star classes compared to standard classes.

Impact on expected earnings. We do not possess individual salary data; however, we have retrieved median salaries disaggregated by graduate school, cohort, and gender for the various STEM graduate schools in our sample from the CTI website. In our STEM graduate sample, the unadjusted gender pay gap is €1,600. After incorporating STEM school fixed effects, the gap reduces to €750, or about 50% of the initial gap. This suggests that an important part of the gender pay gap among engineers is due to differences in access to STEM graduate schools with the best labor market outcomes.

Table C12 in the Appendix presents the effect of studying in a more selective and competitive study environment on the median expected salaries of female and male engineering students, one year after graduation. Columns (1) and (2) use earnings without bonuses, while Columns (3) and (4) use earnings with bonuses. As in our main results in Table 4, we interact all observable characteristics and fixed effects with gender in Columns (2) and (4). The gender gap in expected earnings is approximately 40 to 70 percent larger (depending on the inclusion of bonuses) among alumni of the star class than among those from standard classes. Specifically, our preferred estimates in Column (4)

Figure 6: Probability To Be Among the Top of the Class, by Gender and Star Class Status



Notes: The figure shows the change in the probability of being among the top of the class (top 10%, top 20%, top 30%, top 40%, top 50%) during the prep programs for female students in future star classes on the left-hand side, and in actual star classes on the right-hand side. We use percentile rank weighted GPA to rank students. Demographic controls include geographic origin (Paris or Parisian area), low-income status, socio-economic status of each parent (four categories), French nationality, disability status, high school science academic track, student's option (engineering science or computer science), and the gender composition in the *track* × *program*. The sample is restricted to non-repeaters.

indicate that for female students preparing for entrance exams in a more competitive environment, the gender pay gap — including bonuses — increases by around €900 compared to those in a less competitive learning environment. This represents a 2.0 percent decrease in expected annual earnings, based on a baseline of €42,234. Even right after graduation, and when using aggregate wage data, we find that the impact of the competitive environment created by this selective STEM program on the gender wage gap is significant, albeit relatively small. A different impact might be observed with individual salary data or with salary data later in the career. These results also confirm the prevalence of the gender pay gap at the upper end of the earnings distribution among STEM workers, a phenomenon observed in previous literature (Blau and Kahn, 2017).

Heterogeneity. In this section, we present the heterogeneity of our main results by parental income status, prior academic achievement, and repeater status.

By income status. The gender gap in access to top STEM graduate schools for star class students increases significantly more for low-income students than for high-income students, both in absolute and relative terms (see Table C14).[31] Specifically, low-income female students in a star class are rela-

[31] As our data do not include a detailed measure of parental income, we use need-based scholarship status as a proxy, categorizing need-based scholarship recipients as low-income students and non-need-based scholarship recipients as

tively 5.4 percentage points less likely to gain admission to top STEM graduate schools compared to their low-income male peers, which corresponds to a 39% reduction from baseline. Results regarding school desirability follow a similar pattern.

By initial prior achievement. Table C15 presents the results segmented by initial achievement quartiles of students.[32] The increased gender gap in performance observed in star classes is significant across the achievement distribution but is considerably higher in relative terms at the bottom of the distribution compared to the top. Results for school desirability, detailed in Table C16, follow a similar pattern.

By prep program selectivity. We observe heterogeneity by the selectivity of prep programs (Table C17), defined by the percentile rank of admitted students at the high school graduation exam, similar to graduate schools. The negative effect of star classes on women is more pronounced in absolute terms in mid-selectivity prep programs and in relative terms in less selective prep programs. However, in the most selective prep programs the negative gender effect is more pronounced across the entire sample, not just in star classes (for which the effect is not significantly different than in standard classes), suggesting that all classes in these programs might foster a competitive environment.

By repeater status. Approximately 20 percent of students opt to repeat the second year of the prep program decide to retake the competitive entrance exams in order to improve their admission outcomes. This decision is correlated with gender; around 21 percent of men choose to repeat, compared to only 16 percent of women. While we control for repeater status in the main regression, we verify whether our results are consistent across repeater status. Overall, the outcomes are relatively similar for students attempting the exam for the first or second time (Table C18). This suggests that repeating does not enable female students in star classes to bridge the performance gap with their male counterparts, nor does it exacerbate it. To scrutinize this further, we conduct the same regression as in Equation 2, but with repeater status substituted for star class status, to discern the differential impact of repeating on male and female students. The results are shown in the Appendix in Table C13. These findings imply that there is no differential impact of repeating on male and female students, suggesting that for female who opt to remain in the competitive environment of the prep program, there is no gender-disparate impact from additional time spent in these programs. This could imply that once they become accustomed to this competitive environment, female students may benefit from it as much as their male counterparts do.

high-income students.

[32] Achievement is measured through re-weighted GPA from the high school graduation exam, using weights equivalent to those of the most selective competitive exams, as shown in Table A1.

Robustness checks. In the Appendix, we display various specifications to verify the robustness of our results.

Placebo test. First, we conduct two placebo tests using end-of-high-school GPA and first-year prep program GPA to ensure that our findings are not influenced by pre-existing differences in performance between genders and future star class statuses prior to entering the star class. Tables C19 and C20 show that, before entering star classes, there is no gender disparity in initial performance between students destined for standard and star classes.

Controlling for Program Grades. Table C9 presents our results, adjusting for first-year grades of approximately 10,000 students. Including first-semester grades slightly reduces our baseline estimate, but it remains significant at the 5 percent level. Notably, the sample size for this analysis is smaller. When controlling for end-of-year grades, the gender effect of competitive classes on performance loses statistical significance, though the estimates remain negative and close to initial figures in terms of magnitude. This suggests the gender gap might partly emerges during the first year, becoming more pronounced in the second year as students are tracked into standard and star classes.

Revealed preferences of female students. Table C21 shows results with school selectivity defined based on the revealed preferences of female students. Given that, as discussed in Section 4, the preferences of female students are very similar to those of their male counterparts[33], the results are quite comparable to our baseline estimates, albeit slightly larger.

Other definitions of selectivity. Table C22 tests our results with different definitions of STEM graduate school selectivity. The findings remain consistent whether selectivity is based on average school rankings or the top 20% of schools instead of the top 10%. For female students in star classes compared to their male peers, school selectivity decreases by an average of 0.40 ranks, translating to a 0.5% reduction in average school selectivity. The gender gap in school desirability increases by 0.13 ranks in star classes, a 14% amplification. The gender gap also broadens for access to the top quintile of selective schools, though less than for the top 10% (a 6% reduction in selectivity and an 8% reduction in desirability), indicating the most significant gender gap occurs at the very top of the school selectivity distribution.

Probit model. In Table C23, we show that our findings remain robust when using a probit model instead of a linear probability model.

Interaction of fixed effects. We further ensure the robustness of our results by interacting the year, prep program and track fixed effects together, allowing program quality to vary over time and tracks. The outcomes of this analysis are shown in Table C24 and the impact of being a female student in a

[33]Refer to Table B4 for the list of schools categorized by each selectivity definition

star class on gaining admission to STEM graduate schools in the top 10% of selectivity is very similar to our main estimates in Table 4.

Controlling for the number of selective exams taken. Students in standard and star classes do not take exactly the same exams. To ensure our results are not solely due to differences in exam types, we replicated the results in Table 4, controlling for the number of selective exams taken using fixed effects.[34] The results, presented in Table C27, closely match our baseline estimates. This suggests the observed effect is more attributable to changes in the exam preparation setting rather than changes in the competitiveness of the testing environment.

Results track-by-track. We check the consistency of our results across all tracks. Table C25 and C26 indicates that the observed increase in the gender gap for access to top 10% STEM graduate schools within star classes is present in every track.[35].

Results year-by-year. We verify the consistency of our results across different years. All coefficients are negative and display a similar magnitude, though the negative effects seem somewhat intensified post-Covid — from 2020 to 2023 (refer to Table C28 for school selectivity and Table C29 for school desirability).

In this section, we observed a widening gender gap in the likelihood of accessing top-tier STEM graduate schools among students preparing exams in more competitive environments (star classes) compared to less competitive ones (standard classes). This suggests that the learning environment exacerbates the gender performance gap. The double difference strategy, conducted on the full sample (N=89,065) and closer to an average treatment effect (ATE), cannot control for unobservable factors that may develop differently among male and female students after star class entry; it only controls for previous academic achievement.

In the following section, we present a complementary analysis using a regression discontinuity (RD) approach at the margin of star class entry.

5.2 Regression Discontinuity at the Margin of Star Class Admission

We control for a wide array of observable characteristics in the double difference analysis by gender and star class status. However, to ensure that unobservable characteristics differing by gender are

[34] We include fixed effects for the number of (1) very selective exams (leading to only top-tier schools) and (2) selective exams (leading to at least one top-tier school).

[35] Specifically, for students in track 1 (MP), being a female in the most competitive classes reduces the probability of admission to the most selective STEM graduate schools by 3.9 p.p., corresponding to a 14% reduction. The reductions are 4.3 p.p. (21%) for track 2 (PC) students, 5.3 p.p. (25%) for track 3 (PSI) students, and 3.7 p.p. (74%) for track 4 (PT) students.

not driving the results, we also perform a regression discontinuity analysis at the margin of star class admission. This second methodology provides estimates with greater internal validity, as it compares students who are very similar in both observable and unobservable characteristics, differing only in their access to star classes. Nonetheless, this approach has the drawback of reduced external validity. This is due to (i) the regression discontinuity design yielding a local average treatment effect that is only applicable to students at the margin of star class entry, and (ii) the analysis being limited to the smaller subsample of students for whom we have detailed grade data throughout the program (N=6,782) .

5.2.1 Methodology

We aim to estimate the effect of enrollment in a star class on the probability of gaining admission to one of the top 10% most selective STEM graduate schools, by gender. We leverage the tracking feature present in most prep programs: at the end of the first year, students are divided into star and standard classes based on their performance. In most programs, students are free to accept or decline this offer, creating a context suitable for a fuzzy regression discontinuity design. This analysis is very close to the one performed by Landaud and Maurin (2020). We depart from them by specifically examining the gender differential effect of enrollment in a star class, using a larger and more diverse sample of programs across France.

Unlike most regression discontinuity settings, we do not have direct information on the exact running variable and cutoff used for star class admission.[36] Given the variety of programs and classes in our study, the first step of our analysis is to identify an appropriate running variable and cutoff.

Discussions with prep program teachers confirmed that students gain admission to star classes based on their end-of-first-year class ranking in terms of GPA and the capacity constraints of star classes in their program and track. Therefore, we chose to use students' class rank as the running variable to determine their eligibility for star class enrollment.[37] Consequently, we chose to use students' class rank as the running variable to determine their eligibility for star class enrollment.

Definition of the threshold. We identify the rank threshold at which a significant discontinuity in the probability of accessing a star class is observed in each program. We are reasonably certain that such a discontinuity exists, although its precise value is unknown to us. Therefore, we build on the

[36] We collected data on grades throughout the prep programs but were not provided with the precise running variable and threshold used for admission to star classes.

[37] To rank students, we consider their end-of-first-year weighted GPA, calculated using either weights directly provided by the prep programs or weights equivalent to those of the most selective competitive exams (Table A1).

methodology defined by Hansen (2000) to extract these cutoffs for each class in our sample. This approach has been widely adopted in contexts similar to ours (Hoekstra, 2009; Landaud et al., 2020; Bütikofer et al., 2023), particularly since Porter and Yu (2015) showed that this methodology can be used with a regression discontinuity design. We identify the rank within each class, in each prep program, for each academic year that exhibits the most pronounced discontinuity in the probability of star class admission.[38]

We normalize the cutoff to 0 for all classes. Hence, our running variable is essentially the relative rank distance to the estimated cutoff of star class admission. Following Fort et al. (2022) recommendations for multi-cutoff RD analysis, we include prep program $\times$ class fixed effects in our analyses.

RDD sample. The algorithm used to identify a fuzzy discontinuity in star class admission tends to consistently pinpoint a discontinuity between one student admitted to the star class and one student enrolled in the standard class. These students are outliers: the one on the right side of the threshold is very likely to be an always-taker (enrolling in the star class anyway, even though students ranked above do not), while the one on the left side of the threshold is likely to be a never-taker (not enrolling in the star class, even though students ranked below do). Including these individuals in our estimation sample could bias our estimates (De Chaisemartin and Behaghel, 2020), as they are likely to have specific unobservable characteristics that make them always-takers or never-takers, which could be correlated with their admission to the most selective STEM graduate schools.

Subsequently, we define our RD sample as balanced, retaining an equal number of students from each side of the cutoff for each class.[39] Finally, we exclude observations whose running variables are above 18 or below -18, to omit running variable values for which very few programs are observed.[40] This ensure at least 20 students per running variable value.

First stage. Our method identifies a fuzzy regression discontinuity. Not all students above the estimated threshold enter star classes, nor do all students below it attend standard classes. Table 5 and Figures D11 display the fist stage of the regression discontinuity[41] The first stage is around 0.90, meaning that being above the cutoff increases the probability of star class admission by 90%. This

[38]We select the cutoff that yields the highest R^2 value when we regress a dummy variable taking a value of one if a student accesses a star class on a dummy variable taking a value of one if a student's rank is above the rank under consideration. This process is repeated for all ranks within a given class.

[39]For instance, consider a class of 30 students: the top 10 students enroll in the star class while the remaining 20 attend a standard class. In this case, we would only include the top 20 students in our analysis to have a balanced sample of students around the threshold.

[40]The average class size in our sample is 35 students (with a maximum of 48 and a minimum of 10).

[41]Note that the coefficients in the table and graph differ slightly because the graph does not include program $\times$ class fixed effects, while the table does.

indicates that our method for identifying the running variable and cutoff is effective. The effect is of similar magnitude for both men and women (although noisier for women due to a much smaller sample size). This supports Assumption 1: male and female students are admitted to star classes based on the same criteria and with similar magnitudes around the threshold.

Table 5: First Stage of the Regression Discontinuity, by Gender

	(1) Star Class (All Students)	(2) Star Class (Male Students)	(3) Star Class (Female Students)
Baseline mean (Below cutoff)	0.08	0.09	0.05
Baseline RD estimate	0.90*** (0.06)	0.92*** (0.06)	1.00*** (0.11)
Robust 95% CI	[0.78 ; 1.02]	[0.81 ; 1.04]	[0.78 ; 1.21]
Obs. used in estimation	1,778	1,303	316
Total number of obs.	6,782	4,934	1,848

Notes: This table presents the non-parametric regression discontinuity estimates of the probability of attending a star class, constituting the first stage of our RD analysis. These estimates are based on Calonico et al. (2017, 2019). We display the first stage using the optimal bandwidth selection of our main estimates (Table 7). Selectivity of graduate schools is measured using the percentile rank from high school graduation exam results of admitted students. The running variable is defined the distance between a student's rank and the estimated cutoff rank for star class admission, computed at the *class* $\times$ *track* $\times$ *prep program* $\times$ *year*. We include *program* $\times$ *class* fixed effects. Standard errors are clustered at the level of the running variable.

Density test. Figure D12 in the Appendix shows the density of the running variable around the threshold, based on Cattaneo et al. (2018). The density is continuous and symmetric around the estimated star class entry threshold, due to the sample restrictions described above.

Balance of observable characteristics. We test for the balance of observable characteristics in Table 6. The threshold was set by the algorithm where the most significant disparity in star class admission is evident. This means that "real differences" in observable characteristics may exist at this point. However, for our gender-based analysis, it is important that no gender-differentiated selection in observable characteristics is evident at the threshold. Panel A shows that demographic characteristics are perfectly balanced around the threshold for both male and female students. Panel B shows that previous academic achievement is slightly less balanced. The sample is balanced around the threshold for both male and female students regarding GPA at the end of the first year, just before entering the star classes. This is reassuring since GPA is the basis of our running variable, students' rank in the class. However, there are some gender imbalances around the cutoff for mathematics, physics, and foreign language grades at the end of the first year. The coefficients for male and female students are of the same sign and magnitudes for mathematics and foreign language. For physics, it seems to be imbalanced only for male students. To ensure these imbalances do not affect our re-

Table 6: Balance of Observable Characteristics Around the RD Threshold, by Gender

	All Students		Male Students		Female Students	
	Baseline mean (1)	RD Estimate (2)	Baseline mean (3)	RD Estimate (4)	Baseline mean (5)	RD Estimate (6)
Panel A. Demographics						
Female	0.27	-0.03	–	–	–	–
		(0.09)				
Age	19.34	0.03	19.34	-0.15	19.32	0.35
		(0.14)		(0.24)		(0.31)
Need-based scholarship holder	0.24	0.04	0.24	0.07	0.24	-0.07
		(0.14)		(0.17)		(0.14)
High SES	0.81	0.01	0.82	0.03	0.80	-0.05
		(0.13)		(0.16)		(0.23)
From Paris	0.10	0.02	0.10	-0.06	0.09	-0.00
		(0.10)		(0.12)		(0.11)
Panel B. Previous Academic Achievement						
Highest honor at HS graduation exam	0.76	-0.00	0.73	0.01	0.84	-0.05
		(0.08)		(0.09)		(0.17)
Percentile Rank GPA in 1st Year - 2nd Semester	0.51	-0.01	0.51	0.00	0.50	-0.01
		(0.03)		(0.04)		(0.05)
Percentile Rank in Math in 1st Year - 2nd Semester	0.51	0.14**	0.53	0.14**	0.46	0.14
		(0.06)		(0.07)		(0.11)
Percentile Rank in Physics in 1st Year - 2nd Semester	0.51	0.07	0.53	0.10**	0.48	-0.03
		(0.06)		(0.05)		(0.12)
Percentile Rank in French in 1st Year - 2nd Semester	0.51	-0.14	0.47	-0.13	0.61	-0.05
		(0.09)		(0.10)		(0.09)
Percentile Rank in Foreign Language in 1st Year - 2nd Semester	0.51	-0.25***	0.48	-0.26**	0.59	-0.19
		(0.08)		(0.11)		(0.18)
Panel D. All Baseline Characteristics Jointly						
F-Stat		8.943		6.791		2.914
P-value		0.000		0.000		0.001
N.		6,782		4,934		1,848

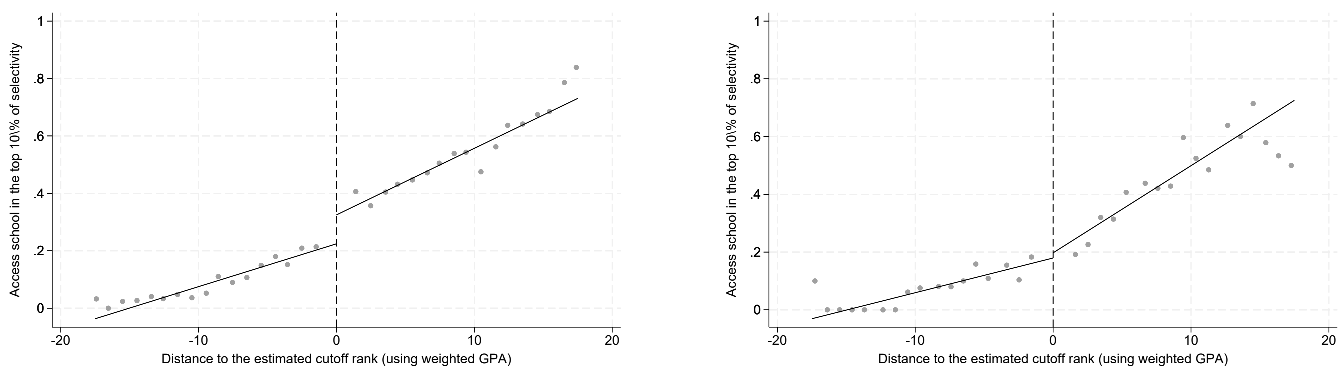
Notes: This table presents non-parametric regression discontinuity biased-corrected estimates based on Calonico et al. (2017, 2019) to compare the characteristics of students near the cutoff for star class admission, by gender. Panel A and Panel B each assess different aspects of students' characteristics at this cutoff. Each coefficient results from a separate regression, where the student's relative distance to the cutoff serves as the running variable. Non-parametric estimates use a triangular kernel, with optimal bandwidths calculated separately for each outcome and sample. Columns 1, 3, and 5 display the mean value of the dependent variable for students below the cutoff. We include prep program $\times$ class fixed effects. Standard errors are in parentheses.

gression discontinuity estimates, we included these variables as controls in a robustness check in Table D37. The results are similar to our main estimates, indicating that our findings are not driven by these imbalances.[42]

5.2.2 Results

Our regression discontinuity analysis reveals a gender discrepancy in outcomes (Figure 7). Male students increase their chances of admission to top 10% STEM schools by attending a star class, while the effect seems null for female students. The graphs show that although accessing star classes seems to have a null effect for female students, the slope is steeper for female star class students compared to female standard class students. This suggests that star classes might slightly improve their performance gain, even if not significantly for students marginally admitted. Additionally, the graphs indicate that gender differences seem to occur at the very top of the distribution: the probability of male students accessing top graduate schools increases exponentially, while it collapses for female students. However, the RDD results discussed here are only informative for marginal students.

Figure 7: Admission to Top 10% Most Selective STEM Graduate Schools, by Gender



(a) Male Students

(b) Female Students

Notes: These figures depict the probability of admission to top-tier STEM graduate schools over the period 2015-2023, around the margin of star class admission, by gender. Selectivity of schools is measured using the percentile rank from high school graduation exam results of admitted students. The running variable is defined as the distance between a student's rank and the estimated threshold rank for star class admission, computed at the *class* $\times$ *track* $\times$ *prep program* $\times$ *year* level.

We perform a fuzzy regression discontinuity analysis on the effect of accessing star classes on admission to the most selective graduate schools (Table 7).[43] Star classes significantly increase the

[42] As an additional check, we regressed the admission to a top 10% STEM graduate schools on the percentile rank of the mathematics grade at the end of the second semester of the first year. We find that a 0.14 higher percentile rank in mathematics at the end of the first year of prep program is associated with an increase of 0.07 in the probability of admission to a top 10% STEM school. This is reassuring because it means that the imbalance in mathematics grades around the cutoff could not explain all the positive effect we observe on admission to a top 10% school at the cutoff.

[43] Figure 7 and Table 7 display slightly different results. Figure 7 plots the pooled intent-to-treat effect of gaining admission to a star class, while Table 7 provides the 2SLS estimates, including prep program $\times$ class fixed effects, following Fort et al. (2022) recommendations.

likelihood of admission to the top 10% of STEM graduate schools. At the margin, the probability increases by 13.1 percentage points, an 120% increase from a baseline of 11%. These findings align with results from Landaud and Maurin (2022) on one preparatory program.

Our findings indicate significant gender-based heterogeneity in the results. For male students, enrolling in a star class significantly increases the likelihood of admission to the top 10% most selective STEM graduate schools by 20 percentage points, a 180% increase from male students below the cutoff, significant at the 1 percent level. In contrast, for female students, the effect is negative but only marginally significant (at the 10 percent level only). Estimates varying bandwidth size show that most intervals are compatible with a null effect for women (Figure D15). The effects for female and male students are statistically different from each other.

Table 7: Admission to Top 10% Most Selectivite STEM Graduate Schools, by Gender

	(1) Top 10 decile of selectivity (All students)	(2) Top 10 decile of selectivity (Male students)	(3) Top 10 decile of selectivity (Female students)
Baseline mean (Below cutoff)	0.11	0.11	0.09
Baseline RD estimate	0.131*** (0.046)	0.202*** (0.048)	-0.119* (0.070)
Robust 95% CI	[0.040 ; 0.221]	[0.108 ; 0.296]	[-0.255 ; 0.018]
Obs. used in estimation	1,778	1,303	814
Total number of obs.	6,782	4,934	1,848

Notes: This table displays the fuzzy non-parametric regression discontinuity estimates of admission probability to top-tier STEM graduate schools. These estimates are based on Calonico et al. (2017, 2019). In alignment with recent advancements in regression discontinuity literature, we include bias-corrected estimates and robust standard errors. School selectivity is measured using the percentile rank from high school graduation exam results of admitted students. The running variable is defined as the distance between a student's rank and the cutoff rank for star class admission, computed at the *class* $\times$ *track* $\times$ *prep program* $\times$ *year*. Our analysis includes controls for program $\times$ class fixed effects. Standard errors are clustered at the level of the running variable.

We run several robustness tests of the main results in Appendix Section D4.3. The results are quite sensitive to the definition of graduate school selectivity, showing a null effect of star class admission on women when using school desirability (Table D30) and even positive effects when using graduate school desirability for female students (Table D31). This highlights the need for cautious interpretation. However, the effects are larger for men in all these specifications.

We find that our main results holds when: (i) we do not cluster at the running variable level (Table D32); (ii), we increase the bandwidth size (Figure D15, Tables D34 and D35); (iii) we take a local polynomial of order two (Table D36); (iv) we control for the unbalanced grades around the cutoff (Table D37); (v) we include prep programs $\times$ class $\times$ year fixed effects (Table D38); (vi) we perform a local randomization analysis,[44] which is an alternative methodology in the case of a regression discontinuity with a discrete running variable (Table D33).

In this section, we use the methodology developed by Hansen (2000) to identify a threshold for star class admission and observe the causal effect of star class enrollment on students' success in competitive STEM graduate school entrance exams. We find that, at the margin, star class admission significantly improves male students' performance but has a null or even negative effect on female students. Consistent with the previous section, star classes are found to increase the gender performance gap. In the next part, we present the mechanisms explaining the increased gender gap in access to top programs in star classes.

5.3 Mechanisms

We have observed a higher gender gap in performance in star classes compared to standard classes, both on average and at the margin of star class entry. Using detailed data on performance over time in the program and on students' preferences, we decompose the factors explaining this increased gender gap in the probability of gaining admission to top-tier graduate schools, as discussed in Section 4 for the raw gender gap. Results are presented in Table 8.

As for the raw gender gap, the star class-specific gender gap is not explained by previous academic achievement (-16%). Beginning of prep program ability explains 26% of this gap (respectively 21% when considering only explaining factors summing to 100 percent), while end of prep program performance explains 37% (resp. 29%). The role played by specific gender differences in star class in the choice of competitive exams is marginal (5%, resp. 4%). In contrast, the specific gender gap in performance on the D-Day of the high-stakes exams for star class students is large, explaining 58% of the star class gender gap (resp. 46%). As with women overall, gender differences in preferences toward STEM graduate schools do not play a role, if not a compensatory factor limiting the underrepresentation of women in the most selective programs (-13%), consistent with results from the estimation of students' preferences (Figure 5).

Overall, the results regarding the specific gender gap in performance in star classes are relatively close to those for the raw gender gap, except that the role of the D-Day of the exam is amplified: around half of the gender gap specific to star class is explained by exam performance, compared to

[44] With this methodology, appropriate for discrete running variables like in our case, we focus on students around the cutoff, retaining the 4 students below and 4 students above, which corresponds to the optimal bandwidth computed by Calonico et al. (2019). We then perform a simple below- and above-mean comparison to test for the local average treatment effect (LATE). The underlying assumption is that these students are essentially randomly assigned just below or just above the cutoff.

only 20% of the raw gender gap. In contrast, the role of exam preparation is less striking, although still important: explaining half of the star class gender gap as opposed to three-fourths of the raw gender gap. These results suggest that the gender gap in high-stakes exam performance may be particularly significant among the high achievers, enrolled in star classes.

In the last section, we explore the impact of exposure to higher-ability peers as a mechanism for the star class effect.

6 The Role of Peers

Studying in star classes can impact various aspects of a student's learning experience, including studying alongside higher-achieving peers, encountering more demanding teachers, delving deeper into the most challenging parts of the curriculum, taking more difficult mock exams (often 6-hour written exams compared to 4-hour ones), and benefiting from smaller class sizes, among others. In this section, we aim to focus on one specific aspect of star classes: exposure to higher-ability peers. Our goal is to quantify the share of the differential impact of star classes by gender that could be explained by the presence of high-achieving peers. To achieve this, we take advantage of the quasi-random year-to-year variations in peer ability within a program to observe the differential impact of peer ability on gender performance in the competitive entrance exams.

We define peer ability as the average percentile rank of peers in the high school graduation exam grades distribution, re-weighted by the coefficients of the most selective competitive exams (as reported in Table A1). Peer achievement is measured before individuals enter prep programs, to avoid conflating peer effects with individuals subjected to the same common shocks (such as different teachers, headmasters, exam subjects in this specific year, etc.). The peer group is defined as the first-year cohort peers in one's track and program. We exclude the individual under consideration from the computation of peer ability by using leave-own-out means. Rather than defining peer ability at the class level, we define it at the level of the entire first-year cohort within a program and track. This approach addresses the potential for endogenous sorting of peers across classrooms within a program starting in the first year, before the tracking system between star and standard classes institutionalizes this sorting. Figure E16 (a) shows the variation in peer ability that we are using, from cohort to cohort within a certain program and track. All cohort characteristics included in the regression[45] are also defined at this level. We include year and program $\times$ track fixed effects in all

[45]Proportion of female students, proportion of French students, proportion of high SES students, proportion of students from Paris, and proportion of students from the Parisian area.

Table 8: Decomposition Gender Gap in Star Class in Access to Top 10% Most Selective Schools

	Receive an offer from top 10% most selective graduate schools						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Female student × Star class	-0.038**	-0.044***	-0.034**	-0.020	-0.018	0.004	-0.001
	(0.017)	(0.016)	(0.016)	(0.016)	(0.015)	(0.013)	(0.013)
Decomposition		-16%	26%	37%	5%	58%	-13%
% explained by additionnal control							
<u>**Controls:**</u>							
Demographics and Fixed-effects	✓	✓	✓	✓	✓	✓	✓
Star class							
Female student							
Demographic controls							
Program, track, and year FE							
Previous ability		✓	✓	✓	✓	✓	✓
HS & MS grade FE							
Beginning of prep program ability			✓	✓	✓	✓	✓
1st year - 1st semester grade FE							
End of prep program ability				✓	✓	✓	✓
2nd year - 2nd semester grade FE							
Star class							
Applications					✓	✓	✓
Apply to top school exams							
D-Day Effect						✓	✓
Percentile rank at top school exams							
Preferences							✓
Top schools in ROLs							
N	8,779	8,778	8,778	8,778	8,778	8,778	8,778
Adj-R^2	0.494	0.513	0.539	0.577	0.590	0.724	0.730

Standard errors in parentheses
$^{*}\ p < 0.1$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$

Notes: This table shows the evolution of the star class specific gender gap in access to the top 10% most selective graduate schools when adding various controls. We include controls for star class and female students in all specifications. We use the subsample of students for whom we have collected grades during the prep program to be able to control precisely for those grades. Table B7 in the Appendix displays the same decomposition on the full sample of students, but without controls for prep program ability. The decomposition raw is computed by observing how much the gender gap is reduced when an additional control is added, compared to the raw gender gap observed in the first column. Selectivity of schools is defined by the average percentile rank of admitted students at the high school graduation exam. Demographic controls include geographic origin (Paris or the Parisian area), low-income status, the socio-economic status of each parent (categorized into four groups), French nationality, repeater status, disability status, whether the student was in a science academic track during high school, the student's option (either engineering science or computer science), and the gender composition in the *track × program*. We include year, track, and program fixed effects. Previous ability is controlled by (i) decile rank dummy variables from both high school and middle school graduation exams' GPAs, and (ii) quintile rank dummy variables for each grade at the high and middle school graduation exams in the subjects studied in prep programs—mathematics, physics and chemistry, engineering science, French (both written and oral), and foreign languages. Prep programs' ability controls include (i) dummy variables for each decile of the weighted GPA, and (ii) dummy variables for each quintile in grades for mathematics, physics, chemistry, engineering science, computer science, French, and foreign languages, during the first semester (1st year, 1st semester) and last semester (2nd year, 2nd semester) in the prep program. Applications to top graduate schools are controlled by dummies for application to each of the exams leading to top graduate schools, defined track by track. Performance during competitive entrance exams is controlled for by (i) dummies for whether or not the applicant is ranked at competitive exams leading to top graduate schools and (ii) the percentile rank of their rank in those exams, controlled linearly for each entrance exam. Preferences are controlled for by adding dummy variables for whether or not each of the top graduate schools is ranked among students' ranked-ordered lists of schools. The order of these variables is chosen to follow the chronological order of decisions made by students. Standard errors are clustered at the *track × program × cohort* level.

specifications to identify the effect of having higher-achieving peers from variation across cohorts within a program $\times$ track. Contrary to the rest of the analyses, these fixed effects are the first year ones, to be consistent with the level of peer variation we are using. We also interact all observable individual characteristics and cohort characteristics with gender, to allow these characteristics to have a gender specific effect. Overall, we conduct the following reduced-form equations:

$$\begin{aligned} y_{ikpc} = \alpha_0 &+ \alpha_F \overline{A}_{-i,kpc} \times F_i + \alpha_M \overline{A}_{-i,kpc} + \gamma_F \boldsymbol{X_i} \times F_i + \gamma_M \boldsymbol{X_i} \\ &+ \pi_F \boldsymbol{\overline{C}_{-i,kpc}} \times F_i + \pi_M \boldsymbol{\overline{C}_{-i,kpc}} + \lambda_k \times \lambda_p + \lambda_c + \epsilon_{ikpc} \end{aligned} \tag{4}$$

And, disaggregating by gender of the high-achieving peers:

$$\begin{aligned} y_{ikpc} = \alpha_0 &+ \beta_F F_i \times \overline{AM}_{-i,kpc} + \beta_M \overline{AM}_{-i,kpc} + \delta_F F_i \times \overline{AF}_{-i,kpc} + \delta_M \overline{AF}_{-i,kpc} \\ &+ \gamma_F \boldsymbol{X_i} \times F_i + \gamma_M \boldsymbol{X_i} + \pi_F \boldsymbol{\overline{C}_{-i,kpc}} \times F_i + \pi_M \boldsymbol{\overline{C}_{-i,kpc}} \\ &+ \lambda_k \times \lambda_p + \lambda_c + \epsilon_{ikpc} \end{aligned} \tag{5}$$

where y_{ikpc} is a dummy variable that takes the value 1 if student i in track k, program p, and cohort c is admitted to a STEM graduate school in the top 10% of selectivity or desirability; F_i is an indicator for individual i being a female student; $\overline{A}_{-i,kpc}$ is the average percentile rank at the high school graduation exam of the first-year peers of individual i in track k, program p, and cohort c, excluding individual i (leave-own-out mean); $\overline{AM}_{-i,kpc}$ is the same statistics for male peers only; $\overline{AF}_{-i,kpc}$ is the same statistics for female peers only; $\boldsymbol{X_i}$ is a vector of characteristics of student i, including detailed previous academic achievement; $\boldsymbol{\overline{C}_{-i,kpc}}$ is a vector of characteristics of the first-year peers of individual i in track k, program p, and cohort c, excluding individual i (leave-own-out mean); λ_k represents first year track fixed effects; λ_p represents first year program fixed effects; and λ_c represents first year cohort fixed effects.

Table 9 displays the coefficients of the interaction between peer ability and gender; school selectivity is examined in the first two columns, and school desirability in the last two. Columns (1) and (3) correspond to Equation 4; Columns (2) and (4) correspond to Equation 5. The results indicate that the presence of higher-achieving peers positively influences male performance more significantly than female performance. Specifically, the estimates in Column (1) suggest that a one percentile increase in the average ability of cohort peers would result in a 0.11 percentage point increase in the probability of males gaining access to top 10% most selective graduate schools, representing a 0.7 percent increase in the baseline probability; whereas for females, this effect is 0.06 percentage point smaller.

Table 9: Effect of Having Higher-Achieving Peers on Access to Top 10% Most Selective STEM Graduate Schools, by Gender

	(1) Top 10% grad. schools (Selectivity)	(2) Top 10% grad. schools (Selectivity)	(3) Top 10% grad. schools (Desirability)	(4) Top 10% grad. schools (Desirability)
Baseline proba. of access	0.15	0.15	0.16	0.16
Average percentile rank of 1st year peers at the HS graduation exam	0.11***		0.10***	
	(0.018)		(0.018)	
Female student × Average percentile rank of 1st year peers	-0.059***		-0.059***	
	(0.018)		(0.018)	
Average percentile rank of 1st year female peers at the HS graduation exam		0.029***		0.027***
		(0.0098)		(0.010)
Female student × Average percentile rank of 1st year female peers		0.017		0.013
		(0.014)		(0.015)
Average percentile rank of 1st year male peers at the HS graduation exam		0.079***		0.073***
		(0.016)		(0.016)
Female student × Average percentile rank of 1st year male peers		-0.075***		-0.071***
		(0.020)		(0.020)
Demographic Controls	✓	✓	✓	✓
MS & HS Exam Score	✓	✓	✓	✓
Cohort Controls	✓	✓	✓	✓
Demographic Controls × Female student	✓	✓	✓	✓
MS & HS Exam Score × Female student	✓	✓	✓	✓
Cohort Controls × Female student	✓	✓	✓	✓
Program × Track FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
N	165,031	165,031	165,031	165,031

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: The table reports the change in the probability of admission to top 10% most selective STEM graduate schools for students with higher-achieving peers, with results presented by gender in Columns (1) and (3), and further broken down by the gender of the high-achieving peers in Columns (2) and (4). Peer achievement is defined based on the average percentile rank of students at the national high school examination, measured before they enter prep programs. Grades at the national high school examination are re-weighted with the coefficients of the most selective competitive exams (see Table A1). The peer group is defined at the 1st year track × year × program level, excluding the individual under consideration (leave-own-out mean). Selectivity of schools is measured in two ways: Columns (1) and (2) use the average percentile rank of admitted students at the high school graduation exam, while Columns (3) and (4) are based on the revealed preferences of applicants (Avery et al., 2013). Demographic controls include geographic origin (Paris or Parisian area), low-income status, the socio-economic status of each parent (into four categories), French nationality, repeater status, disability status, whether the student was in a science academic track during high school, the student's option (either engineering science or computer science), and the gender composition in the *track × program*. Previous ability is controlled by (i) decile rank dummy variables from both high school and middle school graduation exams GPA, and (ii) quintile rank dummy variables for each grades at the high and middle school graduation exams at the subjects studied in prep programs — mathematics, physics and chemistry, engineering science, French (both written and oral), and foreign languages. Cohort controls include the proportion of female students, the proportion of French students, the proportion of high SES students, the proportion of students from Paris and the proportion of students from the Parisian area. Cohort controls and peer characteristics are measured in 1st year of prep program. We include interactions of all demographics, previous achievement and cohort controls with a gender dummy variable, allowing observable characteristics to have gender-specific performance impacts. We do not interact fixed effects with gender since this is the source of variation we are using. We include year, and track × program fixed effects in all specification. Contrary to the rest of the analysis, these fixed effects are determined based on the first year of the program, as this corresponds to the level of peer variation we are using. Standard errors are clustered at the *track × program* level, also corresponding to the first year in the program.

These findings appear to show that, in this student population, male students are more responsive to the presence of higher-achieving peers than female students. Considering that students are frequently ranked alongside their classmates during prep programs, this difference could be attributed to a stronger inclination towards competition among male students. When surrounded by higher-achieving peers, men tend to enhance their performance more than women, who exhibit smaller sensitivity to their peer ability.

We further investigate whether students react differently to the ability of their male and female peers, in Columns (2) and (4) of Table 9. We find that the ability of female peers has a positive effect on male and female students' performance alike. Conversely, we show that the ability of male peers has a positive effect on male students' performance (0.079), but no effect on female students' performance (0.079 - 0.075). Interestingly, our findings also indicate that male students are almost three times more influenced by the abilities of their male peers than by those of their female peers (0.075 vs. 0.029). Results regarding school desirability presented in Columns (3) and (4) are very similar.

In a robustness test presented in the Appendix (Table E39), we interact fixed effects two by two[46] to allow the average quality of programs to vary over time and by track. The variation we are using is thus differential variation in peer ability across tracks within programs over time.[47] Figure E16 (b) shows the variation in peer ability that we are using. Overall, results are very close to our baseline estimates.

In Table E40 in the Appendix, we present the same set of results for the subset of programs that offer both a standard and a star class. These results are closely aligned with those for the full sample presented in Table 9, albeit slightly larger in magnitude. We use these results on the restricted sample to conduct a back-of-the-envelope calculation to assess the share of the star class effect that could be explained solely by the presence of better-achieving peers. On average, within a specific program, year, and track, students in star classes are surrounded by classmates who are ranked 0.13 percentile ranks higher at the high school graduation exam. Multiplying this by the effect of a one-percentile increase in peer ability for female students obtained in the restricted sample of programs with both star and standard classes of 0.14 (Table E40, Column (1)), we obtain the negative effect on admission to top STEM schools attributable solely to the presence of better-achieving peers in star classes: $-0.14 \times 0.13 = -0.0182$. We then divide this by the negative coefficient of star class for female

[46]Track fixed effects are interacted with program fixed effects, year fixed effects are interacted with program fixed effects, and year fixed effects are interacted with track fixed effects

[47]The three fixed effects cannot be interacted together since this interaction represents the level of variation that we are using.

students observed in Table 4, Column (2): $\frac{-0.0182}{-0.033} = 0.551$. This suggests that around 55 percent of the gender differential impact of star classes could be explained by the gender differential impact of having higher-achieving peers, with the rest attributable to other factors such as more demanding teachers, deeper exploration of the most challenging parts of the curriculum, etc.

7 Conclusion

We study the STEM elite higher education programs in France to learn about potential gender gaps in competitive entrance examinations. We provide evidence of the effects of preparing high-stakes exams in a competitive environment on the gender performance gap in a real-life setting. Despite self-selecting more than men into these programs, women benefit less from this selective and competitive environment. We confirm previous findings from the literature on the increased gender performance gap during high-stakes exams. We go beyond that by showing that an important part of the gender gap in performance is explained by the context of the learning environment.

Specifically, studying in a star class compared to a standard one appears to decrease women's chances of being admitted to a top STEM graduate school by about 15 to 25 percent, relative to their male counterparts. This holds true both *on average* — observed through a double difference analysis of male and female students in star and standard classes, while flexibly controlling for previous academic achievement — and *at the margin* of star class admission. At this margin, star classes appear to significantly boost male performance, tripling their admission chances to the top 10% most selective STEM graduate schools, but have a null or even negative effect on the performance of female students. This increased gender gap in performance translates into an increased gender gap in expected earnings among STEM workers just one year after graduation. We uncover that an important mechanism of this result is the differential effect of having higher-achieving peers. Male students appear very responsive to the presence of better-achieving peers, while female students are less so, thus increasing the gender gap in performance. The mechanisms that we unveil could be attributed to two explanations: male students being more willing to engage in competitive behavior with other male students, thereby increasing their performance when exposed to higher-achieving male peers; or male students helping each other more, acting as a "boys' club" mechanism (similar to those described in Cullen and Perez-Truglia (2023)). We cannot distinguish between these two potential explanations in our setting; further research would be needed to do so. We observe, however, that while the raw gender gap in access to the top programs is primarily due to the increasing gender performance gap during exam preparation, the gender gap specific to students in star classes is more closely related to performance differences on the exam day.

The entrance examinations to elite graduate schools are very competitive, thus their preparation as well, and this has a gender differential effect on students' achievement. Improving the representation of female students in the most selective STEM institutions requires to question the organization of their admission processes. Our results suggest that highly competitive admission processes incur equity costs, attributable to gender gaps in high-stakes entrance exams and their preparation. However, further research would be needed to determine the efficiency of selection based on competition before drawing conclusions about the overall effectiveness of these admission processes.

References

Arenas, Andreu and Caterina Calsamiglia, "Gender Differences in High-Stakes Performance and College Admission Policies," 2022. IZA Discussion Paper No. 15550.

Avery, Christopher N., Mark E. Glickman, Caroline M. Hoxby, and Andrew Metrick, "A Revealed Preference Ranking of U.S. Colleges and Universities," *The Quarterly Journal of Economics*, 2013, *128* (1), 425–467.

Azmat, Ghazala, Caterina Calsamiglia, and Nagore Iriberri, "Gender Differences in Response to Big Stakes," *Journal of the European Economic Association*, 2016, *14* (6), 1372–1400.

Bechichi, Nagui and Cosima Bluntz, "Les déterminants de la poursuite d'études en classe préparatoire aux grandes écoles: les enseignements de la procédure APB 2016," 2019. Note d'information du SIES, Ministère de l'Enseignement Supérieur et de la Recherche.

_, Julien Grenet, and Georgia Thebault, "D'Admission post-bac à Parcoursup: quels effets sur la répartition des néo-bacheliers dans les urrmations d'enseignement supérieur?," 2021. *France, portrait social - édition 2021*, collection « Insee Références », p. 105-121.

Beede, David N., Tiffany A. Julian, David Langdon, George McKittrick, Beethika Khan, and Mark E. Doms, "Women in STEM: A Gender Gap to Innovation," 2011.

Betts, Julian R., "The Economics of Tracking in Education," *Handbook of the Economics of Education*, 2011, *3* (11), 341–381.

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

Bonneau, Cécile, Pauline Charousset, Julien Grenet, and Georgia Thebault, *Quelle démocratisation des grandes écoles depuis le milieu des années 2000 ?*, Vol. 30, Institut des Politiques Publiques, 2021.

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

Bütikofer, Aline, Rita Ginja, Katrine V. Løken, and Fanny Landaud, "Higher-Achievement Schools, Peers and Mental Health," *The Economic Journal*, 2023, *133* (655), 2580–2613.

Cai, Xiqian, Yi Lu, Jessica Pan, and Songfa Zhong, "Gender Gap Under Pressure: Evidence From China's National College Entrance Examination," *Review of Economics and Statistics*, 2019, *101* (2), 249–263.

Calonico, Sebastian, Matias D. Cattaneo, Max H. Farrell, and Rocio Titiunik, "Rdrobust: Software for Regression-Discontinuity Designs," *The Stata Journal*, 2017, *17* (2), 372–404.

_ , _ , _ , and _ , "Regression Discontinuity Designs Using Covariates," *Review of Economics and Statistics*, 2019, *101* (3), 442–451.

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

Card, David and Laura Giuliano, "Can Tracking Raise the Test Scores of High-Ability Minority Students?," *American Economic Review*, 2016, *106* (10), 2783–2816.

Cattaneo, Matias D., Michael Jansson, and Xinwei Ma, "Manipulation Testing Based on Density Discontinuity," *The Stata Journal*, 2018, *18* (1), 234–261.

Chaisemartin, Clément De and Luc Behaghel, "Estimating the Effect of Treatments Allocated by Randomized Waiting Lists," *Econometrica*, 2020, *88* (4), 1453–1477.

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

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

Delaney, Judith M. and Paul J. Devereux, "Gender Differences in STEM Persistence After Graduation," *Economica*, 2022, *89* (356), 862–883.

DEPP, "Résultats Définitifs de la Session 2021 du Baccalauréat," *Note d'Information*, 2022, *n°22.10.*

Duflo, Esther, Pascaline Dupas, and Michael Kremer, "Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence From a Randomized Evaluation in Kenya," *American Economic Review*, 2011, *101* (5), 1739–1774.

Ellis, Jessica, Bailey K. Fosdick, and Chris Rasmussen, "Women 1.5 Times More Likely To Leave Stem Pipeline After Calculus Compared to Men: Lack of Mathematical Confidence a Potential Culprit," *PLoS ONE*, 2016, *11* (7), 1–14.

Fack, Gabrielle, Julien Grenet, and Yinghua He, "Beyond Truth-Telling: Preference Estimation with Centralized School Choice and College Admissions," *American Economic Review*, 2019, *109* (4), 1486–1529.

Farré, Lídia and Francesc Ortega, "Selecting talent: Gender differences in success in competitive selection processes," *Journal of Human Resources*, 2023, *58* (6), 1881–1913.

Flory, Jeffrey A., Uri Gneezy, Kenneth L. Leonard, and John A. List, "Gender, Age, and Competition: A Disappearing Gap?," *Journal of Economic Behavior and Organization*, 2018, *150*, 256–276.

Fort, Margherita, Andrea Ichino, Enrico Rettore, and Giulio Zanella, "Multi-Cutoff RD Designs with Observations Located at Each Cutoff: Problems and Solutions," 2022. CEPR Discussion Paper No. DP16974.

Gale, D. and Lloyd S. Shapley, "College Admissions and the Stability of Marriage," *The American Mathematical Monthly*, 1962, *69* (1), 9–15.

Gneezy, Uri, Kenneth L. Leonard, and John A. List, "Gender Differences in Competition: Evidence from a Matrilineal and a Patriarchal Society," *Econometrica*, 2009, *77* (5), 1637–1664.

_ , Muriel Niederle, and Aldo Rustichini, "Performance in Competitive Environments: Gender Differences," *The Quarterly Journal of Economics*, 2003, *118* (3), 1049–1074.

Griffith, Amanda L., "Persistence of Women and Minorities in STEM Field Majors: Is It the School That Matters?," *Economics of Education Review*, 2010, *29* (6), 911–922.

Hansen, Bruce E., "Sample Splitting and Threshold Estimation," *Econometrica*, 2000, *68* (3), 575–603.

Hedges, Larry V. and Amy Nowell, "Sex Differences in Mental Test Scores, Variability, and Numbers of High-Scoring Individuals," *Science*, 1995, *269* (5220), 41–45.

Hoekstra, Mark, "The Effect of Attending the Flagship State University on Earnings: A Discontinuity-Based Approach," *The Review of Economics and Statistics*, 2009, *91* (4), 717–724.

Iriberri, Nagore and Pedro Rey-Biel, "Competitive Pressure Widens the Gender Gap in Performance: Evidence From a Two-Stage Competition in Mathematics," *The Economic Journal*, 2019, *129* (620), 1863–1893.

Johnson, Wendy, Andrew Carothers, and Ian J. Deary, "A Role for the X Chromosome in Sex Differences in Variability in General Intelligence?," *Perspectives on Psychological Science*, 2009, *4* (6), 598–611.

Jurajda, Štěpán and Daniel Münich, "Gender Gap in Performance Under Competitive Pressure: Admissions to Czech Universities," *American Economic Review, Papers and Proceedings*, 2011, *101* (3), 514–518.

Kahn, Shulamit and Donna Ginther, "Women and Science, Technology, Engineering, and Mathematics (STEM): Are Differences in Education and Careers Due to Stereotypes, Interests, or Family?," in "The Oxford Handbook of Women and the Economy," Oxford University Press, 2018.

Kugler, Adriana D., Catherine H. Tinsley, and Olga Ukhaneva, "Choice of Majors: Are Women Really Different From Men?," *Economics of Education Review*, 2021, *81*, 102079.

Landaud, Fanny and Éric Maurin, "Aim High and Persevere ! Competitive Pressure and Access Gaps in Top Science Graduate Programs," 2020. PSE Working paper No 2020 - 79.

_ and _ , "Tracking When Ranking Matters," 2022. THEMA Working Paper n°2023-12.

_ , Son Thierry Ly, and Éric Maurin, "Competitive Schools and the Gender Gap in the Choice of Field of Study," *Journal of Human Resources*, 2020, *55* (1), 278–308.

Machin, Stephen and Tuomas Pekkarinen, "Global Sex Differences in Test Score Variability," *Science*, 2008, *322* (5906), 1331–1332.

Montalbán, José and Almundena Sevilla, "Gender Differences in Performance: The Role of External Testing Environments," 2023. Working Paper.

Montolio, Daniel and Pere A. Taberner, "Gender Differences Under Test Pressure and Their Impact on Academic Performance: A Quasi-Experimental Design," *Journal of Economic Behavior & Organization*, 2021, *191*, 1065–1090.

Morin, Louis-Philippe, "Do Men and Women Respond Differently to Competition? Evidence From a Major Education Reform," *Journal of Labor Economics*, 2015, *33* (2), 443–491.

Niederle, Muriel and Lise Vesterlund, "Do Women Shy Away From Competition? Do Men Compete Too Much?," *Quarterly Journal of Economics*, 2007, *122* (3), 1067–1101.

_ and _ , "Gender and Competition," *Annual Review of Economics*, 2011, *3*, 601–630.

OECD, *Education at a Glance 2019*, OECD Publishing, 2019.

Ors, Evren, Frédéric Palomino, and Eloic Peyrache, "Performance Gender Gap: Does Competition Matter?," *Journal of Labor Economics*, 2013, *31* (3), 443–499.

Parodi, Maxime, Hélène Périvier, and Fabrice Larat, "De Sciences Po à l'ENA, la voie étroite vers les sommets de la fonction publique," 2022. Working Paper OFCE n°15/2022.

Pekkarinen, Tuomas, "Gender Differences in Behaviour Under Competitive Pressure: Evidence on Omission Patterns in University Entrance Examinations," *Journal of Economic Behavior & Organization*, 2015, *115*, 94–110.

Porter, Jack and Ping Yu, "Regression Discontinuity Designs With Unknown Discontinuity Points: Testing and Estimation," *Journal of Econometrics*, 2015, *189* (1), 132–147.

Saccardo, Silvia, Aniela Pietrasz, and Uri Gneezy, "On the Size of the Gender Difference in Competitiveness," *Management Science*, 2018, *64* (4), 1541–1554.

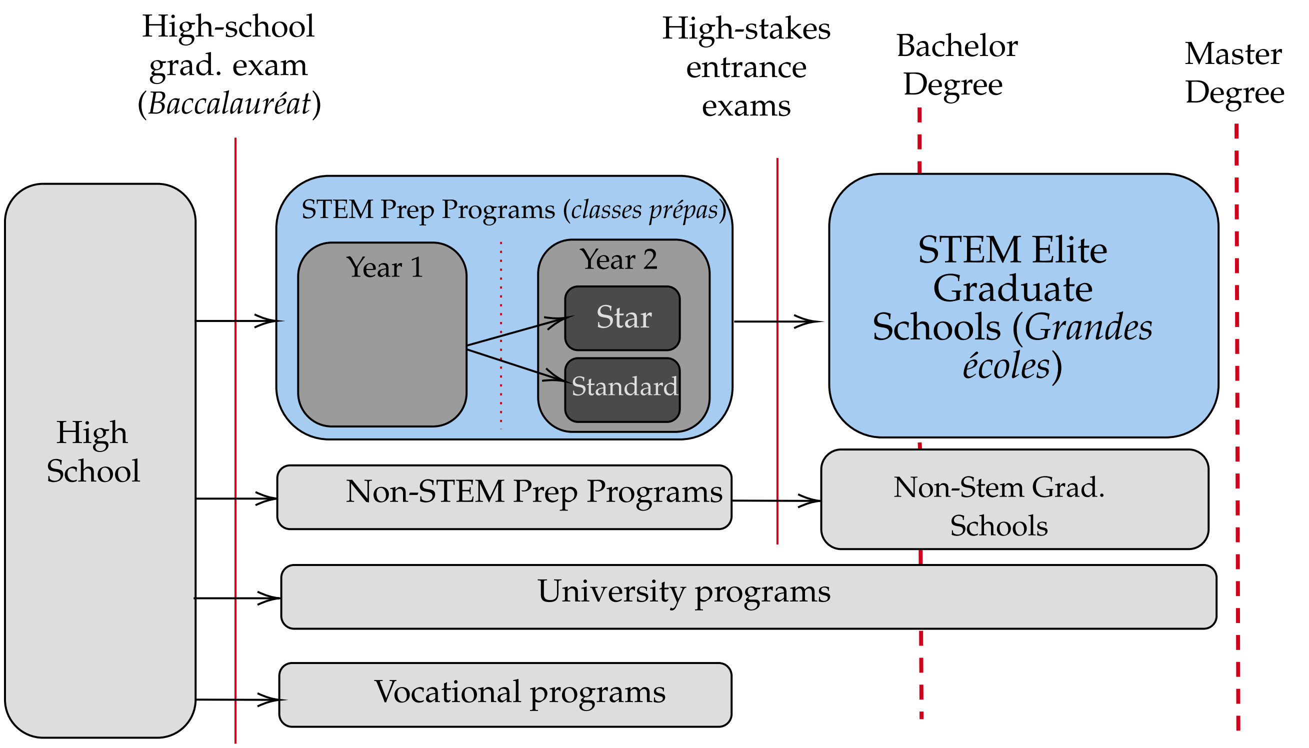
Schlosser, Analia, Zvika Neeman, and Yigal Attali, "Differential Performance in High Versus Low Stakes Tests: Evidence From the GRE Test," *The Economic Journal*, 2019, *129* (623), 2916–2948.

Sousa, José De and Guillaume Hollard, "From Micro to Macro Gender Differences: Evidence from Field Tournaments," *Management Science*, 2023, *69* (6), 3358–3399.

Appendix

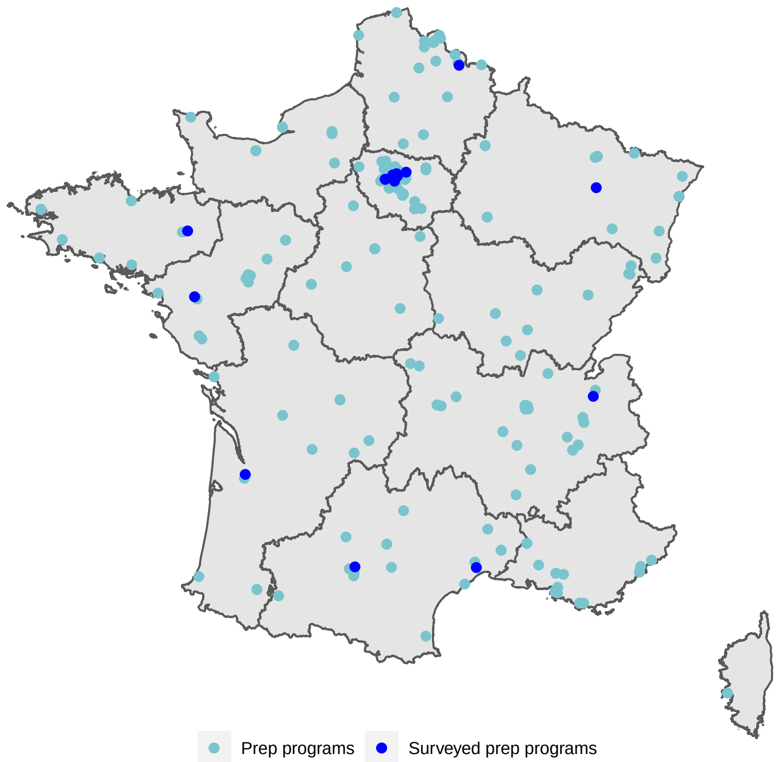
A1 Detailed Institutional Background

Figure A1: Simplified Diagram of French Higher Education

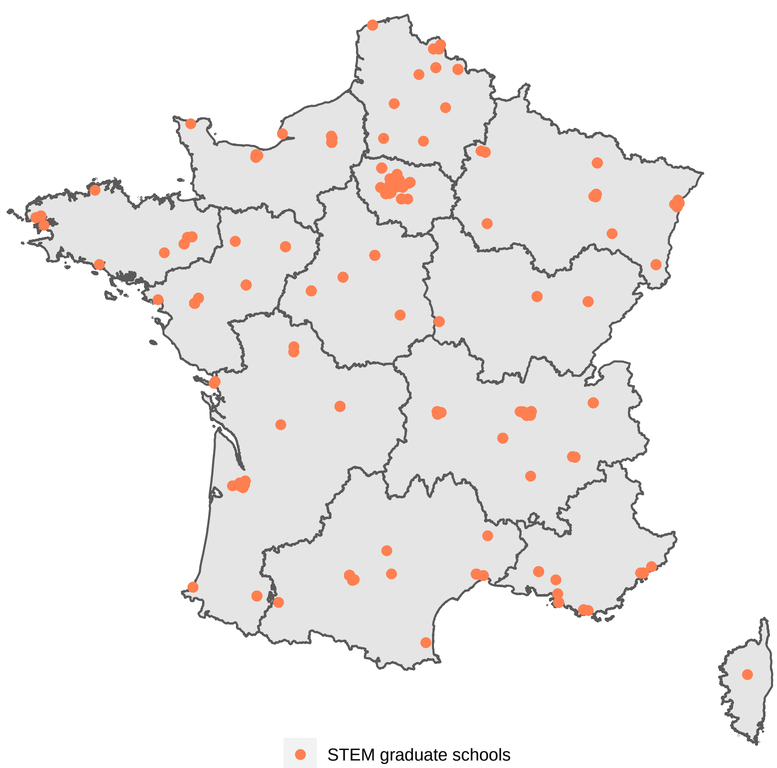


Notes: This diagram shows the organization of France's main higher education tracks. The programs we study here are STEM prep programs and STEM elite graduate schools, highlighted in blue.

Figure A2: Maps of Preparatory Programs and STEM Graduate Schools



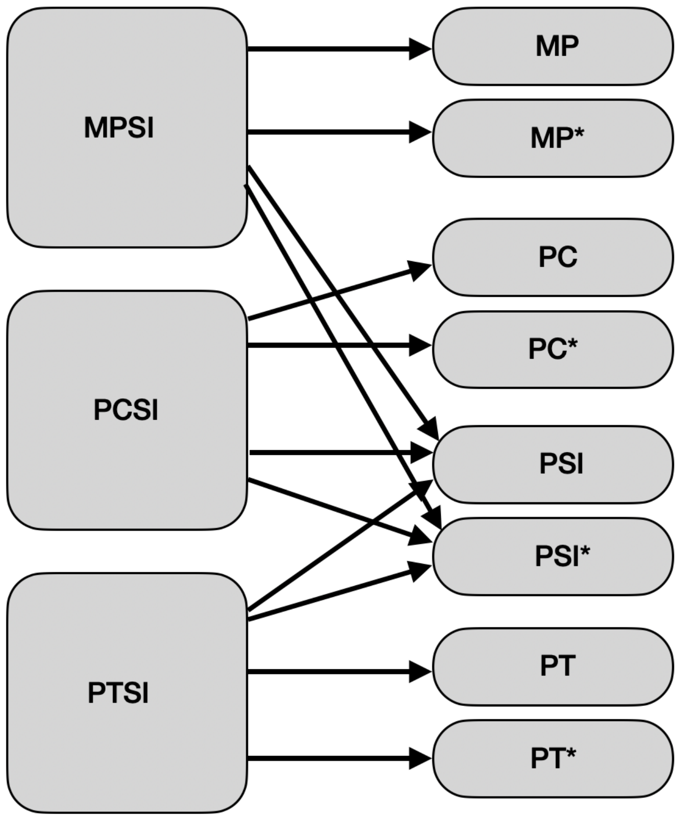
(a) STEM Preparatory Programs



(b) STEM Graduate Schools

Notes: These maps illustrate the geographical distribution across France of STEM preparatory programs (Panel a) and of STEM graduate schools (Panel b).

Figure A3: Organization Chart of Science CPGE Programs



Notes: This diagram illustrates the various tracks within the science preparatory programs. The boxes positioned on the left represent the first year of the program, while those on the right depict the second year. Although there are other science prep classes (such as *BCPST*, *TB*, *TPC*, *TSI*), they are excluded from this study due to the absence of tracking between standard and star classes in their second year.

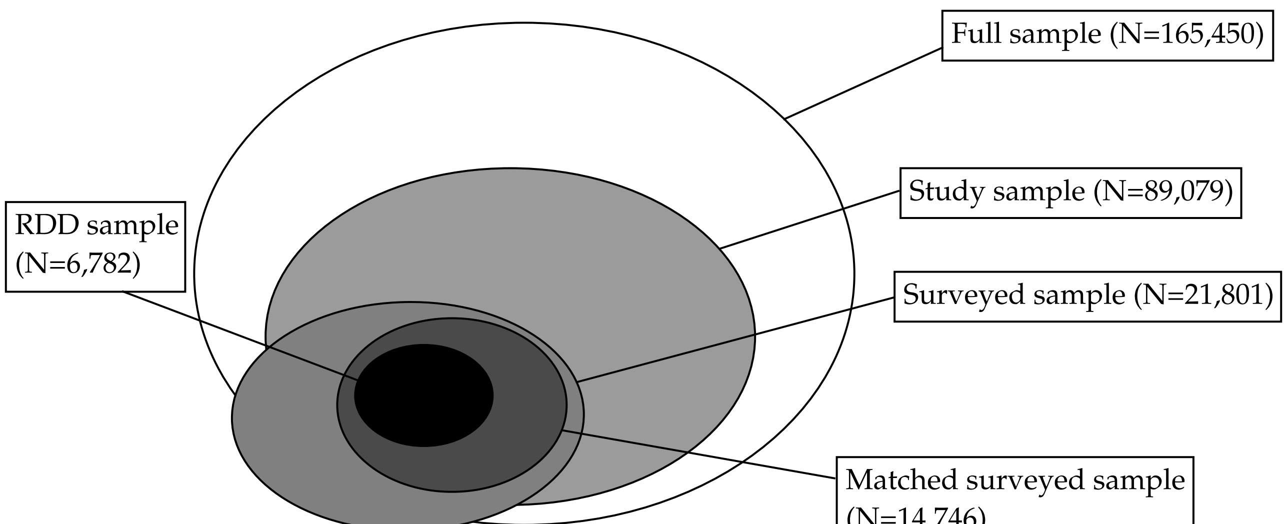
Table A1: Most Selective Competitive Exam Coefficients

	Track 1 (MP)	Track 2 (PC)	Track 3 (PSI)	Track 4 (PT)
Math.	0.36	0.24	0.28	0.18
Physics & Chemistry	0.31	–	0.32	0.18
Physics	–	0.31	–	–
Chemistry	–	0.17	–	–
Engineering Science	0.03	–	0.13	0.31
Computer Science	0.06	0.03	0.04	0.06
French	0.13	0.13	0.11	0.15
Foreign Language	0.12	0.12	0.13	0.13

Notes: This table presents the coefficients from competitive exams used to calculate the high school graduation exam weighted GPA. These coefficients are based on the mean values from the three most selective exam clusters (*Banque X-ENS*, *Banque Mines-Ponts* and *Banque Centrale*). When particular high school grades, such as computer science or engineering science, are missing, the coefficients are modified to account for the absence of these subjects.

B2 Complementary Descriptive Statistics

Figure B4: Diagram of the Different Samples of Analysis



Notes: This diagram depicts the various subsets of data used in our study. The 'full sample' corresponds to the universe of applicants to STEM graduate schools from 2015 to 2023, excluding the biology track. Only applicants enrolled in prep programs located in France are included. We retain only those prep programs that were present throughout the nine-year period covered by our data and had more than 10 students per year. The 'study sample' corresponds to the restriction of the full sample to combinations of *program* × *track* that have both a standard and a star class in the second year of the program, which we use in our main analysis concerning star classes. The 'surveyed sample' corresponds to the data on programs we surveyed to collect data on grades earned during the prep programs. The location of these programs can be found on the map in Figure A2. Part of the individuals surveyed were not included in our full sample, that is why the two circles do not fully overlap. Specifically, it concerns students who dropped out of programs after the first year and those who were not in *program* × *track* with both standard and star classes, despite these programs being specifically targeted during our data collection. The 'matched surveyed sample' corresponds to students that we surveyed and were able to statistically match with the administrative data from applications to STEM graduate schools. One prep program opposed this statistical matching. They are thus not matched, but are included in the analyses which only use surveyed prep programs data. The 'RDD sample' corresponds to the sample used for the regression discontinuity analysis, where we only include students for whom we could link to the administrative data on school admission, who had recorded grades during their first year in the prep programs, who were in a class in which the R^2 of the estimation to identify the cutoff was above 0.75 and who were within the vicinity of [-18;+18] of the identified threshold for star class admission.

Table B2: Descriptive Statistics for Male and Female Students (2015-2023)

	Sample		
	All Students (1)	Male Students (2)	Female Students (3)
A. Students			
Female	0.25	0.00	1.00
Age	19.6	19.6	19.5
	(0.8)	(0.8)	(0.7)
Need-based scholarship holder	0.28	0.28	0.29
Mother is high SES	0.48	0.48	0.49
Father is high SES	0.62	0.62	0.62
High School Graduation Exam			
Highest honors	0.50	0.47	0.59
High honors	0.32	0.34	0.27
Honors	0.14	0.16	0.11
Percentile rank	0.78	0.77	0.82
	(0.18)	(0.18)	(0.18)
Percentile rank re-weighted with comp. exams. coeffs.	0.87	0.87	0.88
	(0.16)	(0.16)	(0.16)
From Paris	0.08	0.08	0.09
From Parisian area - outside Paris	0.21	0.20	0.21
Star class	0.36	0.36	0.33
Repeater	0.19	0.20	0.16
In a prep program in Paris	0.18	0.18	0.20
In a prep program in Parisian area - outside Paris	0.13	0.13	0.13
B. Prep Programs			
Number of prep programs	200	200	197
in Paris	23	23	23
in Parisian area (outside Paris)	31	31	30
Number of classes	546	546	542
Star classes	185	185	184
Number of students	165,450	124,272	41,178

Notes: This table displays the descriptive statistics for all the students in our sample, and then divide our sample between male and female students. The sample is constructed from the SCEI administrative datasets, which cover the entire universe of applicants to elite STEM graduate schools from 2015 to 2023, excluding the biology track. Only applicants enrolled in prep programs located in France are included. We retain only those prep programs that were present throughout the nine-year period covered by our data and had more than 10 students per year. Age is the age of the candidates when taking the competitive entrance exam. For socio-economic status, we rely on the Department of Education's statistical service (DEPP), which classifies occupations into four groups by socio-economic status (SES): High SES includes white-collar professionals, managers, CEOs, teachers, and artists; Upper-middle SES includes intermediate occupations, technicians, foremen, and supervisors; Lower-middle SES includes farmers, artisans, shopkeepers, and white-collar employees; Low SES includes blue-collar workers and non-working people. The SES of the child's legal representative is used for classification. Percentile rank at the high school graduation exam are computed using the scores re-weighted with the coefficients of the most selective competitive entrance exams to STEM graduate schools.

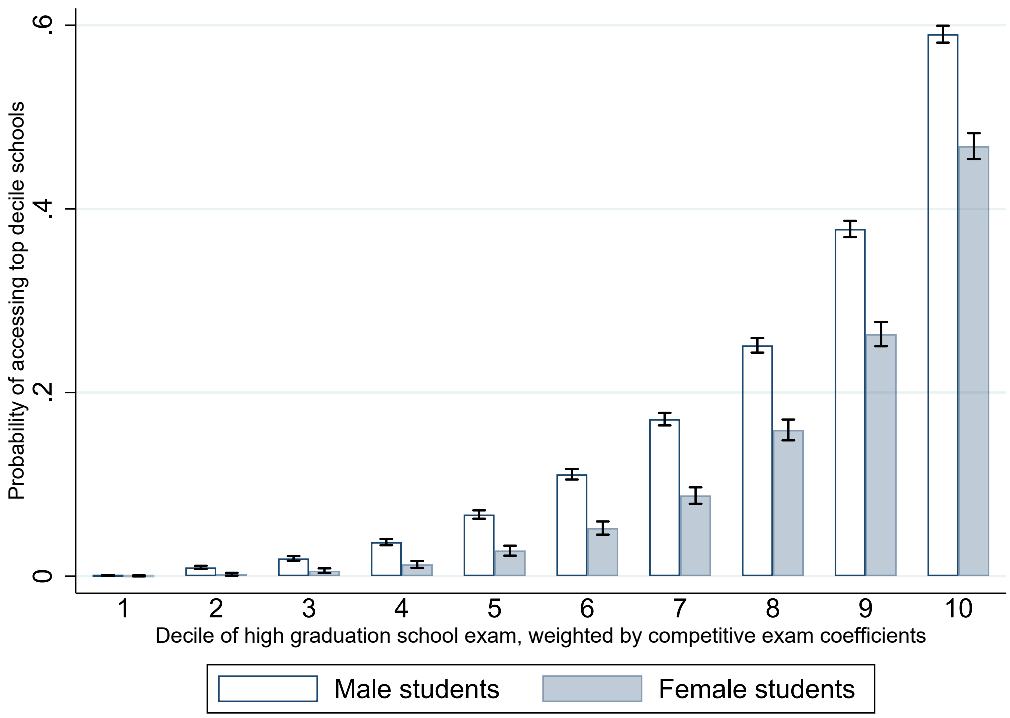
Table B3: Comparison of Samples (2015-2023)

	Sample		
	All (1)	Study (2)	Survey (3)
A. Students			
Female	0.25	0.26	0.26
Age	19.6	19.5	19.5
	(0.8)	(0.8)	(0.7)
Need-based scholarship holder	0.28	0.26	0.25
Mother is high SES	0.48	0.52	0.55
Father is high SES	0.62	0.67	0.71
High School Graduation Exam			
Highest honors	0.50	0.64	0.73
High honors	0.32	0.27	0.20
Honors	0.14	0.08	0.06
From Paris	0.08	0.09	0.12
From Parisian area - outside Paris	0.21	0.19	0.20
Star class	0.36	0.47	0.51
Repeater	0.19	0.18	0.15
In a prep program in Paris	0.18	0.23	0.27
In a prep program in Parisian area - outside Paris	0.13	0.10	0.17
Year of Exams			
2015	0.11	0.11	0.07
2016	0.11	0.11	0.10
2017	0.11	0.11	0.11
2018	0.11	0.11	0.13
2019	0.11	0.11	0.14
2020	0.11	0.11	0.19
2021	0.11	0.11	0.22
2022	0.11	0.11	0.21
2023	0.10	0.10	0.06
B. Prep Programs			
Number of prep programs	200	66	18
in Paris	23	13	5
in Parisian area (outside Paris)	31	7	4
Number of classes	546	273	80
Star classes	185	133	43
Number of students	165,450	89,079	14,746

Notes: This table compares the students across our three samples. In Column (1), we show the sample constructed from the SCEI administrative datasets. This dataset covers the whole universe of applicants to elite STEM graduate schools from 2015 to 2023, excluding the biology track. Only applicants enrolled in prep programs located in France are included. We retain only those prep programs that were present throughout the nine-year period covered by our data and had more than 10 students per year. In Column (2), we display our study sample, a subset of the sample in Column (1). This subset only includes *programs* $\times$ *track* that have both a standard class and a star class. In Column (3), we show the subsample of prep programs that we surveyed in 2022 and 2023 to gather data on students' grades during the prep programs.

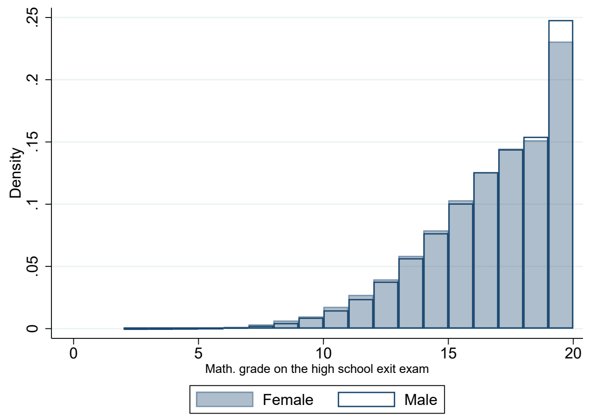
B2.1 Previous Academic Achievement

Figure B5: Probability of Gaining Admission to a Top 10% Most Selective STEM Graduate School, by High School Graduation Exam Results and Gender

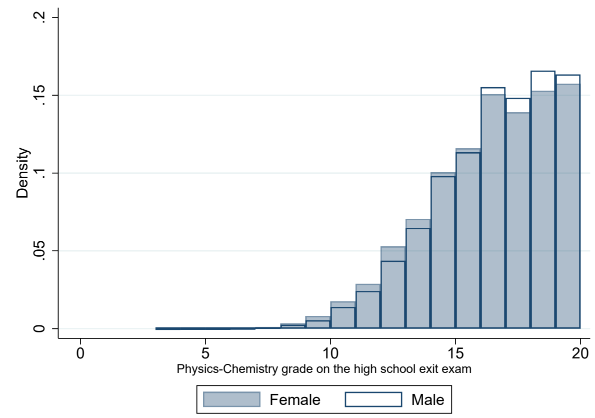


Notes: This figure illustrates the probability of gaining admission to elite STEM schools within the top 10% most selective ones, by gender and prior academic performance. Prior academic achievement is measured through results obtained at the high school graduation exam, with weights based on coefficients of the entrance examination to the most selective STEM graduate schools (refer to Table A1).

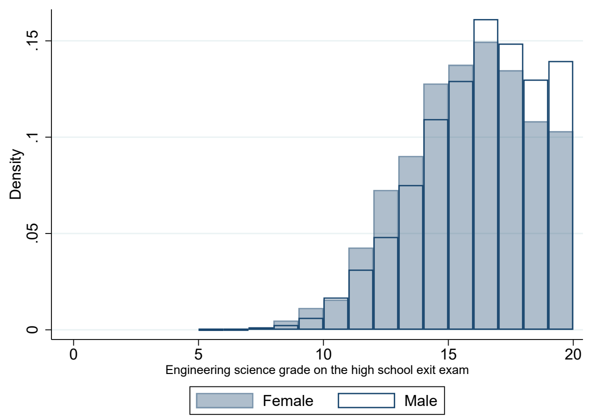
Figure B6: Density of Raw Grades at the High School Graduation Exam of CPGE Students, by Gender



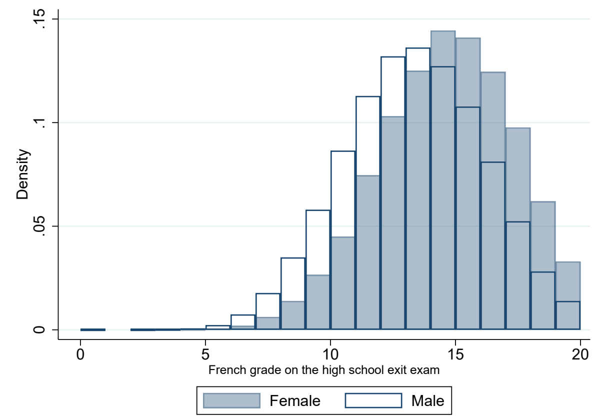
(a) Math.



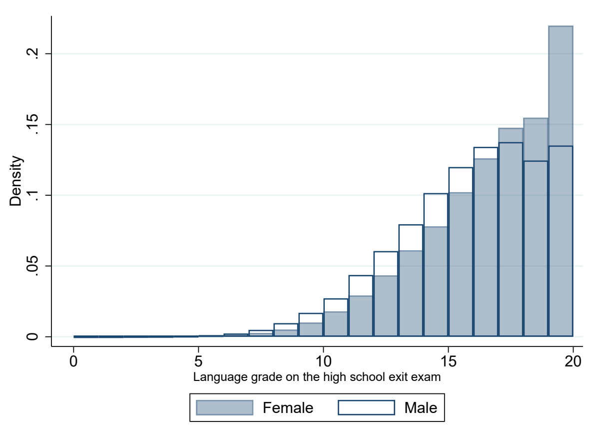
(b) Physics and Chemistry



(c) Engineering Science



(d) French



(e) Foreign Language

Notes: These histograms depict the distribution of raw grades achieved at the high school graduation exam by prep program students and by gender. Grades in engineering science are available for only approximately 20 percent of the entire sample; this is because the majority of students do not take up engineering studies in high school.

B2.2 School Selectivity

Table B4: STEM Graduate Schools in the Top 10% of Selectivity or Desirability

	Selectivity				Desirability (all applicants)				Desirability (female applicants)			
	MP	PC	PSI	PT	MP	PC	PSI	PT	MP	PC	PSI	PT
Centrale Lyon	X	X	X	X	X	X	X	X	X	X	X	X
CentraleSupélec	X	X	X	X	X	X	X	X	X	X	X	X
ISAE - SUPAERO Toulouse	X	X	X	X	X	X	X	X	X	X	X	X
Mines de Paris	X	X	X	X	X	X	X	X	X	X	X	X
Polytechnique	X	X	X	X	X	X	X	X	X	X	X	X
Ponts ParisTech	X	X	X	X	X	X	X	X	X	X	X	X
ENSTA ParisTech	X	X	X	X	X	X	X	X	X	X	X	
ENS Ulm	X	X	X		X	X	X		X	X		
ENSAE Paris	X	X	X									
ENS de Lyon	X	X			X	X			X	X		
Télécom ParisTech	X		X		X	X	X		X	X	X	
Mines de Nancy	X			X			X				X	
Ecole Météorologie	X											
Centrale Nantes		X	X	X	X	X	X	X	X	X	X	X
ESPCI Paris		X				X				X		
Chimie ParisTech		X										
ENS Cachan Paris-Saclay			X		X	X			X	X		
Centrale Lille					X		X	X	X		X	X
Mines de Saint-Etienne											X	
Arts et Métiers												X
Total number of schools	130	134	128	96	130	134	128	96	130	134	128	96

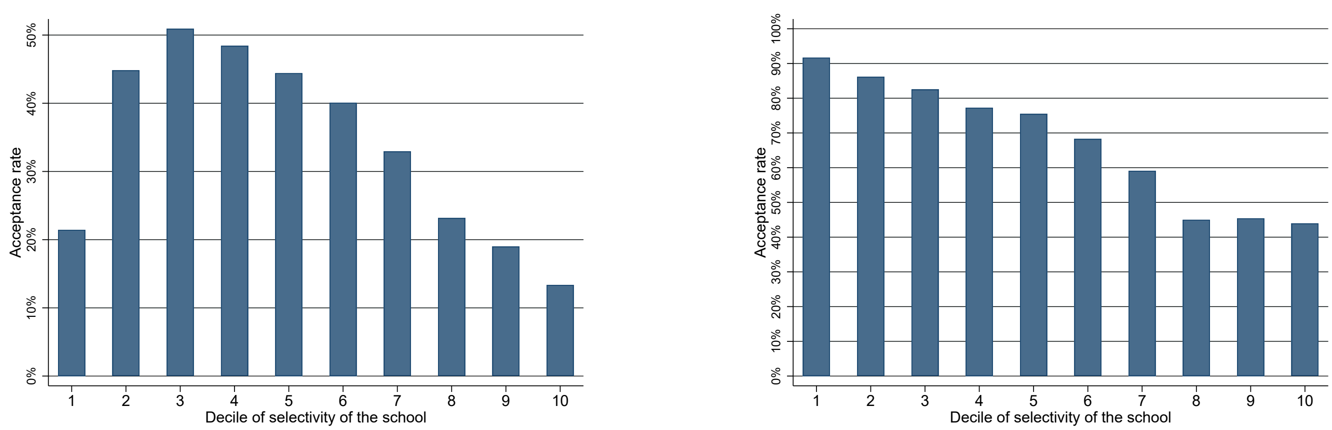
Notes: This table enumerates the STEM graduate schools which are in the top 10% of selectivity, desirability and desirability for female students. In the first four columns, selectivity of schools is measured using the average percentile rank of admitted students at the high school graduation exam. The subsequent eight columns use the revealed preferences of applicants as a metric for desirability, following the approach of Avery et al. (2013). This method captures the desirability of STEM graduate schools, where more sought-after institutions garner more applications and more often appear higher on students' rank-ordered lists. The final four columns consider only female applicants when calculating school desirability. For all the definitions, they are consistently defined on a track-by-track basis, acknowledging that not all tracks lead to the exact same STEM graduate schools, even if many overlap.

Table B5: Average Earnings of STEM Elite Grad. School Graduates, One Year After Graduation, by School Selectivity Decile

Decile of selectivity	Median earnings (male alumni)	Median earnings (female alumni)	Median earnings (all alumni)
1	39,395	38,383	39,096
2	38,799	38,099	38,617
3	36,501	35,682	36,246
4	37,822	36,847	37,516
5	37,501	37,859	37,604
6	36,549	36,178	36,458
7	36,653	36,225	36,541
8	37,420	37,016	37,321
9	39,557	38,924	39,390
10	42,872	41,912	42,662

Notes: This table displays the average median earnings, by decile of selectivity, of STEM graduate schools segmented by gender for students who graduated between 2018 and 2023. Salaries are determined based on each STEM graduate school's median from the previous year and include bonuses. Alumni currently enrolled in doctoral programs are excluded from this analysis. If a salary was missing for a specific year, we inferred it from the adjacent cohort. For the most recent cohorts that haven't yet graduated, we inferred the salary based on the previous cohorts and the average annual salary increase for that specific school. When gender-segregated median salaries were not available, the general median was applied to both male and female graduates. It's noteworthy that some students from science preparatory classes choose to study at non-engineering STEM graduate schools, such as *ENS Paris*, *ENS Paris-Saclay*, *ENS Lyon*, and *ENS Rennes*, or STEM institutions abroad. We don't have earnings information for these institutions. In total, earnings data is missing for 5.6% of the schools and for 3.8% of individuals who are admitted to a STEM graduate school. The data comes from SCEI (2015-2023) and the CTI wage data (2018-2022), available on the CTI website.

Figure B7: Acceptance Rate, by Grad. School Selectivity Decile

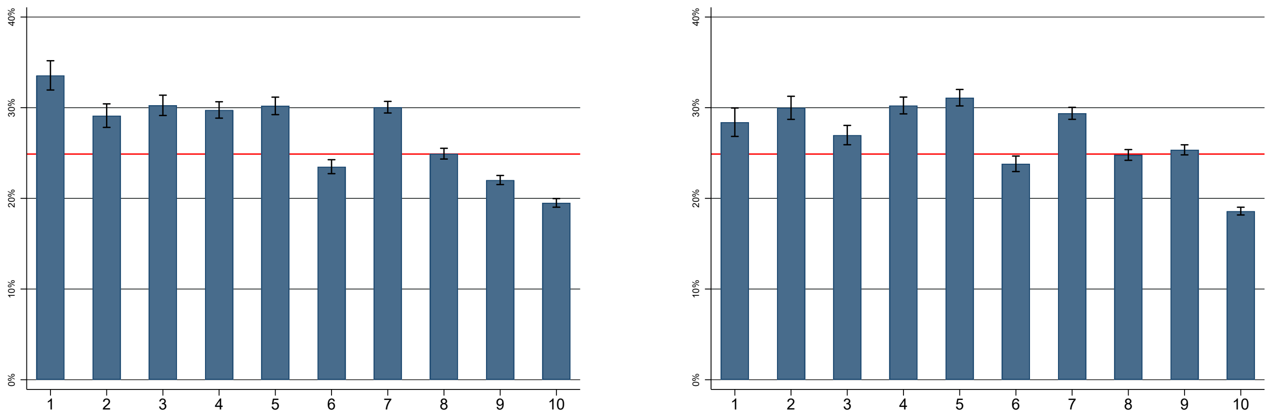


(a) All applications

(b) Students ranked at the entrance exam

Notes: These figures illustrates the acceptance rate of schools, by schools selectivity decile. The acceptance rate is computed as the ratio of candidates who met or exceeded the admission criteria — meaning their rank at the entrance exam was below the admission threshold — over the overall number of applicants to the school (Panel a) or the number of applicants ranked at one of the entrance exam to the school (Panel b). Acceptance rate captures offers of admission and not actual enrollment. Consequently, while a student could theoretically receive offers from multiple schools, they will ultimately register at only one institution.

Figure B8: Proportion of Female Students in STEM graduate schools, by Decile of Desirability of STEM Graduate Schools

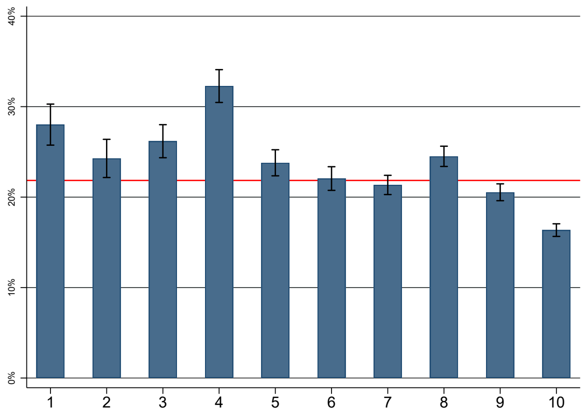


(a) All applicants

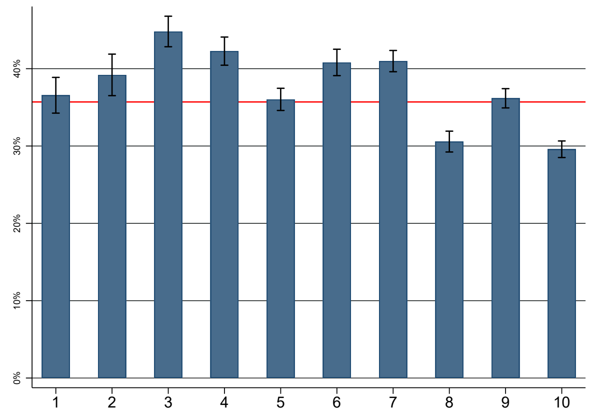
(b) Female applicants only

Notes: This figure shows the proportion of female students in STEM graduate schools, by decile of desirability of STEM graduate schools. The average number of female students in these schools is represented with the red line. Desirability of schools is defined track-by-track based on students' revealed preferences (Avery et al. (2013)). This method captures the desirability of STEM graduate schools because more sought-after institutions garner more applications and more often appear higher on students' rank-ordered lists. Figure (a) uses all application to define school desirability while Figure (b) retains application of female students only to define school desirability.

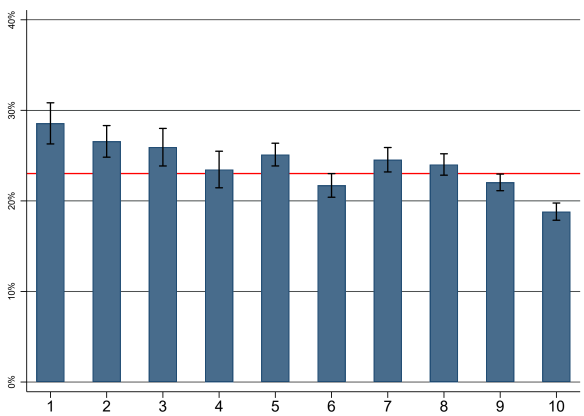
Figure B9: Proportion of Female Students, by Decile of Selectivity of STEM Graduate Schools and Tracks



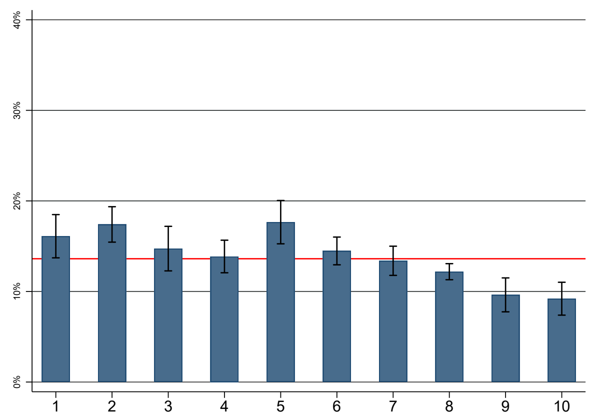
(a) Track 1 (MP)



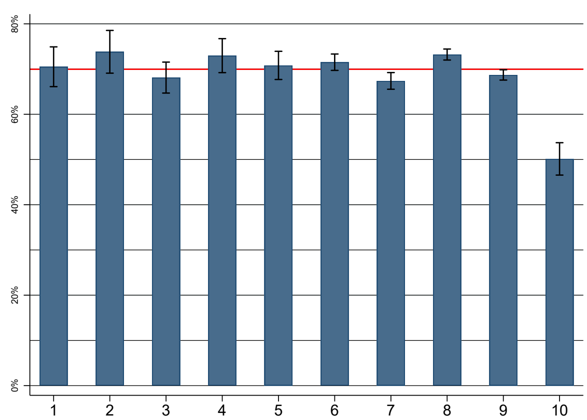
(b) Track 2 (PC)



(c) Track 3 (PSI)



(d) Track 4 (PT)

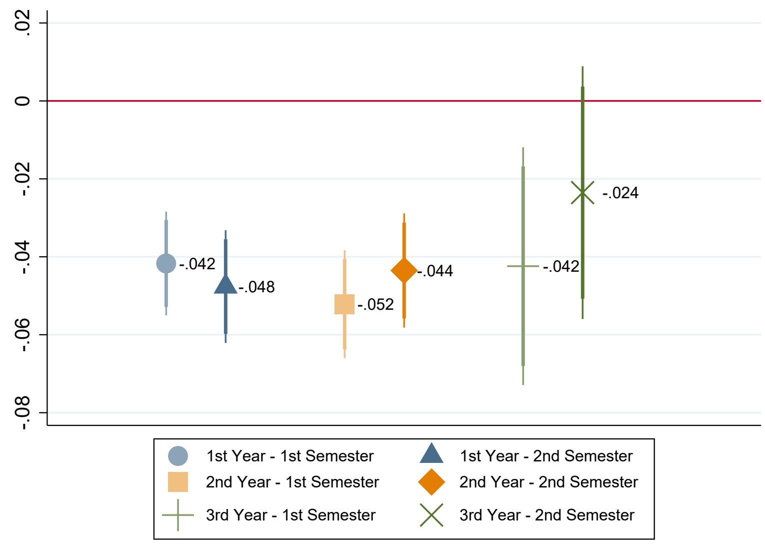


(e) Track 5 (BCPST)

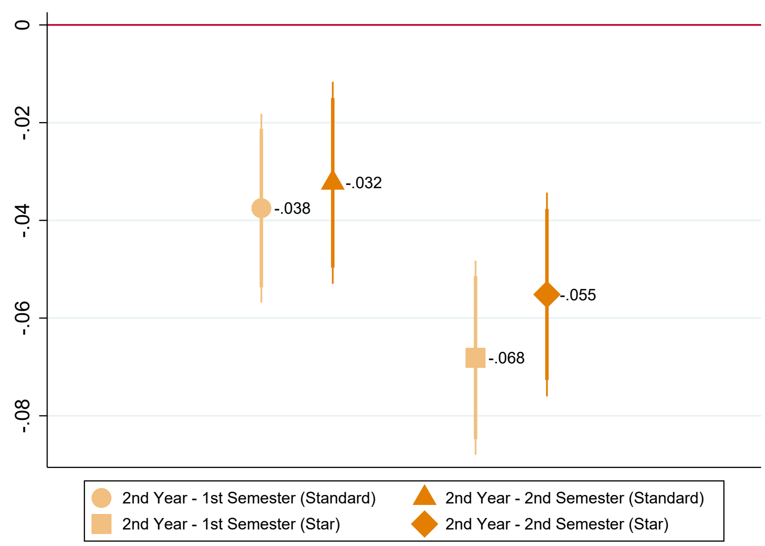
Notes: This figure depicts the proportion of female students in STEM graduate schools, by decile of school selectivity, across the four main tracks under study, as well as the biology track. The average number of female students in these schools is represented with the red line. Selectivity of schools is measured using the average percentile rank of admitted students at the high school graduation exam and is defined on a track-by-track basis. The fifth track represents the biology track. In this track, female students constitute the majority of students (70%). However, it is excluded from the primary analysis because, unlike the other four tracks, the biology track does not differentiate between star and non-star classes in the second year.

B2.3 Performance During Prep Programs

Figure B10: Gender Gap in the Probability of Being in the Top Quintile of the Class



(a) First, Second and Third Years



(b) Focus on the Second Year, by Class Type

Notes: This figure displays the gender gap in the probability of being among high-achievers during prep programs (defined as being in the top quintile of the class. Indeed, the top 10% of most selective STEM graduate schools actually account for roughly 20% of students since they are bigger than the less selective schools.). To rank students, we use their weighted GPA, which was either provided directly by the prep programs or computed using students' grades in each subject and competitive exam coefficients between subjects.
N = 21,801.

B2.4 Complementary Decomposition Results

Table B6: Full Sample: Decomposition Gender Gap in Access to Top 10% Most Selective Schools

	Receive an offer from top 10% most selective graduate schools					
	(1)	(2)	(3)	(4)	(5)	(6)
Female student	-0.045***	-0.051***	-0.037***	-0.026***	-0.001	-0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)
Decomposition		-13%	31%	24%	56%	0%
% explained by additionnal control						
Controls:						
Demographics and Fixed-effects	✓	✓	✓	✓	✓	✓
Demographic characteristics						
Program, track, and year FE						
Previous ability		✓	✓	✓	✓	✓
HS & MS grade FE						
Prep program ability			✓	✓	✓	✓
Star class						
Applications				✓	✓	✓
Apply to top school exams						
D-Day Effect					✓	✓
Percentile rank at top school exams						
Preferences						✓
Top schools in ROLs						
N	165,418	165,418	165,418	165,418	165,418	165,418
Adj-R^2	0.306	0.358	0.416	0.451	0.716	0.724

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: This table shows the evolution of the gender gap in access to the top 10% most selective graduate schools when adding various controls. We use the full sample of students taking high-stakes entrance exams in the SCEI data. Table 2 displays the same decomposition on the subsample of students for whom we have prep program grades, thereby controlling for prep program ability. The decomposition raw is computed by observing how much the gender gap is reduced when an additional control is added, compared to the raw gender gap observed in the first column. Selectivity of schools is defined by the average percentile rank of admitted students at the high school graduation exam. Demographic controls include geographic origin (Paris or the Parisian area), low-income status, the socio-economic status of each parent (categorized into four groups), French nationality, repeater status, disability status, whether the student was in a science academic track during high school, the student's option (either engineering science or computer science), and the gender composition in the *track* $\times$ *program*. We include year, track, and program fixed effects. Previous ability is controlled by (i) decile rank dummy variables from both high school and middle school graduation exams' GPAs, and (ii) quintile rank dummy variables for each grade at the high and middle school graduation exams in the subjects studied in prep programs—mathematics, physics and chemistry, engineering science, French (both written and oral), and foreign languages. Prep programs' ability controls account for whether or not the applicants are in a star class, which is the only variable we have regarding academic performance in the prep program for the full sample. Applications to top graduate schools are controlled by dummies for application to each of the exams leading to top graduate schools, defined track by track. Performance during competitive entrance exams is controlled for by (i) dummies for whether or not the applicant is ranked at competitive exams leading to top graduate schools and (ii) the percentile rank of their rank in those exams, controlled linearly for each entrance exam. Preferences are controlled for by adding dummy variables for whether or not each of the top graduate schools is ranked among students' ranked-ordered lists of schools. The order of these variables is chosen to follow the chronological order of decisions made by students. Standard errors are clustered at the *track* $\times$ *program* $\times$ *cohort* level.

Table B7: Full Sample: Decomposition Gender Gap in Star Class in Access to Top 10% Most Selective Schools

	Receive an offer from top 10% most selective graduate schools					
	(1)	(2)	(3)	(4)	(5)	(6)
Female student × Star class	-0.039***	-0.046***	-0.046***	-0.040***	-0.002	-0.004*
	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)
Decomposition		-18%	0%	15%	97%	-5%
% explained by additionnal control						
Controls:						
Demographics and Fixed-effects	✓	✓	✓	✓	✓	✓
Star class						
Female student						
Demographic controls						
Program, track, and year FE						
Previous ability		✓	✓	✓	✓	✓
HS & MS grade FE						
Prep program ability			✓	✓	✓	✓
Star class						
Applications				✓	✓	✓
Apply to top school exams						
D-Day Effect					✓	✓
Percentile rank at top school exams						
Preferences						✓
Top schools in ROLs						
N	165,418	165,418	165,418	165,418	165,418	165,418
Adj-R^2	0.391	0.417	0.417	0.451	0.716	0.724

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: This table shows the evolution of the gender-specific gap in star classes in access to the top 10% most selective graduate schools when adding various controls. We include controls for star class and female students in all specifications. We use the full sample of students taking high-stakes entrance exams in the SCEI data. Table 8 displays the same decomposition on the subsample of students for whom we have prep program grades, thereby controlling for prep program ability. The decomposition raw is computed by observing how much the gender gap is reduced when an additional control is added, compared to the raw gender gap observed in the first column. Selectivity of schools is defined by the average percentile rank of admitted students at the high school graduation exam. Demographic controls include geographic origin (Paris or the Parisian area), low-income status, the socio-economic status of each parent (categorized into four groups), French nationality, repeater status, disability status, whether the student was in a science academic track during high school, the student's option (either engineering science or computer science), and the gender composition in the *track* × *program*. We include year, track, and program fixed effects. Previous ability is controlled by (i) decile rank dummy variables from both high school and middle school graduation exams' GPAs, and (ii) quintile rank dummy variables for each grade at the high and middle school graduation exams in the subjects studied in prep programs—mathematics, physics and chemistry, engineering science, French (both written and oral), and foreign languages. Applications to top graduate schools are controlled by dummies for application to each of the exams leading to top graduate schools, defined track by track. Performance during competitive entrance exams is controlled for by (i) dummies for whether or not the applicant is ranked at competitive exams leading to top graduate schools and (ii) the percentile rank of their rank in those exams, controlled linearly for each entrance exam. Preferences are controlled for by adding dummy variables for whether or not each of the top graduate schools is ranked among students' ranked-ordered lists of schools. The order of these variables is chosen to follow the chronological order of decisions made by students. Standard errors are clustered at the *track* × *program* × *cohort* level.

C3 Comparison of Star and Standard Classes: Complementary Results

C3.1 Complementary Results Using Surveyed Data in Prep Programs

Table C8: Access to Star Class, by Gender

	(1)	(2)	(3)
	Access to star class	Access to star class	Access to star class
Baseline proba. of access	.52	.52	.52
Female student	-0.086***	-0.047***	-0.011
	(0.013)	(0.0079)	(0.0077)
Rank in the class at the end of first year		-0.029***	-0.014***
		(0.00056)	(0.0013)
First year second semester grades FE			✓
Year FE	✓	✓	✓
Track FE	✓	✓	✓
Program FE	✓	✓	✓
N	10,538	10,332	10,332

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: The table reports the effect of gender on access to star class. The rank in the class corresponds to the rank of student in their class at the end of their first year in the prep program, just before the decision of admission to a star class or a standard class. In the third column, we control for academic achievement at the second semester in the first year by adding dummy variables for each quintile in grades for mathematics, physics, chemistry, engineering science, computer science, French, and foreign language. We include year, track and program fixed effects. Standard errors are clustered at the *track* $\times$ *program* $\times$ *cohort* level.

Table C9: Access to STEM Graduate Schools in the Top 10 Percent of Selectivity : Controlling for Grades During the Prep Programs

	(1) Top 10% grad. schools (Selectivity)	(2) Top 10% grad. schools (Selectivity)	(3) Top 10% grad. schools (Selectivity)
Baseline proba. of access	0.31	0.31	0.31
Female student × Star class	-0.040**	-0.034**	-0.027
	(0.016)	(0.017)	(0.017)
Star class	0.38***	0.25***	0.17***
	(0.016)	(0.018)	(0.019)
Female student	-0.0098	-0.018	-0.021*
	(0.0089)	(0.011)	(0.011)
Demographic controls	✓	✓	✓
First year first semester grades FE		✓	
First year second semester grades FE			✓
Year FE	✓	✓	✓
Track FE	✓	✓	✓
Program FE	✓	✓	✓
N	12,559	10,060	9,863

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: The table displays the change in probability of admission to top-tier STEM graduate schools for female students in star classes. Selectivity of schools is measured based on percentile rank at the high school graduation exam of individuals admitted at the school. The sample is composed of all students in the prep programs we surveyed in 2022 and 2023 who we could statistically matched to SCEI outcome data (2015-2023). In this sample, we exclusively include non-repeaters to maximize the sample size. Studying repeaters would require tracking the programs over three years, significantly reducing the sample size. In Table C18 we show that in the full sample, effects for reapeaters and non-repeaters are very similar. Demographic control variables include geographic origin (Paris or Parisian area), low-income status, the socio-economic status of each parent (into four categories), French nationality, disability status, whether the student was in a science academic track during high school, the student's option (either engineering science or computer science), and the gender composition in the *track* × *program*. Previous ability is controlled by (i) decile rank dummy variable for the average GPA, and (ii) quintile rank dummy variables for each grades in mathematics, physics, chemistry, engineering science, computer science, French, and foreign languages. We include year, track and program fixed effects. Standard errors are clustered at the *track* × *program* × *cohort* level.

Table C10: Probability To Be Among the Top 20% of the Class, by Gender and Star Class Status

	Before star classes		During star classes	
	Top 20% 1st Year 1st Semester	Top 20% 1st Year 2nd Semester	Top 20% 2nd Year 1st Semester	Top 20% 2nd Year 2nd Semester
Baseline proba. to be in top 20%	0.21	0.21	0.18	0.18
Female student × Star class	0.0016	-0.011	-0.046***	-0.047***
	(0.017)	(0.017)	(0.014)	(0.016)
Star class	0.39***	0.41***	0.0096*	0.013**
	(0.012)	(0.014)	(0.0051)	(0.0059)
Female student	0.00057	0.0025	-0.017*	-0.0065
	(0.0041)	(0.0042)	(0.010)	(0.011)
Demographic controls	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
Track FE	✓	✓	✓	✓
Program FE	✓	✓	✓	✓
N	11,300	11,061	14,299	13,006

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: The table reports the change in the probability to be among the top 20% of the class during the prep programs for female students in future star classes in Columns (1) and (2), and in actual star classes in Columns (3) and (4). We use percentile rank weighted GPA to rank students. Demographic controls include geographic origin (Paris or Parisian area), low-income status, the socio-economic status of each parent (into four categories), French nationality, disability status, whether the student was in a science academic track during high school, the student's option (either engineering science or computer science), and the gender composition in the *track* × *program*. The analyses in Columns (3) and (4) is restricted to non-repeaters. Standard errors are clustered at the *track* × *program* × *cohort* level.

C3.2 Detailed Control Variables

Table C11: Access to STEM graduate schools in the Top 10 Percent of Selectivity

	(1) Top 10% grad. schools (Selectivity)	(2) Top 10% grad. schools (Selectivity)	(3) Top 10% grad. schools (Desirability)	(4) Top 10% grad. schools (Desirability)
Baseline proba. of access	0.22	0.22	0.23	0.23
Female student × Star class	-0.031***	-0.042***	-0.048***	-0.058***
	(0.0063)	(0.0061)	(0.0061)	(0.0060)
Star class	0.31***	0.27***	0.32***	0.29***
	(0.0062)	(0.0061)	(0.0063)	(0.0061)
Female student	-0.021***	-0.027***	-0.023***	-0.028***
	(0.0029)	(0.0031)	(0.0029)	(0.0030)
From Paris		0.052***		0.049***
		(0.0065)		(0.0064)
From parisian area		0.019***		0.016***
		(0.0056)		(0.0054)
Need-based scholarship students		-0.0093***		-0.0087***
		(0.0028)		(0.0028)
Repeater		0.055***		0.061***
		(0.0036)		(0.0037)
Disabled student		-0.0076		-0.0016
		(0.0087)		(0.0087)
Scientific high-school background		-0.0070		0.0031
		(0.0080)		(0.0086)
Student with French nationality		0.0016		0.010
		(0.0063)		(0.0065)
Engineering science option		0.031**		0.050***
		(0.013)		(0.014)
Computer science option		0.026*		0.052***
		(0.013)		(0.015)
Proportion of female students in the track		-0.00014		-0.00043
		(0.00032)		(0.00033)
Father High SES		0.023***		0.031***
		(0.0056)		(0.0054)
Father Medium High SES		0.010		0.019***
		(0.0063)		(0.0064)
Father Medium Low SES		-0.0027		0.0027
		(0.0057)		(0.0057)
Father Low SES		0.0042		0.0074
		(0.0061)		(0.0061)
Mother High SES		0.022***		0.021***
		(0.0044)		(0.0043)
Mother Medium High SES		-0.00048		0.0010
		(0.0050)		(0.0050)
Mother Medium Low SES		-0.0033		-0.0064
		(0.0046)		(0.0045)
Mother Low SES		-0.0066		-0.0066
		(0.0052)		(0.0053)
MS & HS Exam Score		✓		✓
Year FE	✓	✓	✓	✓
Track FE	✓	✓	✓	✓
Program FE	✓	✓	✓	✓
N	89,079	89,065	89,079	89,065

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: The table displays the change in probability of admission to top-tier STEM graduate schools for female students in star classes. Selectivity of schools is measured in two ways: Columns (1) and (2) use the average percentile rank of admitted students at the high school graduation exam, while Columns (3) and (4) are based on the revealed preferences of applicants (Avery et al. (2013)). Previous ability is controlled by (i) decile rank dummy variables from both high school and middle school graduation exams GPA, and (ii) quintile rank dummy variables for each grades at the high and middle school graduation exams at the subjects studied in prep programs — mathematics, physics and chemistry, engineering science, French (both written and oral), and foreign languages. We include year, track and program fixed effects. Standard errors are clustered at the *track* × *program* × *cohort* level.

C3.3 Effect on Expected Earnings

Table C12: Effect on Expected Earnings

	(1) Aggregated earnings (without bonuses)	(2) Aggregated earnings (without bonuses)	(3) Aggregated earnings (with bonuses)	(4) Aggregated earnings (with bonuses)
Baseline	38,836	38,836	42,021	42,021
Female student × Star class	-348.7***	-296.3***	-1038.1***	-873.7***
	(61.9)	(65.8)	(80.8)	(81.4)
Star class	1788.1***	1778.2***	3216.2***	3160.5***
	(51.2)	(51.4)	(85.6)	(83.8)
Female student	-677.3***		-1068.8***	
	(41.5)		(44.8)	
Demographic Controls	✓	✓	✓	✓
MS & HS Exam Score	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
Track FE	✓	✓	✓	✓
Program FE	✓	✓	✓	✓
Demographic Controls × Female student		✓		✓
MS & HS Exam Score × Female student		✓		✓
Year Fixed-Effects × Female student		✓		✓
Track Fixed-Effects × Female student		✓		✓
Program Fixed-Effects × Female student		✓		✓
N	61,235	61,235	76,545	76,545

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: The table reports the change in expected earnings, one year after graduation. It corresponds to the median gross annual salary of the most recent cohort of alumni of the school. Earnings are defined at the *school* × *cohort* × *gender* level and are sourced from the (CTI website). When the salary was missing for a specific year, we inferred it from the adjacent cohort. For the most recent cohorts that haven't graduated yet, we inferred the salary based on the previous cohort and the average salary increase per year for this specific school. Columns (1) and (2) correspond to salary without bonuses, while Columns (3) and (4) correspond to salary with bonuses. In Columns (2) and (4), we include interactions of all controls and fixed-effects with a gender dummy variable, allowing observable characteristics to have gender-specific effects . Demographic controls include geographic origin (Paris or Parisian area), low-income status, the socio-economic status of each parent (into four categories), French nationality, repeater status, disability status, whether the student was in a science academic track during high school, the student's option (either engineering science or computer science), and the gender composition in the *track* × *program*. Previous ability is controlled by (i) decile rank dummy variables from both high school and middle school graduation exams GPA, and (ii) quintile rank dummy variables for each grades at the high and middle school graduation exams at the subjects studied in prep programs — mathematics, physics and chemistry, engineering science, French (both written and oral), and foreign languages. We include year, track and program fixed effects. Standard errors are clustered at the *track* × *program* × *cohort* level.

C3.4 Effect of Repeating, by Gender

Table C13: Access to STEM Graduate Schools in the Top 10 Percent of Selectivity, by Gender and Repeater Status

	(1) Top 10% grad. schools (Selectivity)	(2) Top 10% grad. schools (Selectivity)	(3) Top 10% grad. schools (Desirability)	(4) Top 10% grad. schools (Desirability)
Baseline proba. of access	0.15	0.15	0.16	0.16
Female student × Repeater	0.0033	0.0011	-0.0015	-0.0041
	(0.0043)	(0.0041)	(0.0043)	(0.0041)
Repeater	0.024***	0.037***	0.029***	0.043***
	(0.0025)	(0.0024)	(0.0027)	(0.0025)
Female student	-0.046***	-0.038***	-0.052***	-0.043***
	(0.0020)	(0.0019)	(0.0021)	(0.0019)
Demographic Controls		✓		✓
MS & HS Exam Score		✓		✓
Year FE	✓	✓	✓	✓
Track FE	✓	✓	✓	✓
Program FE	✓	✓	✓	✓
N	165,450	165,450	165,450	165,450

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: The table illustrates the change in the probability of admission to top-tier STEM graduate schools for female students repeating in prep programs over the period 2015-2023. Selectivity of schools is measured in two ways: Columns (1) and (2) use the average percentile rank of admitted students at the high school graduation exam, while Columns (3) and (4) are based on the revealed preferences of applicants (Avery et al. (2013)). Demographic controls include geographic origin (Paris or Parisian area), low-income status, the socio-economic status of each parent (into four categories), French nationality, star class, disability status, whether the student was in a science academic track during high school, the student's option (either engineering science or computer science), and the gender composition in the *track* × *program*. Previous ability is controlled by (i) decile rank dummy variables from both high school and middle school graduation exams GPA, and (ii) quintile rank dummy variables for each grades at the high and middle school graduation exams at the subjects studied in prep programs — mathematics, physics and chemistry, engineering science, French (both written and oral), and foreign languages. We include year, track and program fixed effects. Standard errors are clustered at the *track* × *program* × *cohort* level. For a comprehensive list of coefficients on the demographic control variables, refer to Table C11 in the Appendix.

C3.5 Heterogeneity

Table C14: Heterogeneity by Income Status: Access to STEM Graduate Schools in the Top 10% of Selectivity

	(1) Top 10% grad. schools (Selectivity)	(2) Top 10% grad. schools (Selectivity)	(3) Top 10% grad. schools (Desirability)	(4) Top 10% grad. schools (Desirability)
Income status	Low-income	High-income	Low-income	High-income
Baseline proba. of access	0.14	0.25	0.16	0.26
Female student × Star class	-0.054***	-0.038***	-0.072***	-0.053***
	(0.010)	(0.0070)	(0.011)	(0.0070)
Star class	0.23***	0.29***	0.24***	0.30***
	(0.0075)	(0.0064)	(0.0079)	(0.0063)
Female student	-0.022***	-0.030***	-0.020***	-0.032***
	(0.0040)	(0.0037)	(0.0041)	(0.0037)
Demographic Controls	✓	✓	✓	✓
MS & HS Exam Score	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
Track FE	✓	✓	✓	✓
Program FE	✓	✓	✓	✓
N	23,409	65,656	23,409	65,656

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: The table reports the change in probability of admission to top-tier STEM graduate schools for female students in star classes, by parental income status. Low income students are need-based scholarships students and high-income students are non need-based scholarships students. Selectivity of schools is measured using the average percentile rank of admitted students at the high school graduation exam. Demographic controls include geographic origin (Paris or Parisian area), the socio-economic status of each parent (into four categories), French nationality, repeater status, disability status, whether the student was in a science academic track during high school, the student's option (either engineering science or computer science), and the gender composition in the *track* × *program*. Previous ability is controlled by (i) decile rank dummy variables from both high school and middle school graduation exams GPA, and (ii) quintile rank dummy variables for each grades at the high and middle school graduation exams at the subjects studied in prep programs — mathematics, physics and chemistry, engineering science, French (both written and oral), and foreign languages. We include year, track and program fixed effects. Standard errors are clustered at the *track* × *program* × *cohort* level.

Table C15: Heterogeneity by Ability: Access to STEM Graduate Schools in the Top 10% of Selectivity

	(1) Top 10% grad. schools (Selectivity)	(2) Top 10% grad. schools (Selectivity)	(3) Top 10% grad. schools (Selectivity)	(4) Top 10% grad. schools (Selectivity)
Quartile of ability	Q1	Q2	Q3	Q4
Baseline proba. of access	0.02	0.11	0.25	0.49
Female student × Star class	-0.036***	-0.081***	-0.082***	-0.035**
	(0.0076)	(0.0099)	(0.012)	(0.013)
Star class	0.078***	0.21***	0.34***	0.44***
	(0.0051)	(0.0071)	(0.0077)	(0.0087)
Female student	-0.0048***	-0.0070*	-0.016**	-0.057***
	(0.0016)	(0.0040)	(0.0062)	(0.010)
Demographic Controls	✓	✓	✓	✓
MS & HS Exam Score	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
Track FE	✓	✓	✓	✓
Program FE	✓	✓	✓	✓
N	20,587	20,591	20,585	20,593

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: The table reports the change in probability of admission to top-tier STEM graduate schools for female students in star classes, by prior ability status. We define ability based on re-weighted GPA at the high school graduation exam (with weights equal to competitive exams coefficients), and present results for each quartile. Selectivity of schools is measured using the average percentile rank of admitted students at the high school graduation exam. Demographic controls include geographic origin (Paris or Parisian area), low-income status, the socio-economic status of each parent (into four categories), French nationality, repeater status, disability status, whether the student was in a science academic track during high school, the student's option (either engineering science or computer science), and the gender composition in the *track* × *program*. Previous ability is controlled by (i) decile rank dummy variables from both high school and middle school graduation exams GPA, and (ii) quintile rank dummy variables for each grades at the high and middle school graduation exams at the subjects studied in prep programs — mathematics, physics and chemistry, engineering science, French (both written and oral), and foreign languages. We include year, track and program fixed effects. Standard errors are clustered at the *track* × *program* × *cohort* level.

Table C16: Heterogeneity by Ability: Access to STEM Graduate Schools in the Top 10% of Desirability

	(1) Top 10% grad. schools (Desirability)	(2) Top 10% grad. schools (Desirability)	(3) Top 10% grad. schools (Desirability)	(4) Top 10% grad. schools (Desirability)
Quartile of ability	Q1	Q2	Q3	Q4
Baseline proba. of access	0.03	0.12	0.27	0.51
Female student × Star class	-0.042***	-0.092***	-0.11***	-0.049***
	(0.0081)	(0.0100)	(0.012)	(0.013)
Star class	0.092***	0.22***	0.35***	0.45***
	(0.0056)	(0.0072)	(0.0078)	(0.0087)
Female student	-0.0063***	-0.013***	-0.016**	-0.051***
	(0.0018)	(0.0042)	(0.0067)	(0.0099)
Demographic Controls	✓	✓	✓	✓
MS & HS Exam Score	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
Track FE	✓	✓	✓	✓
Program FE	✓	✓	✓	✓
N	20,587	20,591	20,585	20,593

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: The table reports the change in probability of admission to top-tier STEM graduate schools for female students in star classes, by prior ability status. We define ability based on re-weighted GPA at the high school graduation exam (with weights equal to competitive exams coefficients), and present results for each quartile. School desirability is based on the revealed preferences of applicants (Avery et al. (2013)). Demographic controls include geographic origin (Paris or Parisian area), low-income status, the socio-economic status of each parent (into four categories), French nationality, repeater status, disability status, whether the student was in a science academic track during high school, the student's option (either engineering science or computer science), and the gender composition in the *track* × *program*. Previous ability is controlled by (i) decile rank dummy variables from both high school and middle school graduation exams GPA, and (ii) quintile rank dummy variables for each grades at the high and middle school graduation exams at the subjects studied in prep programs — mathematics, physics and chemistry, engineering science, French (both written and oral), and foreign languages. We include year, track and program fixed effects. Standard errors are clustered at the *track* × *program* × *cohort* level.

Table C17: Heterogeneity by Selectivity of the Prep Programs: Access to STEM Graduate Schools in the Top 10% of Selectivity

	(1) Top 10% grad. schools (Selectivity)	(2) Top 10% grad. schools (Selectivity)	(3) Top 10% grad. schools (Selectivity)	(4) Top 10% grad. schools (Selectivity)
Quartile of selectivity of prep. program	Q1	Q2	Q3	Q4
Baseline proba. of access	0.05	0.13	0.19	0.47
Female student × Star class	-0.051***	-0.079***	-0.075***	0.00016
	(0.0069)	(0.011)	(0.012)	(0.013)
Star class	0.090***	0.21***	0.31***	0.43***
	(0.0062)	(0.0100)	(0.0093)	(0.0095)
Female student	-0.0030	-0.015***	-0.020***	-0.047***
	(0.0019)	(0.0038)	(0.0045)	(0.0090)
Demographic Controls	✓	✓	✓	✓
MS & HS Exam Score	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
Track FE	✓	✓	✓	✓
Program FE	✓	✓	✓	✓
N	19,320	20,712	24,134	24,899

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: The table reports the change in probability of admission to top-tier STEM graduate schools for female students in star classes, by selectivity of the prep program. Selectivity of prep programs and graduate schools is measured using the average percentile rank of admitted students at the high school graduation exam. Demographic controls include geographic origin (Paris or Parisian area), low-income status, the socio-economic status of each parent (into four categories), French nationality, repeater status, disability status, whether the student was in a science academic track during high school, the student's option (either engineering science or computer science), and the gender composition in the *track* × *program*. Previous ability is controlled by (i) decile rank dummy variables from both high school and middle school graduation exams GPA, and (ii) quintile rank dummy variables for each grades at the high and middle school graduation exams at the subjects studied in prep programs — mathematics, physics and chemistry, engineering science, French (both written and oral), and foreign languages. We include year, track and program fixed effects. Standard errors are clustered at the *track* × *program* × *cohort* level.

Table C18: Access to STEM Graduate Schools in the Top 10% of Selectivity, by Repeater Status

	(1) Top 10% grad. schools (Selectivity)	(2) Top 10% grad. schools (Selectivity)	(3) Top 10% grad. schools (Desirability)	(4) Top 10% grad. schools (Desirability)
Repeater status	Non-repeaters	Repeaters	Non-repeaters	Repeaters
Baseline proba. of access	0.22	0.23	0.23	0.25
Female student $\times$ Star class	-0.042***	-0.046***	-0.057***	-0.062***
	(0.0065)	(0.014)	(0.0063)	(0.014)
Star class	0.27***	0.28***	0.29***	0.29***
	(0.0064)	(0.0088)	(0.0064)	(0.0091)
Female student	-0.026***	-0.030***	-0.026***	-0.033***
	(0.0032)	(0.0065)	(0.0033)	(0.0071)
Demographic Controls	✓	✓	✓	✓
MS & HS Exam Score	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
Track FE	✓	✓	✓	✓
Program FE	✓	✓	✓	✓
N	72,727	16,268	72,727	16,268

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: The table reports the change in probability of admission to top-tier STEM graduate schools for female students in star classes, by repeater status. Selectivity of schools is measured in two ways: Columns (1) and (2) use the average percentile rank of admitted students at the high school graduation exam, while Columns (3) and (4) are based on the revealed preferences of applicants (Avery et al. (2013)). Demographic controls include geographic origin (Paris or Parisian area), low-income status, the socio-economic status of each parent (into four categories), French nationality, disability status, whether the student was in a science academic track during high school, the student's option (either engineering science or computer science), and the gender composition in the *track* $\times$ *program*. Previous ability is controlled by (i) decile rank dummy variables from both high school and middle school graduation exams GPA, and (ii) quintile rank dummy variables for each grades at the high and middle school graduation exams at the subjects studied in prep programs — mathematics, physics and chemistry, engineering science, French (both written and oral), and foreign languages. We include year, track and program fixed effects. Standard errors are clustered at the *track* $\times$ *program* $\times$ *cohort* level.

C3.6 Robustness Checks

Table C19: Placebo Test: High School Graduation Exam Percentile Ranks

	(1) GPA	(2) GPA	(3) Re-weighted GPA	(4) Re-weighted GPA
Baseline percentile rank	0.84	0.84	0.92	0.92
Female student × Star class	0.00042	0.00030	0.0021	0.0019
	(0.0018)	(0.0017)	(0.0015)	(0.0015)
Star class	0.077***	0.077***	0.058***	0.057***
	(0.0017)	(0.0017)	(0.0016)	(0.0016)
Female student	0.043***	0.041***	0.011***	0.0097***
	(0.0013)	(0.0013)	(0.0012)	(0.0012)
Demographic Controls		✓		✓
Year FE	✓	✓	✓	✓
Track FE	✓	✓	✓	✓
Program FE	✓	✓	✓	✓
N	82,352	82,352	82,356	82,356

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: The table reports the change in percentile rank at the high school graduation exam for female students in star classes. In Columns (1) and (2), percentile ranks are computed using raw GPA among all high school graduation exam recipients while in Columns (3) and (4), percentile ranks are computed using re-weighted GPA, with weight equal to the most selective competitive exam coefficients, as presented in Table A1. Demographic controls include geographic origin (Paris or Parisian area), low-income status, the socio-economic status of each parent (into four categories), French nationality, repeater status, disability status, whether the student was in a science academic track during high school, the student's option (either engineering science or computer science), and the gender composition in the *track* × *program*. We include year, track and program fixed effects. Standard errors are clustered at the *track* × *program* × *cohort* level.

Table C20: Placebo Test: Prep Programs GPA

	(1) Pct. Rank GPA in 1st Year 1st Semester	(2) Pct. Rank GPA in 1st Year 1st Semester	(3) Pct. Rank GPA in 1st Year 2nd Semester	(4) Pct. Rank GPA in 1st Year 2nd Semester
Baseline percentile rank	0.51	0.51	0.52	0.52
Female Student in Star class	0.011	0.0080	0.0049	0.00079
	(0.0092)	(0.0087)	(0.0095)	(0.0089)
Star class	0.41***	0.39***	0.43***	0.41***
	(0.0050)	(0.0046)	(0.0047)	(0.0045)
Female	0.0080	0.011*	0.011*	0.012**
	(0.0064)	(0.0063)	(0.0060)	(0.0059)
Demographic controls		✓		✓
Year FE	✓	✓	✓	✓
Track FE	✓	✓	✓	✓
Program FE	✓	✓	✓	✓
N	11,300	11,300	11,061	11,061

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: The table reports the change in percentile rank at the weighted GPA during the prep programs for female students in (future) star classes. In Columns (1) and (2), we use the weighted GPA at the end of the first semester in the first year, while in Columns (3) and (4), we use the weighted GPA at the end of the second semester in the first year. Demographic controls include geographic origin (Paris or Parisian area), low-income status, the socio-economic status of each parent (into four categories), French nationality, repeater status, disability status, whether the student was in a science academic track during high school, the student's option (either engineering science or computer science), and the gender composition in the *track* × *program*. Standard errors are clustered at the *track* × *program* × *cohort* level.

Table C21: Access to STEM Graduate Schools in the Top 10 Percent of Desirability for Female Students

	(1) Top 10% grad. schools (Desirability)	(2) Top 10% grad. schools (Desirability)	(3) Top 10% grad. schools (Desirability)
Baseline proba. of access	0.26	0.26	0.26
Female student × Star class	-0.060***	-0.068***	-0.055***
	(0.0062)	(0.0062)	(0.0066)
Star class	0.35***	0.31***	0.30***
	(0.0057)	(0.0057)	(0.0058)
Female student	-0.016***	-0.022***	
	(0.0030)	(0.0031)	
Demographic Controls		✓	✓
MS & HS Exam Score		✓	✓
Year FE	✓	✓	✓
Track FE	✓	✓	✓
Program FE	✓	✓	✓
Demographic Controls × Female student			✓
MS & HS Exam Score × Female student			✓
Year Fixed-Effects × Female student			✓
Track Fixed-Effects × Female student			✓
Program Fixed-Effects × Female student			✓
N	89,079	89,065	89,065

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: The table reports the change in probability of admission to top-tier STEM graduate schools for female students in star classes. Selectivity of schools is based on the revealed preferences of female applicants only (Avery et al. (2013)). In Columns (3), we include interactions of all controls and fixed-effects with a gender dummy variable, allowing observable characteristics to have gender-specific performance impacts. Demographic controls include geographic origin (Paris or Parisian area), low-income status, the socio-economic status of each parent (into four categories), French nationality, repeater status, disability status, whether the student was in a science academic track during high school, the student's option (either engineering science or computer science), and the gender composition in the *track* × *program*. Previous ability is controlled by (i) decile rank dummy variables from both high school and middle school graduation exams GPA, and (ii) quintile rank dummy variables for each grades at the high and middle school graduation exams at the subjects studied in prep programs — mathematics, physics and chemistry, engineering science, French (both written and oral), and foreign languages. We include year, track and program fixed effects. Standard errors are clustered at the *track* × *program* × *cohort* level.

Table C22: Robustness Check: Other Measures of School Selectivity

	(1) Average selectivity	(2) Top 20% of selectivity	(3) Average desirability	(4) Top 20% of desirability
Baseline proba. of access	83	0.39	0.92	0.42
Female student $\times$ Star class	-0.40***	-0.022***	-0.13***	-0.034***
	(0.095)	(0.0064)	(0.012)	(0.0065)
Star class	6.91***	0.34***	0.88***	0.35***
	(0.073)	(0.0054)	(0.012)	(0.0050)
Female student	-0.69***	-0.033***	-0.11***	-0.038***
	(0.076)	(0.0037)	(0.0072)	(0.0038)
Demographic Controls	✓	✓	✓	✓
MS & HS Exam Score	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
Track FE	✓	✓	✓	✓
Program FE	✓	✓	✓	✓
N	78,001	89,065	78,001	89,065

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: The table reports the change in average selectivity of admission school and the probability of admission to top quintile STEM graduate schools for female students in star classes. Selectivity of schools is measured in two ways: Columns (1) and (2) use the average percentile rank of admitted students at the high school graduation exam, while Columns (3) and (4) are based on the revealed preferences of applicants (Avery et al. (2013)). Demographic controls include geographic origin (Paris or Parisian area), low-income status, the socio-economic status of each parent (into four categories), French nationality, repeater status, disability status, whether the student was in a science academic track during high school, the student's option (either engineering science or computer science), and the gender composition in the *track* $\times$ *program*. Previous ability is controlled by (i) decile rank dummy variables from both high school and middle school graduation exams GPA, and (ii) quintile rank dummy variables for each grades at the high and middle school graduation exams at the subjects studied in prep programs — mathematics, physics and chemistry, engineering science, French (both written and oral), and foreign languages. We include year, track and program fixed effects. Standard errors are clustered at the *track* $\times$ *program* $\times$ *cohort* level.

Table C23: Probit Model: Access to STEM Graduate Schools in the Top 10 Percent of Selectivity

	(1) Top 10% grad. schools (Selectivity)	(2) Top 10% grad. schools (Selectivity)	(3) Top 10% grad. schools (Desirability)	(4) Top 10% grad. schools (Desirability)
Baseline proba. of access	0.22	0.22	0.23	0.23
Female student × Star class	-0.058*	-0.095***	-0.083***	-0.12***
	(0.032)	(0.034)	(0.030)	(0.032)
Star class	1.62***	1.47***	1.62***	1.47***
	(0.019)	(0.020)	(0.018)	(0.019)
Female student	-0.13***	-0.21***	-0.16***	-0.24***
	(0.027)	(0.030)	(0.025)	(0.027)
Demographic Controls		✓		✓
MS & HS Exam Score		✓		✓
Year FE	✓	✓	✓	✓
Track FE	✓	✓	✓	✓
Program FE	✓	✓	✓	✓
N	88,438	88,424	88,438	88,424

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: The table displays the change in probability of admission to top-tier STEM graduate schools for female students in star classes. Selectivity of schools is measured in two ways: Columns (1) and (2) use the average percentile rank of admitted students at the high school graduation exam, while Columns (3) and (4) are based on the revealed preferences of applicants (Avery et al. (2013)). Previous ability is controlled by (i) decile rank dummy variables from both high school and middle school graduation exams GPA, and (ii) quintile rank dummy variables for each grades at the high and middle school graduation exams at the subjects studied in prep programs — mathematics, physics and chemistry, engineering science, French (both written and oral), and foreign languages. We include year, track and program fixed effects. Standard errors are clustered at the *track* × *program* × *cohort* level.

Table C24: Robustness Check: Interaction of Fixed Effects

	(1) Top 10% grad. schools (Selectivity)	(2) Top 10% grad. schools (Selectivity)	(3) Top 10% grad. schools (Desirability)	(4) Top 10% grad. schools (Desirability)
Baseline proba. of access	0.22	0.22	0.23	0.23
Female student × Star class	-0.042***	-0.040***	-0.058***	-0.057***
	(0.0061)	(0.0061)	(0.0060)	(0.0061)
Star class	0.27***	0.28***	0.29***	0.29***
	(0.0061)	(0.0062)	(0.0061)	(0.0062)
Female student	-0.027***	-0.052***	-0.028***	-0.056***
	(0.0031)	(0.0087)	(0.0030)	(0.0087)
Demographic Controls	✓	✓	✓	✓
MS & HS Exam Score	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
Track FE	✓	✓	✓	✓
Program FE	✓	✓	✓	✓
Year × Track FE		✓		✓
Year × Program FE		✓		✓
Track × Program FE		✓		✓
Year × Track × Program FE		✓		✓
N	89,065	89,065	89,065	89,065

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: The table reports the change in probability of admission to top-tier STEM graduate schools for female students in star classes. Selectivity of schools is measured in two ways: Columns (1) and (2) use the average percentile rank of admitted students at the high school graduation exam, while Columns (3) and (4) are based on the revealed preferences of applicants (Avery et al. (2013)). Demographic controls include geographic origin (Paris or Parisian area), low-income status, the socio-economic status of each parent (into four categories), French nationality, repeater status, disability status, whether the student was in a science academic track during high school, the student's option (either engineering science or computer science), and the gender composition in the *track* × *program*. Previous ability is controlled by (i) decile rank dummy variables from both high school and middle school graduation exams GPA, and (ii) quintile rank dummy variables for each grades at the high and middle school graduation exams at the subjects studied in prep programs — mathematics, physics and chemistry, engineering science, French (both written and oral), and foreign languages. We include year, track and program fixed effects. We include *year* × *track*, *year* × *program*, *track* × *program*, and *year* × *track* × *program* fixed effects in Columns (2) and (4). Standard errors are clustered at the *track* × *program* × *cohort* level.

Table C25: Access to STEM Graduate Schools in the Top 10% of Selectivity, by Track

	(1) Top 10% grad. schools (Selectivity)	(2) Top 10% grad. schools (Selectivity)	(3) Top 10% grad. schools (Selectivity)	(4) Top 10% grad. schools (Selectivity)
Track	(MP)	(PC)	(PSI)	(PT)
Baseline proba. of access	0.28	0.20	0.21	0.05
Female student × Star class	-0.039***	-0.043***	-0.053***	-0.037***
	(0.011)	(0.0086)	(0.014)	(0.012)
Star class	0.33***	0.26***	0.29***	0.078***
	(0.0086)	(0.011)	(0.014)	(0.012)
Female student	-0.033***	-0.012***	-0.038***	-0.0046
	(0.0052)	(0.0045)	(0.0067)	(0.0037)
Demographic Controls	✓	✓	✓	✓
MS & HS Exam Score	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
Track FE				
Program FE	✓	✓	✓	✓
N	36,061	28,155	17,474	7,375

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: The table reports the change in probability of admission to top-tier STEM graduate schools for female students in star classes, by track. Selectivity of schools is measured using the average percentile rank of admitted students at the high school graduation exam. Demographic controls include geographic origin (Paris or Parisian area), low-income status, the socio-economic status of each parent (into four categories), French nationality, repeater status, disability status, whether the student was in a science academic track during high school, the student's option (either engineering science or computer science), and the gender composition in the *track* × *program*. Previous ability is controlled by (i) decile rank dummy variables from both high school and middle school graduation exams GPA, and (ii) quintile rank dummy variables for each grades at the high and middle school graduation exams at the subjects studied in prep programs — mathematics, physics and chemistry, engineering science, French (both written and oral), and foreign languages. We include year and prep program fixed effects. Standard errors are clustered at the *track* × *program* × *cohort* level.

Table C26: Access to STEM Graduate Schools in the Top 10% of Desirability, by Track

	(1) Top 10% grad. schools (Desirability)	(2) Top 10% grad. schools (Desirability)	(3) Top 10% grad. schools (Desirability)	(4) Top 10% grad. schools (Desirability)
Track	(MP)	(PC)	(PSI)	(PT)
Baseline proba. of access	0.31	0.19	0.23	0.05
Female student × Star class	-0.041***	-0.063***	-0.056***	-0.038***
	(0.010)	(0.0084)	(0.014)	(0.011)
Star class	0.36***	0.25***	0.30***	0.079***
	(0.0085)	(0.011)	(0.013)	(0.012)
Female student	-0.035***	-0.017***	-0.037***	-0.0064
	(0.0049)	(0.0045)	(0.0063)	(0.0040)
Demographic Controls	✓	✓	✓	✓
MS & HS Exam Score	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
Track FE				
Program FE	✓	✓	✓	✓
N	36,061	28,155	17,474	7,375

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: The table reports the change in probability of admission to top-tier STEM graduate schools for female students in star classes, by track. School desirability is based on the revealed preferences of applicants (Avery et al. (2013)). Demographic controls include geographic origin (Paris or Parisian area), low-income status, the socio-economic status of each parent (into four categories), French nationality, repeater status, disability status, whether the student was in a science academic track during high school, the student's option (either engineering science or computer science), and the gender composition in the *track* × *program*. Previous ability is controlled by (i) decile rank dummy variables from both high school and middle school graduation exams GPA, and (ii) quintile rank dummy variables for each grades at the high and middle school graduation exams at the subjects studied in prep programs — mathematics, physics and chemistry, engineering science, French (both written and oral), and foreign languages. We include year and prep program fixed effects. Standard errors are clustered at the *track* × *program* × *cohort* level.

Table C27: Access to STEM Graduate Schools in the Top 10 Percent of Selectivity, Controlling for Number of Selective Exams Taken

	(1) Top 10% grad. schools (Selectivity)	(2) Top 10% grad. schools (Selectivity)	(3) Top 10% grad. schools (Desirability)	(4) Top 10% grad. schools (Desirability)
Baseline proba. of access	0.22	0.22	0.23	0.23
Female students × Star class	-0.044***	-0.036***	-0.060***	-0.050***
	(0.0058)	(0.0061)	(0.0058)	(0.0062)
Star class	0.23***	0.23***	0.25***	0.24***
	(0.0062)	(0.0063)	(0.0063)	(0.0064)
Female students	-0.022***		-0.022***	
	(0.0030)		(0.0030)	
Demographic Controls	✓	✓	✓	✓
MS & HS Exam Score	✓	✓	✓	✓
FE Number of Selective Exams Taken	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
Track FE	✓	✓	✓	✓
Program FE	✓	✓	✓	✓
Demographic Controls × Female student		✓		✓
MS & HS Exam Score × Female student		✓		✓
Year Fixed-Effects × Female student		✓		✓
Track Fixed-Effects × Female student		✓		✓
Program Fixed-Effects × Female student		✓		✓
N	89,015	89,015	89,015	89,015

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: The table illustrates the change in the probability of admission to top-tier STEM graduate schools for female students in star classes during the period from 2015 to 2023, while controlling for the number of selective the exams taken. We include fixed effects for the number of (1) very selective exams taken (i.e., those leading to only top-tier schools) and (2) selective exams taken (i.e., those leading to at least one top-tier school). Selectivity of schools is measured in two ways: Columns (1) and (2) use the average percentile rank of admitted students at the high school graduation exam, while Columns (3) and (4) are based on the revealed preferences of applicants (Avery et al., 2013). In Columns (2) and (4), we include interactions of all controls and fixed-effects with a gender dummy variable, allowing observable characteristics to have gender-specific performance impacts. Demographic controls include geographic origin (Paris or Parisian area), low-income status, the socio-economic status of each parent (into four categories), French nationality, repeater status, disability status, whether the student was in a science academic track during high school, the student's option (either engineering science or computer science), and the gender composition in the *track* × *program*. Previous ability is controlled by (i) decile rank dummy variables from both high school and middle school graduation exams GPA, and (ii) quintile rank dummy variables for each grades at the high and middle school graduation exams at the subjects studied in prep programs — mathematics, physics and chemistry, engineering science, French (both written and oral), and foreign languages. We include year, track and program fixed effects. Standard errors are clustered at the *track* × *program* × *cohort* level. For a comprehensive list of coefficients on the demographic control variables, refer to Table C11 in the Appendix.

Table C28: Access to STEM graduate schools in the Top 10% of Selectivity, by Year

	(1) Top 10% grad. schools (Selectivity)	(2) Top 10% grad. schools (Selectivity)	(3) Top 10% grad. schools (Selectivity)	(4) Top 10% grad. schools (Selectivity)	(5) Top 10% grad. schools (Selectivity)	(6) Top 10% grad. schools (Selectivity)	(7) Top 10% grad. schools (Selectivity)	(8) Top 10% grad. schools (Selectivity)
Year	(2015)	(2016)	(2017)	(2018)	(2019)	(2020)	(2021)	(2022)
Baseline	0.22	0.21	0.22	0.21	0.22	0.22	0.22	0.23
Female student × Star class	-0.033*	-0.044**	-0.026	-0.039**	-0.021	-0.056***	-0.058***	-0.058***
	(0.019)	(0.017)	(0.021)	(0.017)	(0.019)	(0.019)	(0.017)	(0.019)
Star class	0.26***	0.25***	0.26***	0.25***	0.27***	0.29***	0.27***	0.30***
	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.019)	(0.018)	(0.019)
Female student	-0.028***	-0.023**	-0.029**	-0.022**	-0.035***	-0.018**	-0.012	-0.035***
	(0.010)	(0.0094)	(0.012)	(0.0100)	(0.0096)	(0.0081)	(0.0098)	(0.0074)
Demographic Controls	✓	✓	✓	✓	✓	✓	✓	✓
MS & HS Exam Score	✓	✓	✓	✓	✓	✓	✓	✓
Year Fixed Effects								
Track Fixed Effects	✓	✓	✓	✓	✓	✓	✓	✓
Program Fixed Effects	✓	✓	✓	✓	✓	✓	✓	✓
N	9,872	10,068	9,998	9,963	10,047	9,969	9,960	9,968

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: The table reports the change in probability of admission to top-tier STEM graduate schools for female students in star classes, by year. Selectivity of schools is measured using the average percentile rank of admitted students at the high school graduation exam. Demographic controls include geographic origin (Paris or Parisian area), low-income status, the socio-economic status of each parent (into four categories), French nationality, repeater status, disability status, whether the student was in a science academic track during high school, the student's option (either engineering science or computer science), and the gender composition in the *track* × *program*. Previous ability is controlled by (i) decile rank dummy variables from both high school and middle school graduation exams GPA, and (ii) quintile rank dummy variables for each grades at the high and middle school graduation exams at the subjects studied in prep programs — mathematics, physics and chemistry, engineering science, French (both written and oral), and foreign languages. We include track and prep program fixed effects. Standard errors are clustered at the *track* × *program* × *cohort* level.

Table C29: Access to STEM graduate schools in the Top 10% of Desirability, by Year

	(1) Top 10% grad. schools (Desirability)	(2) Top 10% grad. schools (Desirability)	(3) Top 10% grad. schools (Desirability)	(4) Top 10% grad. schools (Desirability)	(5) Top 10% grad. schools (Desirability)	(6) Top 10% grad. schools (Desirability)	(7) Top 10% grad. schools (Desirability)	(8) Top 10% grad. schools (Desirability)
Year	(2015)	(2016)	(2017)	(2018)	(2019)	(2020)	(2021)	(2022)
Baseline	0.23	0.23	0.23	0.22	0.23	0.23	0.23	0.24
Female student × Star class	-0.057***	-0.056***	-0.051**	-0.054***	-0.043**	-0.059***	-0.059***	-0.075***
	(0.020)	(0.016)	(0.021)	(0.019)	(0.019)	(0.015)	(0.017)	(0.019)
Star class	0.27***	0.26***	0.27***	0.28***	0.27***	0.30***	0.28***	0.32***
	(0.018)	(0.018)	(0.016)	(0.019)	(0.018)	(0.020)	(0.018)	(0.019)
Female student	-0.026**	-0.024**	-0.029***	-0.024**	-0.030***	-0.023***	-0.026***	-0.036***
	(0.011)	(0.0092)	(0.010)	(0.0098)	(0.0090)	(0.0086)	(0.0092)	(0.0083)
Demographic Controls	✓	✓	✓	✓	✓	✓	✓	✓
MS & HS Exam Score	✓	✓	✓	✓	✓	✓	✓	✓
Year Fixed Effects								
Track Fixed Effects	✓	✓	✓	✓	✓	✓	✓	✓
Program Fixed Effects	✓	✓	✓	✓	✓	✓	✓	✓
N	9,872	10,068	9,998	9,963	10,047	9,969	9,960	9,968

Standard errors in parentheses

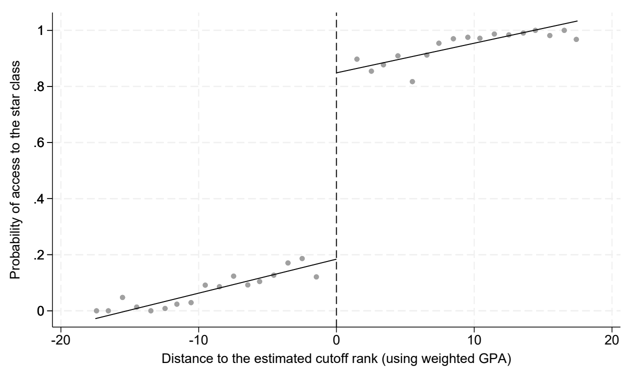
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: The table reports the change in probability of admission to top-tier STEM graduate schools for female students in star classes, by year. School desirability is based on the revealed preferences of applicants (Avery et al. (2013)). Demographic controls include geographic origin (Paris or Parisian area), low-income status, the socio-economic status of each parent (into four categories), French nationality, repeater status, disability status, whether the student was in a science academic track during high school, the student's option (either engineering science or computer science), and the gender composition in the *track* × *program*. Previous ability is controlled by (i) decile rank dummy variables from both high school and middle school graduation exams GPA, and (ii) quintile rank dummy variables for each grades at the high and middle school graduation exams at the subjects studied in prep programs — mathematics, physics and chemistry, engineering science, French (both written and oral), and foreign languages. We include track and prep program fixed effects. Standard errors are clustered at the *track* × *program* × *cohort* level.

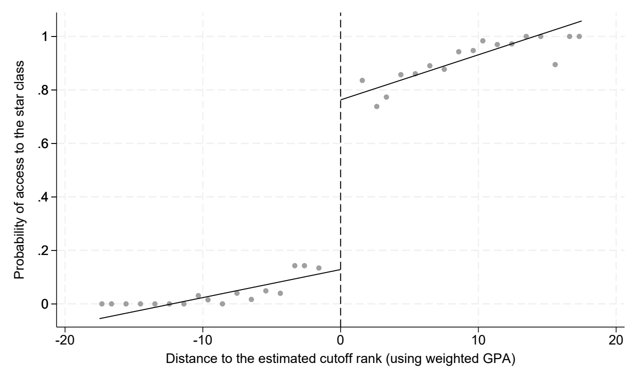
D4 Regression Discontinuity: Complementary Results

D4.1 Validity of the Regression Discontinuity

Figure D11: First Stage of the Regression Discontinuity, by Gender



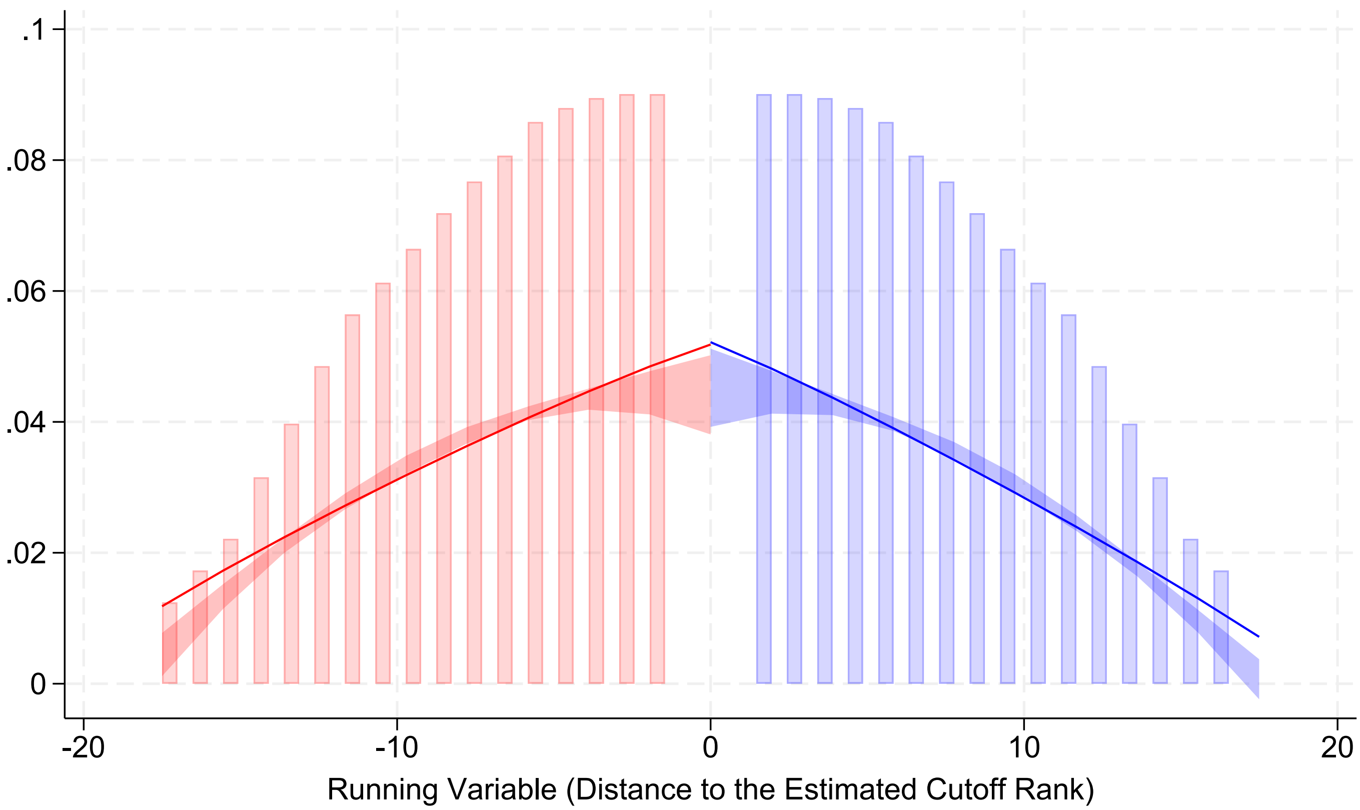
(a) Male Students



(b) Female Students

Notes: These figures display the probability of admission to a star class over the period 2015-2023, around the threshold for entering a star class. The running variable is defined as the distance between a student's weighted GPA and the cutoff weighted GPA for star class admission, computed at the *class* $\times$ *track* $\times$ *prep program* $\times$ *year* level following the algorithm used in Hoekstra (2009) and in Bütikofer et al. (2023). These figures are drawn using the *rdplot* Stata package developed by Calonico et al. (2017).

Figure D12: Density of the Running Variable Around the Threshold



Notes: The graph illustrates the density of the running variable (students' rank in the 1st-year class) around the star class admission threshold, assessing potential manipulation at the cutoff via an RDDensity test. The running variable is defined as the distance between a student's weighted GPA and the cutoff weighted GPA for star class admission, computed at the *class* × *track* × *prep program* × *year* level using the algorithm from Hoekstra (2009) and Bütikofer et al. (2023). These figures are drawn using the *rdddensity* Stata package developed by Cattaneo et al. (2018). We exclude individuals immediately below and above the threshold due to the characteristics of our algorithm used to determine the cutoff. By construction of our sample, we have a balanced sample around the threshold (see Section 5.2.1).

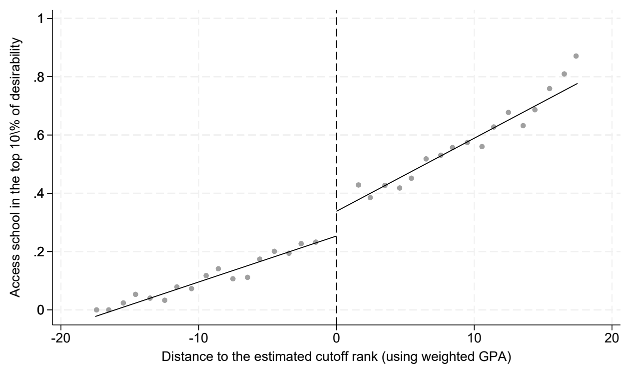
D4.2 Complementary Results

Table D30: Admission to Top 10% Most Desirable STEM Graduate Schools, by Gender

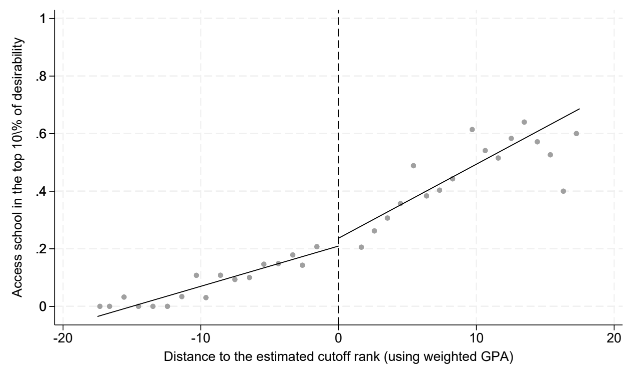
	(1) Top 10 decile of desirability (All students)	(2) Top 10 decile of desirability (Male students)	(3) Top 10 decile of desirability (Female students)
Baseline mean (Below cutoff)	0.13	0.14	0.10
Baseline RD estimate	0.160***	0.226***	0.010
	(0.033)	(0.035)	(0.083)
Robust 95% CI	[0.096 ; 0.224]	[0.157 ; 0.294]	[-0.152 ; 0.172]
Obs. used in estimation	2,358	1,303	475
Total number of obs.	6,782	4,934	1,848

Notes: This table displays the fuzzy non-parametric regression discontinuity estimates of the probability of admission to top-tier STEM graduate schools. These estimates are based on Calonico et al. (2017, 2019). Desirability of schools is measured using applicants' revealed preferences (Avery et al. (2013)). The running variable is defined as the distance between a student's rank and the cutoff rank for star class admission, computed at the *class* $\times$ *track* $\times$ *prep program* $\times$ *year*. We control for program $\times$ class fixed effects. Standard errors are clustered at the level of the running variable.

Figure D13: Admission to Top 10% Most Desirable STEM Graduate Schools, by Gender



(a) Male Students



(b) Female Students

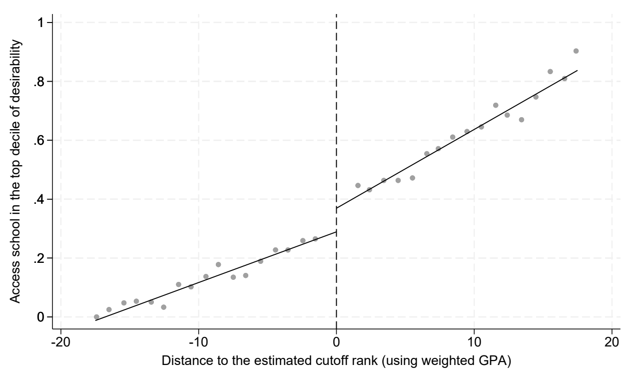
Notes: These figures display the probability of admission to top-tier STEM graduate schools over the period 2015-2023, around the margin of star class admission. Desirability of schools is measured using applicants' revealed preferences Avery et al. (2013). The running variable is defined as the distance between a student's rank and the cutoff rank for star class admission, computed at the *class* $\times$ *track* $\times$ *prep program* $\times$ *year* level following the algorithm used in Hoekstra (2009) and in Bütikofer et al. (2023). These figures are drawn using the *rdplot* Stata package developed by Calonico et al. (2017).

Table D31: Admission to Top 10% Most Desirable (For Female Students) STEM Graduate Schools, by Gender

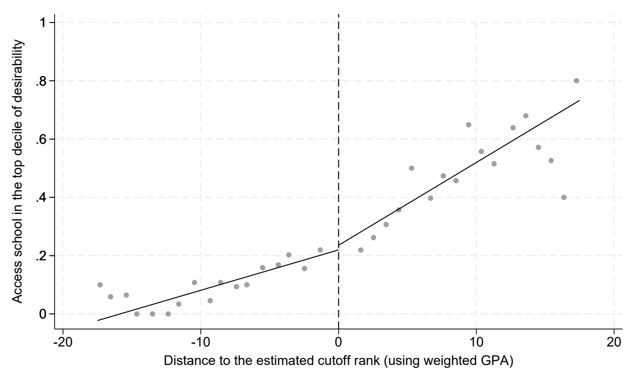
	(1) Top 10 decile of desirability for female (All students)	(2) Top 10 decile of desirability for female (Male students)	(3) Top 10 decile of desirability for female (Female students)
Baseline mean (Below cutoff)	0.15	0.16	0.12
Baseline RD estimate	0.133*** (0.023)	0.180*** (0.019)	0.149* (0.082)
Robust 95% CI	[0.087 ; 0.179]	[0.143 ; 0.216]	[-0.012 ; 0.311]
Obs. used in estimation	2,358	2,110	475
Total number of obs.	6,782	4,934	1,848

Notes: This table displays the fuzzy non-parametric regression discontinuity estimates of admission probability to top-tier STEM graduate schools. These estimates are based on Calonico et al. (2017, 2019). Desirability of schools is measured using applicants' revealed preferences, defined with preferences of female students only (Avery et al. (2013)). The running variable is defined as the distance between a student's rank and the cutoff rank for star class admission, computed at the *class* × *track* × *prep program* × *year*. We control for program × class fixed effects. Standard errors are clustered at the level of the running variable.

Figure D14: Admission to Top 10% Most Desirable (For Female Students) STEM Graduate Schools, by Gender



(a) Male Students



(b) Female Students

Notes: These figures display the probability of admission to top-tier STEM graduate schools over the period 2015-2023, around the margin of star class admission. Desirability of schools is measured using applicants' revealed preferences, defined with preferences of female students only (Avery et al. (2013)). The running variable is defined as the distance between a student's rank and the cutoff rank for star class admission, computed at the *class* × *track* × *prep program* × *year* level following the algorithm used in Hoekstra (2009) and in Bütikofer et al. (2023). These figures are drawn using the *rdplot* Stata package developed by Calonico et al. (2017).

D4.3 Robustness Checks

Table D32: RDD Robustness Check: No Clustering at the Running Variable

	(1) Top 10 decile of selectivity (All students)	(2) Top 10 decile of selectivity (Male students)	(3) Top 10 decile of selectivity (Female students)
Baseline mean (Below cutoff)	0.11	0.11	0.09
Baseline RD estimate	0.117	0.199**	-0.103
	(0.078)	(0.098)	(0.154)
Robust 95% CI	[-0.036 ; 0.270]	[0.006 ; 0.392]	[-0.405 ; 0.200]
Obs. used in estimation	2,358	1,303	646
Total number of obs.	6,782	4,934	1,848

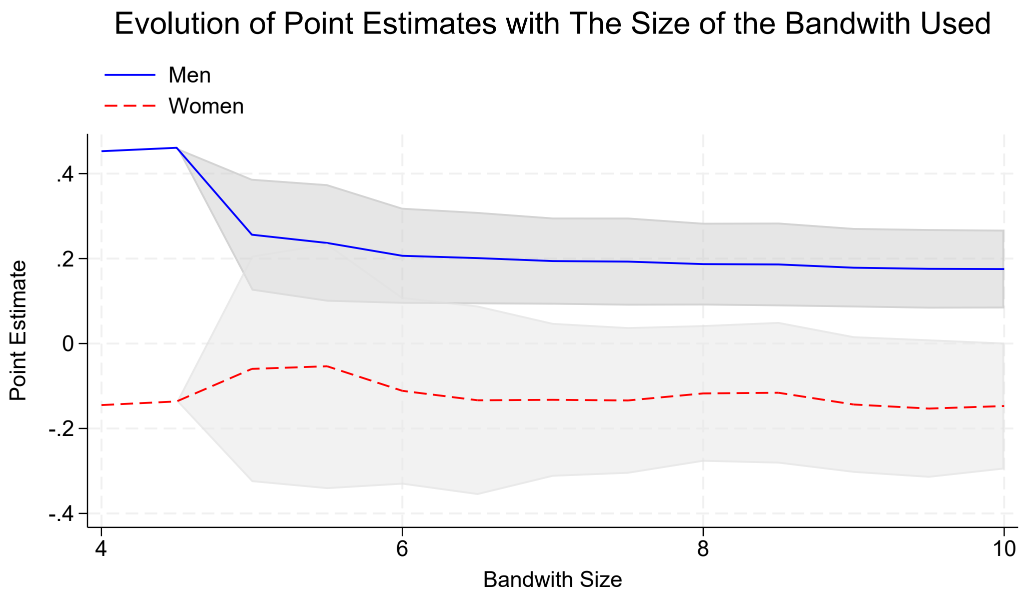
Notes: This table displays the fuzzy non-parametric regression discontinuity estimates of admission probability to top-tier STEM graduate schools. These estimates are based on Calonico et al. (2017, 2019). Selectivity of schools is measured using the percentile rank from high school graduation exam results of admitted students. The running variable is defined as the distance between a student's rank and the cutoff rank for star class admission, computed at the *class* $\times$ *track* $\times$ *prep program* $\times$ *year*. We include controls for percentile rank in mathematics, physics and foreign language at the second semester first year in prep program, and prep program fixed effects. Unlike our main analysis, here we do not cluster standard errors at the running variable level.

Table D33: RDD Robustness Check: Local Randomization

	(1) Top 10 decile of selectivity (All students)	(2) Top 10 decile of selectivity (Male students)	(3) Top 10 decile of selectivity (Female students)
Local randomization estimate	0.21***	0.23***	0.13***
Number of obs.	2372	1743	629

Notes: This table displays the local randomization estimates of admission probability to top-tier STEM graduate schools. These estimates are based on Cattaneo et al. (2016). Due to its very local nature, we must include students just around the cutoff for this robustness check. We use only the sample of students whose distance to the estimated cutoff is between -4 and 4. Selectivity of schools is measured using the percentile rank from high school graduation exam results of admitted students. The running variable is defined as the distance between a student's rank and the cutoff rank for star class admission, computed at the *class* $\times$ *track* $\times$ *prep program* $\times$ *year*.

Figure D15: RDD Robustness Check: Evolution of the Point Estimates with the Bandwidth Size



Notes: The graph illustrates how our main point estimates depends on the bandwidth. These point estimates are based on Calonico et al. (2017, 2019). In alignment with recent advancements in regression discontinuity literature, we rely on bias-corrected estimates and robust standard errors. School selectivity is measured using the percentile rank from high school graduation exam results of admitted students. The running variable is defined as the distance between a student's rank and the cutoff rank for star class admission, computed at the *class* $\times$ *track* $\times$ *prep program* $\times$ *year*. Our analysis includes controls for program fixed effects. Standard errors are clustered at the level of the running variable.

Table D34: RDD Robustness Check: Twice the Bandwidth Size

	(1) Top 10 decile of selectivity (All students)	(2) Top 10 decile of selectivity (Male students)	(3) Top 10 decile of selectivity (Female students)
Baseline mean (Below cutoff)	0.11	0.11	0.09
Baseline RD estimate	0.136*** (0.047)	0.193*** (0.053)	-0.134 (0.088)
Robust 95% CI	[0.043 ; 0.229]	[0.089 ; 0.296]	[-0.307 ; 0.039]
Obs. used in estimation	3,962	2,509	947
Total number of obs.	6,782	4,934	1,848

Notes: This table displays the fuzzy non-parametric regression discontinuity estimates of admission probability to top-tier STEM graduate schools. These estimates are based on Calonico et al. (2017, 2019). Selectivity of schools is measured using the percentile rank from high school graduation exam results of admitted students. The running variable is defined as the distance between a student's rank and the cutoff rank for star class admission, computed at the *class* $\times$ *track* $\times$ *prep program* $\times$ *year*. We control for prep program fixed effects. Unlike our main analysis, here we take twice the optimal bandwidth selected by the procedure defined by Calonico et al. (2017, 2019) for local polynomial regression discontinuity estimators.

Table D35: RDD Robustness Check: Thrice the Bandwidth Size

	(1) Top 10 decile of selectivity (All students)	(2) Top 10 decile of selectivity (Male students)	(3) Top 10 decile of selectivity (Female students)
Baseline mean (Below cutoff)	0.11	0.11	0.09
Baseline RD estimate	0.090* (0.053)	0.154*** (0.054)	-0.135** (0.069)
Robust 95% CI	[-0.014 ; 0.194]	[0.049 ; 0.259]	[-0.270 ; -0.000]
Obs. used in estimation	5,650	3,806	1,472
Total number of obs.	6,782	4,934	1,848

Notes: This table displays the fuzzy non-parametric regression discontinuity estimates of admission probability to top-tier STEM graduate schools. These estimates are based on Calonico et al. (2017, 2019). Selectivity of schools is measured using the percentile rank from high school graduation exam results of admitted students. The running variable is defined as the distance between a student's rank and the cutoff rank for star class admission, computed at the *class* × *track* × *prep program* × *year*. We control for prep program fixed effects. Unlike our main analysis, here we take thrice the optimal bandwidth selected by the procedure defined by Calonico et al. (2017, 2019) for local polynomial regression discontinuity estimators.

Table D36: RDD Robustness Check: Polynomial of Order 2 for the Point Estimator

	(1) Top 10 decile of selectivity (All students)	(2) Top 10 decile of selectivity (Male students)	(3) Top 10 decile of selectivity (Female students)
Baseline mean (Below cutoff)	0.11	0.11	0.09
Baseline RD estimate	0.285*** (0.076)	0.212*** (0.060)	-0.082 (0.095)
Robust 95% CI	[0.135 ; 0.435]	[0.094 ; 0.330]	[-0.268 ; 0.103]
Obs. used in estimation	1,778	2,110	814
Total number of obs.	6,782	4,934	1,848

Notes: This table displays the fuzzy non-parametric regression discontinuity estimates of admission probability to top-tier STEM graduate schools. These estimates are based on Calonico et al. (2017, 2019). Selectivity of schools is measured using the percentile rank from high school graduation exam results of admitted students. The running variable is defined as the distance between a student's rank and the cutoff rank for star class admission, computed at the *class* × *track* × *prep program* × *year*. We control for prep program fixed effects. Unlike our main analysis, here we use a polynomial of order 2 (instead of 1 in our main analysis) for the local polynomial used to compute the point estimator.

Table D37: RDD Robustness Check: Control for Unbalanced Grades and Demographics

	(1) Top 10 decile of selectivity (All students)	(2) Top 10 decile of selectivity (Male students)	(3) Top 10 decile of selectivity (Female students)
Baseline mean (Below cutoff)	0.11	0.11	0.09
Baseline RD estimate	0.120** (0.050)	0.187*** (0.055)	-0.104* (0.063)
Robust 95% CI	[0.021 ; 0.219]	[0.079 ; 0.294]	[-0.227 ; 0.020]
Obs. used in estimation	1,777	1,302	646
Total number of obs.	6,782	4,934	1,848

Notes: This table displays the fuzzy non-parametric regression discontinuity estimates of admission probability to top-tier STEM graduate schools. These estimates are based on Calonico et al. (2017, 2019). Selectivity of schools is measured using the percentile rank from high school graduation exam results of admitted students. The running variable is defined as the distance between a student's rank and the cutoff rank for star class admission, computed at the *class* $\times$ *track* $\times$ *prep program* $\times$ *year*. We include controls for percentile rank in mathematics, physics and foreign language at the second semester first year in prep program,and prep program $\times$ class fixed effects. Standard errors are clustered at the level of the running variable.

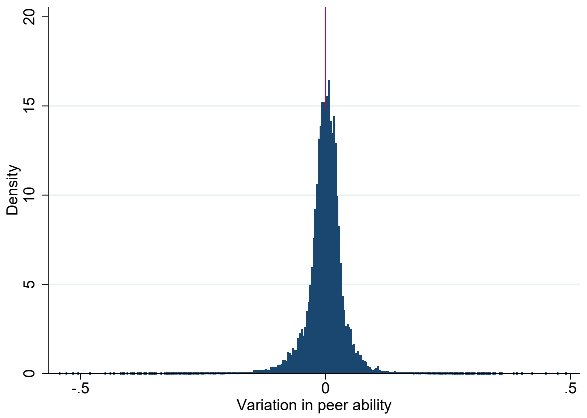
Table D38: RDD Robustness Check: Within Class and Year RDD

	(1) Top 10 decile of selectivity (All students)	(2) Top 10 decile of selectivity (Male students)	(3) Top 10 decile of selectivity (Female students)
Baseline mean (Below cutoff)	0.11	0.11	0.09
Baseline RD estimate	0.141*** (0.046)	0.268*** (0.047)	-0.094* (0.048)
Robust 95% CI	[0.050 ; 0.231]	[0.176 ; 0.360]	[-0.188 ; 0.000]
Obs. used in estimation	1,778	872	316
Total number of obs.	6,782	4,934	1,848

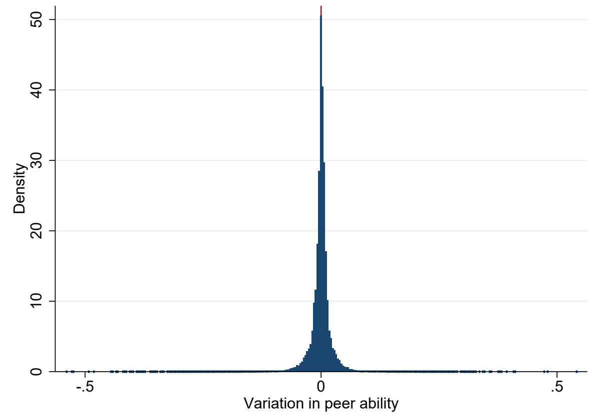
Notes: This table displays the fuzzy non-parametric regression discontinuity estimates of admission probability to top-tier STEM graduate schools. These estimates are based on Calonico et al. (2017, 2019). Selectivity of schools is measured using the percentile rank from high school graduation exam results of admitted students. The running variable is defined as the distance between a student's rank and the cutoff rank for star class admission, computed at the *class* $\times$ *track* $\times$ *prep program* $\times$ *year*. Unlike our main analysis, here we do not include prep programs fixed effect, but we include prep programs $\times$ class $\times$ year fixed effects, since it is the level of our analysis and the one used to identify the RD threshold.

E5 Peer Effects: Complementary Results

Figure E16: Variation in Peer Ability Used



(a) Using year and track × program FE



(b) Using track × program, year × program, and year × track FE

Notes: These histograms displays the year-to-year variation in peer ability within a program and a track. This is the residuals of a regression of peer ability on year, and track × program fixed effects (Figure (a)) or on track × program, year × program, and year × track fixed effects (Figure (b)). Peer ability is defined based on leave-own-out reweighted GPA obtained at the high school graduation exam of first year peers with the most selective competitive exam coefficients, measured before individuals enter the prep programs.

Table E39: Peer Effect on Admission to Top STEM Graduate Schools, by Gender: Robustness Test Interacting the Fixed Effects

	(1) Top 10% grad. schools (Selectivity)	(2) Top 10% grad. schools (Selectivity)	(3) Top 10% grad. schools (Desirability)	(4) Top 10% grad. schools (Desirability)
Baseline proba. of access	0.15	0.15	0.16	0.16
Average percentile rank of 1st year peers at the HS graduation exam	0.11***	0.051**	0.10***	0.042
	(0.018)	(0.024)	(0.018)	(0.026)
Female student × Average percentile rank of 1st year peers	-0.059***	-0.056***	-0.059***	-0.050***
	(0.018)	(0.018)	(0.018)	(0.018)
Demographic Controls	✓	✓	✓	✓
MS & HS Exam Score	✓	✓	✓	✓
Cohort Controls	✓	✓	✓	✓
Demographic Controls × Female student	✓	✓	✓	✓
MS & HS Exam Score × Female student	✓	✓	✓	✓
Cohort Controls × Female student	✓	✓	✓	✓
Year FE	✓		✓	
Program × Track FE	✓	✓	✓	✓
Year × Track FE		✓		✓
Year × Program FE		✓		✓
N	165,031	165,031	165,031	165,031

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: The table reports the change in the probability of admission to top 10% most selective STEM graduate schools for students with higher-achieving peers, by gender. Peer achievement is defined based on the average percentile rank of students at the national high school examination, measured before they enter prep programs. Grades at the national high school examination are re-weighted with the coefficients of the most selective competitive exams (see Table A1). The peer group is defined at the 1st year track × year × program level, excluding the individual under consideration (leave-own-out mean). Selectivity of schools is measured in two ways: Columns (1) and (2) use the average percentile rank of admitted students at the high school graduation exam, while Columns (3) and (4) are based on the revealed preferences of applicants (Avery et al. (2013)). Demographic controls include geographic origin (Paris or Parisian area), low-income status, the socio-economic status of each parent (into four categories), French nationality, repeater status, disability status, whether the student was in a science academic track during high school, the student's option (either engineering science or computer science), and the gender composition in the *track × program*. Previous ability is controlled by (i) decile rank dummy variables from both high school and middle school graduation exams GPA, and (ii) quintile rank dummy variables for each grades at the high and middle school graduation exams at the subjects studied in prep programs — mathematics, physics and chemistry, engineering science, French (both written and oral), and foreign languages. Cohort controls include the proportion of female students, the proportion of French students, the proportion of high SES students, the proportion of students from Paris and the proportion of students from the Parisian area. Cohort controls and peer characteristics are measured in 1st year of prep program. We include interactions of all demographics, previous achievement and cohort controls with a gender dummy variable, allowing observable characteristics to have gender-specific performance impacts. We do not interact fixed effects with gender since this is the source of variation we are using. We include year, and track × program fixed effects in all specification. In Columns (2) and (4) track fixed effects are interacted with programs fixed effects, year fixed effects are interacted with programs fixed effects and year fixed effects are interacted with track fixed effects to allow the average quality of programs to vary over time and over track. Contrary to the rest of the analysis, these fixed effects are determined based on the first year of the program, as this corresponds to the level of peer variation we are using. Standard errors are clustered at the *track × program* level, also corresponding to the first year in the program.

Table E40: Effect of Having Higher-Achieving Peers on Access to Top STEM Graduate Schools, by Gender on Restricted Sample

	(1) Top 10% grad. schools (Selectivity)	(2) Top 10% grad. schools (Selectivity)	(3) Top 10% grad. schools (Desirability)	(4) Top 10% grad. schools (Desirability)
Baseline proba. of access	0.22	0.22	0.23	0.23
Average percentile rank of 1st year peers at the HS graduation exam	0.35***		0.28***	
	(0.046)		(0.047)	
Female student × Average percentile rank of 1st year peers	-0.14***		-0.12***	
	(0.042)		(0.042)	
Average percentile rank of 1st year female peers at the HS graduation exam		0.024		0.0088
		(0.033)		(0.033)
Female student × Average percentile rank of 1st year female peers		0.089*		0.096*
		(0.046)		(0.049)
Average percentile rank of 1st year male peers at the HS graduation exam		0.32***		0.28***
		(0.047)		(0.048)
Female student × Average percentile rank of 1st year male peers		-0.23***		-0.21***
		(0.057)		(0.060)
Demographic Controls	✓	✓	✓	✓
MS & HS Exam Score	✓	✓	✓	✓
Cohort Controls	✓	✓	✓	✓
Demographic Controls × Female student	✓	✓	✓	✓
MS & HS Exam Score × Female student	✓	✓	✓	✓
Cohort Controls × Female student	✓	✓	✓	✓
Program × Track FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
N	89,065	89,065	89,065	89,065

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: The table reports the change in the probability of admission to 10% most selective STEM graduate schools for students with higher-achieving peers, with results presented by gender in Columns (1) and (3), and further broken down by the gender of the high-achieving peers in Columns (2) and (4). It displays the same results as in Table 9 but on the restricted sample of programs having both a star and a standard class. Peer achievement is defined based on the average percentile rank of students at the national high school examination, measured before they enter prep programs. Grades at the national high school examination are re-weighted with the coefficients of the most selective competitive exams (see Table A1). The peer group is defined at the 1st year track × year × program level, excluding the individual under consideration (leave-own-out mean). Selectivity of schools is measured in two ways: Columns (1) and (2) use the average percentile rank of admitted students at the high school graduation exam, while Columns (3) and (4) are based on the revealed preferences of applicants (Avery et al. (2013)). Demographic controls include geographic origin (Paris or Parisian area), low-income status, the socio-economic status of each parent (into four categories), French nationality, repeater status, disability status, whether the student was in a science academic track during high school, the student's option (either engineering science or computer science), and the gender composition in the *track × program*. Previous ability is controlled by (i) decile rank dummy variables from both high school and middle school graduation exams GPA, and (ii) quintile rank dummy variables for each grades at the high and middle school graduation exams at the subjects studied in prep programs — mathematics, physics and chemistry, engineering science, French (both written and oral), and foreign languages. Cohort controls include the proportion of female students, the proportion of French students, the proportion of high SES students, the proportion of students from Paris and the proportion of students from the Parisian area. Cohort controls and peer characteristics are measured in 1st year of prep program. We include interactions of all demographics, previous achievement and cohort controls with a gender dummy variable, allowing observable characteristics to have gender-specific performance impacts. We do not interact fixed effects with gender since this is the source of variation we are using. We include year, and track × program fixed effects in all specification. Contrary to the rest of the analysis, these fixed effects are determined based on the first year of the program, as this corresponds to the level of peer variation we are using. Standard errors are clustered at the *track × program* level, also corresponding to the first year in the program.

Appendix References

Avery, Christopher N., Mark E. Glickman, Caroline M. Hoxby, and Andrew Metrick, "A Revealed Preference Ranking of U.S. Colleges and Universities," *The Quarterly Journal of Economics*, 2013, *128* (1), 425–467.

Bütikofer, Aline, Rita Ginja, Katrine V. Løken, and Fanny Landaud, "Higher-Achievement Schools, Peers and Mental Health," *The Economic Journal*, 2023, *133* (655), 2580–2613.

Calonico, Sebastian, Matias D. Cattaneo, Max H. Farrell, and Rocio Titiunik, "Rdrobust: Software for Regression-Discontinuity Designs," *The Stata Journal*, 2017, *17* (2), 372–404.

_ , _ , _ , and _ , "Regression Discontinuity Designs Using Covariates," *Review of Economics and Statistics*, 2019, *101* (3), 442–451.

Cattaneo, Matias D., Michael Jansson, and Xinwei Ma, "Manipulation Testing Based on Density Discontinuity," *The Stata Journal*, 2018, *18* (1), 234–261.

_ , Rocio Titiunik, and Gonzalo Vazquez-Bare, "Inference in Regression Discontinuity Designs under Local Randomization," *The Stata Journal*, 2016, *16* (2), 331–367.

Hoekstra, Mark, "The Effect of Attending the Flagship State University on Earnings: A Discontinuity-Based Approach," *The Review of Economics and Statistics*, 2009, *91* (4), 717–724.

F6 List of Tables and Figures

List of Figures

List of Tables