

Gendered Choices in Gender-Blind University
Admissions
Online Appendix

Online Appendices

A	Additional results on data and empirical strategy	2
B	Additional results on gender differences in university choices	4
C	Additional results on economic consequences of gendered university choices	19
D	Additional results on what explains gendered patterns in university choices	22
E	Survey instruments	39

A Additional results on data and empirical strategy

Table A1: Comparison of sample county with national and provincial averages

Variable	National		Henan		Sample County
	Mean (1)	Median (2)	Mean (3)	Median (4)	
GDP per capita	3.97	3.17	3.55	3.02	4.30
Secondary industry share	32.16	31.38	38.31	38.56	39.50
No. of above-scale industrial enterprises	73.74	43.00	117.40	105.00	46.00
Savings balance per capita	3.35	3.05	2.83	2.61	3.84
Average years of education	8.33	8.46	8.60	8.46	8.82
Illiteracy rate	7.03	4.22	3.62	3.39	4.44
Gender ratio (females = 100)	105.97	105.04	100.82	100.29	99.68
Share of population aged 0–14	20.33	20.36	24.59	25.04	21.75
Share of population aged 15–64	65.86	65.39	60.77	60.04	62.90
Share of population aged 65+	13.81	14.12	14.64	14.53	15.35

Notes: This table compares key economic and demographic indicators of the sample county with averages across all counties in China and Henan Province. Columns (1) and (2) report mean and median values at the national level; columns (3) and (4) report mean and median values for Henan Province; column (5) presents statistics for the sample county. Economic indicators (rows 1–4) are for 2020. Demographic indicators (rows 5–10) are from the 2020 National Census. “GDP per capita” and “Savings balance per capita” are measured in 10,000 Chinese Yuan.

Table A2: Survey respondents versus administrative sample

	Total			Male			Female			Diff-in-Diff
	Admin (1)	Survey (2)	Difference (3)	Admin (4)	Survey (5)	Difference (6)	Admin (7)	Survey (8)	Difference (9)	(10)
STEM track	0.806 (0.396)	0.826 (0.379)	0.021* (0.012)	0.927 (0.261)	0.937 (0.244)	0.010 (0.011)	0.683 (0.465)	0.720 (0.449)	0.037* (0.019)	0.027 (0.022)
Repeat exam-taker	0.065 (0.246)	0.012 (0.109)	-0.053*** (0.005)	0.075 (0.263)	0.017 (0.130)	-0.058*** (0.008)	0.054 (0.227)	0.007 (0.084)	-0.047*** (0.006)	0.010 (0.010)
Local Hukou	0.974 (0.158)	0.975 (0.155)	0.001 (0.005)	0.976 (0.152)	0.978 (0.147)	0.002 (0.006)	0.972 (0.164)	0.973 (0.162)	0.001 (0.007)	-0.001 (0.009)
Age	18.138 (0.636)	18.080 (0.586)	-0.058*** (0.018)	18.179 (0.637)	18.118 (0.591)	-0.061** (0.026)	18.096 (0.632)	18.042 (0.579)	-0.054** (0.026)	0.007 (0.037)
Poverty status	0.074 (0.262)	0.092 (0.290)	0.018** (0.009)	0.077 (0.267)	0.092 (0.289)	0.014 (0.012)	0.071 (0.256)	0.093 (0.291)	0.023* (0.012)	0.008 (0.017)
Gaokao percentile rank	0.216 (0.158)	0.217 (0.154)	0.001 (0.005)	0.211 (0.155)	0.212 (0.149)	0.001 (0.007)	0.222 (0.161)	0.223 (0.158)	0.001 (0.007)	0.000 (0.009)
Observations	3,112	1,668		1,566	819		1,546	849		

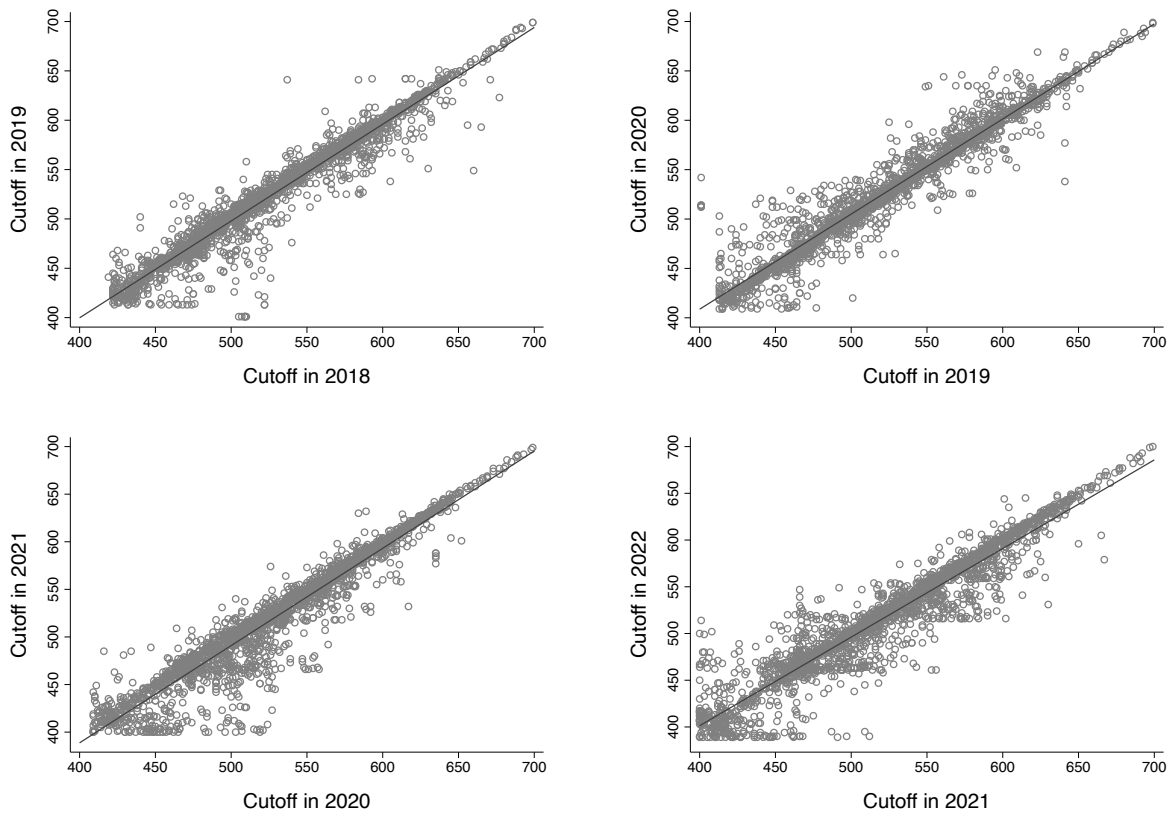
Notes: This table compares individual characteristics between the survey and administrative samples. The administrative sample is restricted to 2022–2024 to match the survey period. Columns (1), (4), and (7) report the administrative sample; columns (2), (5), and (8) report the survey sample; columns (3), (6), and (9) report differences between the two. Column (10) reports the difference-in-differences estimator (Female Diff – Male Diff), testing for gender heterogeneity in sample selection. “STEM track” is an indicator for students in the STEM track. “Repeat exam-taker” indicates students who took the Gaokao more than once. “Local Hukou” indicates students with local residency. “Poverty status” indicates students with official poverty status. “Gaokao percentile rank” is the student’s percentile rank on the Gaokao. Standard deviations in parentheses for columns (1)–(2), (4)–(5), and (7)–(8); standard errors in parentheses for columns (3), (6), (9), and (10). *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

B Additional results on gender differences in university choices

B.1 Estimation of admission probability

This section outlines our method for estimating a student's admission probability to a university, which serves as a measure of the riskiness of students' ROLs. We begin by examining the predictive power of historical admission cutoffs, then model these cutoffs using normal distributions, estimate their parameters, and calculate admission probabilities for each student-university pair. Finally, we validate our estimates against actual admission outcomes.

Figure B1: Admission cutoffs across years



Notes: This figure shows the relationship between university admission cutoffs in consecutive years from 2018 to 2022. Each panel compares cutoffs for two adjacent years; each dot represents a university. All cutoffs are converted to 2018 rank-preserving equivalents to account for year-to-year fluctuations in overall Gaokao performance. Top-left panel: 2019 vs. 2018; top-right: 2020 vs. 2019; bottom-left: 2021 vs. 2020; bottom-right: 2022 vs. 2021.

To estimate admission probabilities, we rely on the predictive power of past admission cutoffs. This approach is valid only if previous years' cutoffs reliably predict current cutoffs. Figure B1 displays this relationship for all universities in our sample from 2018 to 2022, plotting each university's admission cutoff against its cutoff from the previous year. Because average Gaokao

scores fluctuate annually, we normalize all scores to their rank-preserving equivalents in 2018—adjusting scores so that a score at a given percentile in any year corresponds to the score at the same percentile in 2018, ensuring comparability across years. The strong correlation (approximately 0.95) between consecutive years’ adjusted cutoffs validates our use of historical data in predicting current admission thresholds.

To model admission cutoffs, we, like Wang et al. (2025), assume that the annual admission cutoff c_{jt} for university j in year t follows a normal distribution:

$$c_{jt} \sim N(\mu_j, \sigma_j^2).$$

We estimate μ_j using the average admission cutoff for university j during 2015–2023. We do not claim that cutoffs are actually normally distributed; rather, we adopt the perspective of students who observe year-to-year fluctuations without knowing their underlying causes. From this perspective, treating cutoffs as random draws from a normal distribution is intuitive and provides a reasonable approximation of students’ uncertainty. To assess this approximation, we examine the distribution of deviations between realized cutoffs and predicted values in Figure B2. The distribution is bell-shaped and centered around zero with relatively thin tails, suggesting that a normal distribution reasonably captures the variation students face.

We estimate σ_j^2 as the sample variance of university j ’s admission cutoffs over the estimation window. Since each variance is based on only a few annual observations, individual estimates can be noisy. To improve precision, we apply empirical Bayes shrinkage, which pulls each university’s variance toward the cross-sectional mean:

$$\hat{\sigma}_j^2 = w \cdot s_j^2 + (1 - w) \cdot \bar{s}^2$$

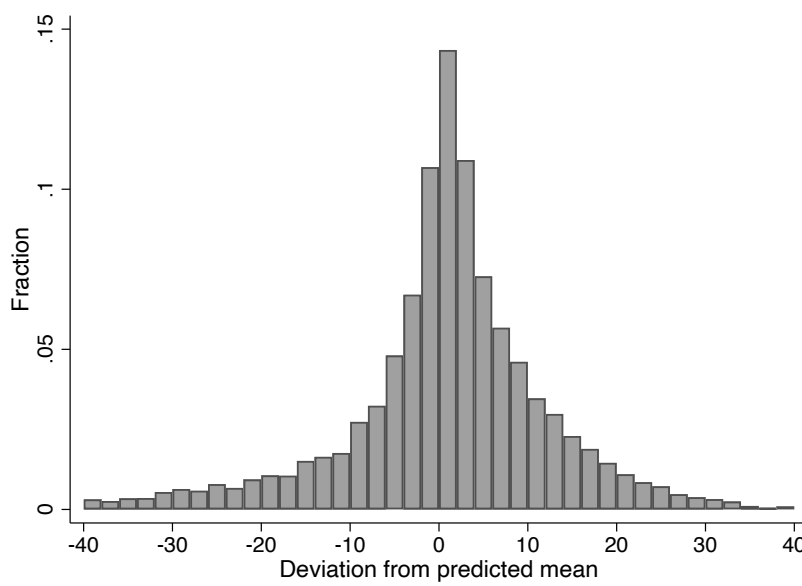
where s_j^2 is the sample variance for university j , $\bar{s}^2$ is the grand mean variance across all universities, and $w \in [0, 1]$ is a shrinkage weight estimated via method of moments under an inverse-gamma prior for the true variance distribution. Universities with variance estimates far from the cross-sectional mean are pulled toward the overall average, reducing the influence of sampling noise on the estimated probabilities.

Given the estimated distribution of admission cutoffs for university j , the probability of student i with Gaokao score s_i being admitted to university j is:

$$\hat{p}_{ij} \equiv \Phi\left(\frac{s_i - \mu_j}{\hat{\sigma}_j}\right)$$

where $\Phi(\cdot)$ is the CDF of the standard normal distribution.

Figure B2: Distribution of realized cutoff minus predicted cutoff



Notes: This figure shows the distribution of deviations between realized admission cutoffs and predicted values for all universities in our sample. The predicted cutoff for each university is the average of its admission cutoffs from 2015 to 2023, with all scores standardized to 2018 rank-preserving equivalents to ensure comparability across years. Extreme outliers deviating more than 40 points from the predicted mean are excluded.

To assess the validity of these estimates, we regress actual admission outcomes on estimated probabilities:

$$\text{Admitted}_{ij} = \alpha + \theta \hat{p}_{ij} + \epsilon_{ij}$$

where Admitted_{ij} is an indicator for student i being admitted to the university listed as their j th choice. If our estimated admission probabilities are valid, then $\theta = 1$.

Table B1 shows that our estimates provide a reasonable approximation of actual admission probabilities. Columns (1)–(4) regress admission to first-choice universities on estimated probability. The estimated θ is approximately 0.94–0.95 and significantly different from zero. Columns (5)–(8) regress admission to second-choice universities on estimated probability, using the subsample of students not admitted to their first choice. The estimated θ is again significantly different from zero, typically around 0.88. These results indicate that our model captures the empirical patterns well.

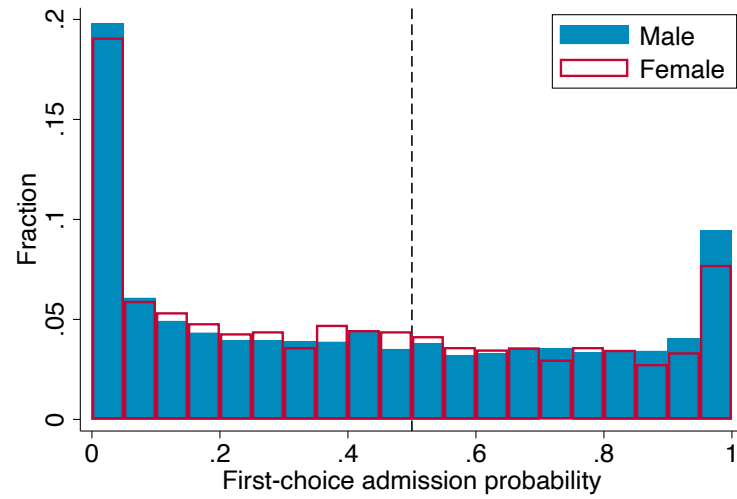
Table B1: Validity of estimated admission probabilities

Dependent variable	Admitted to first choice				Admitted to second choice			
	Full (1)	Tier-1 & 2 (2)	Tier-1 (3)	Tier-2 (4)	Full (5)	Tier-1 & 2 (6)	Tier-1 (7)	Tier-2 (8)
Panel (a): Full sample								
Estimated probability	0.952*** (0.011)	0.939*** (0.013)	0.940*** (0.021)	0.936*** (0.017)	0.876*** (0.017)	0.876*** (0.022)	0.945*** (0.031)	0.813*** (0.030)
Observations	10,121	8,372	3,591	4,781	5,611	4,171	1,916	2,255
R-squared	0.411	0.381	0.365	0.389	0.310	0.284	0.331	0.245
Panel (b): STEM track								
Estimated probability	0.940*** (0.012)	0.918*** (0.014)	0.919*** (0.022)	0.911*** (0.019)	0.885*** (0.019)	0.885*** (0.024)	0.945*** (0.032)	0.817*** (0.035)
Observations	8,351	6,825	3,146	3,679	4,451	3,218	1,625	1,593
R-squared	0.416	0.384	0.363	0.397	0.324	0.304	0.349	0.258
Panel (c): Non-STEM track								
Estimated probability	1.007*** (0.030)	1.031*** (0.035)	1.125*** (0.070)	1.008*** (0.040)	0.870*** (0.042)	0.886*** (0.051)	0.915*** (0.102)	0.891*** (0.059)
Observations	1,770	1,547	445	1,102	1,160	953	291	662
R-squared	0.392	0.366	0.371	0.364	0.275	0.243	0.219	0.260

Notes: This table tests the validity of estimated admission probabilities. Columns (1)–(4) regress an indicator for admission to first-choice universities on estimated admission probability. Columns (5)–(8) regress an indicator for admission to second-choice universities on estimated probability, using the subsample not admitted to their first choice. “Full” includes all students; “Tier-1 & 2” restricts to students in Tier-1 or Tier-2; “Tier-1” and “Tier-2” further restrict to each tier separately. Panel (a) includes all students; panels (b) and (c) restrict to STEM and non-STEM track students, respectively. Standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

B.2 Additional results on application strategies

Figure B3: Gender differences in first-choice risk-taking



Notes: This figure reports the distribution of admission probabilities for first-choice universities by gender.

B.3 Alternative measure of risk-taking: Score margin

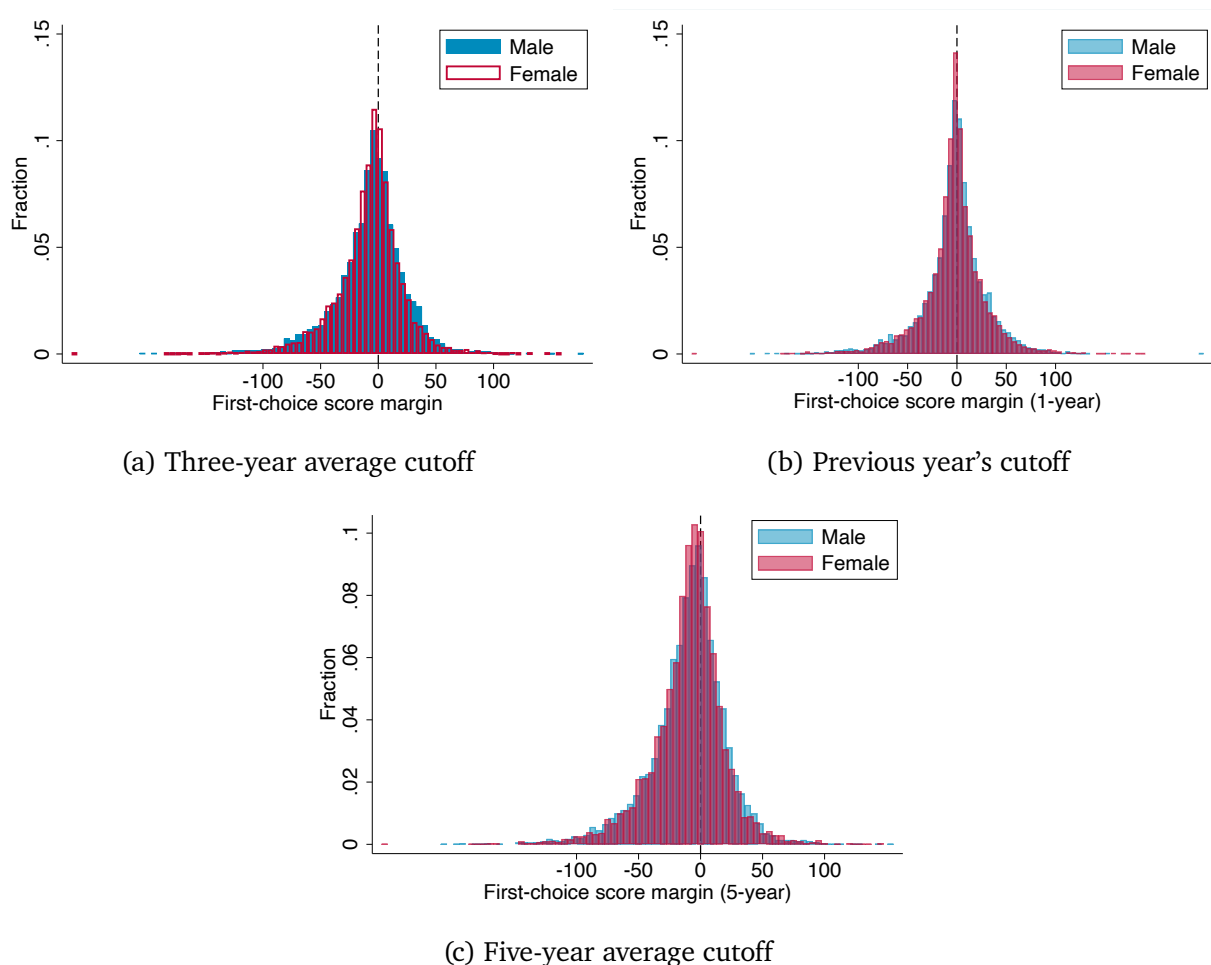
As a model-free alternative to the admission probability measure used in the main analysis, we also assess risk-taking using the *score margin*—the raw difference between a student’s exam score and a university’s historical admission cutoff. Unlike the probability measure, the score margin requires no distributional assumptions and is directly interpretable in score points. We define it as:

$$\text{Score Margin}_{ij} = \text{Gaokao Score}_i^{\text{std}} - \text{Cutoff}_j^{\text{std}}$$

where $\text{Cutoff}_j^{\text{std}}$ is university j ’s average standardized admission cutoff over the previous three years, with all scores standardized to 2018 equivalents to ensure comparability across cohorts. A negative margin indicates a “reach” school where admission is unlikely; a positive margin indicates relative safety.

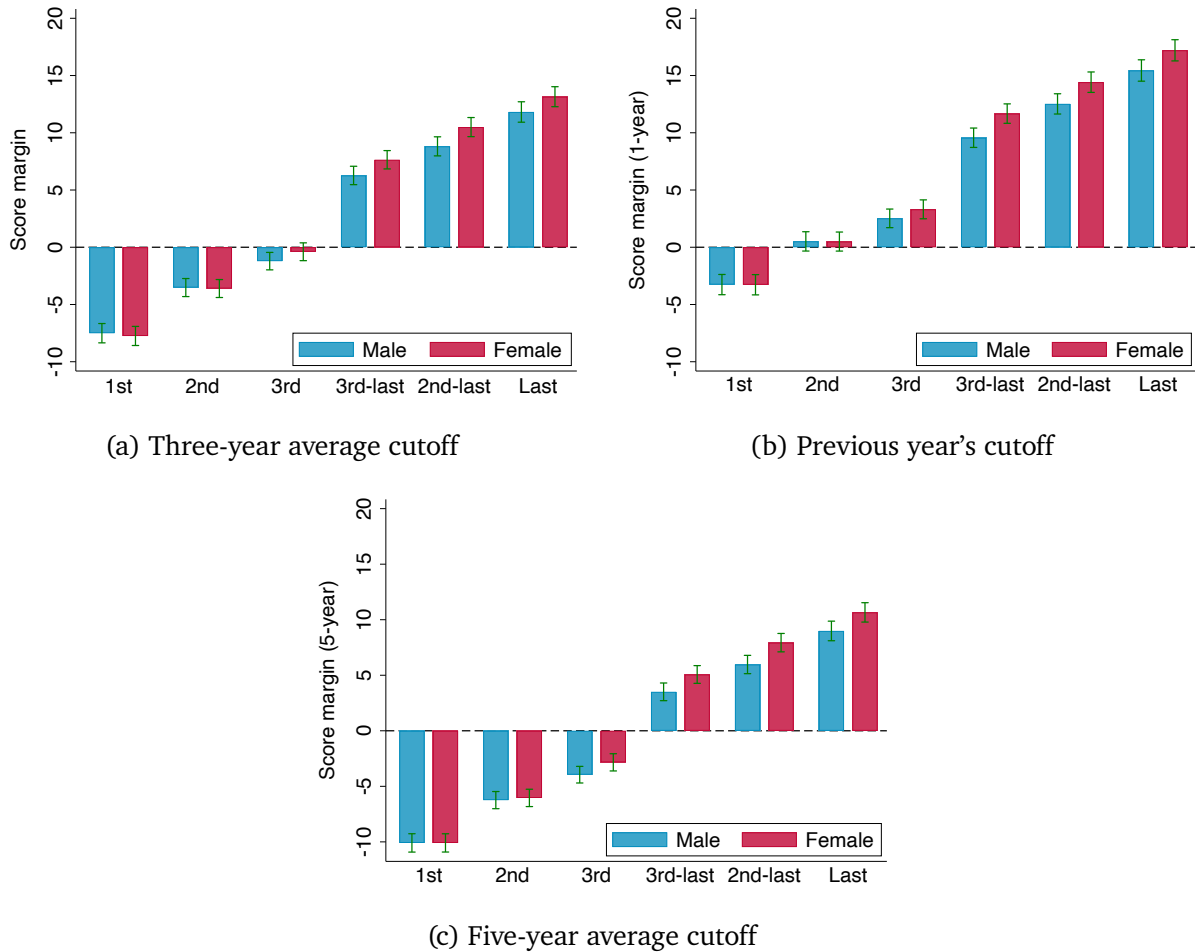
Figure B4 shows the distribution of first-choice score margins by gender. Under all three cutoff definitions—three-year average (Panel a), previous year (Panel b), and five-year average (Panel c)—the distributions for men and women overlap almost entirely, confirming the null result from the probability analysis. Figure B5 extends the analysis to the first three and last three positions on the rank-order list. Men and women target universities with nearly identical score margins across all positions, regardless of the cutoff window used. Together, these results confirm that the absence of gender differences in risk-taking documented in the main text is robust to an alternative, assumption-free measure of admission difficulty.

Figure B4: Gender differences in first-choice score margin: alternative cutoff windows



Notes: This figure shows the distribution of score margins for first-choice universities by gender using different cutoff windows. Score margin is the difference between a student's standardized Gaokao score and the university's historical admission cutoff, with all scores standardized to 2018 equivalents. Panel (a) uses cutoff over the previous three years; Panel (b) use the previous year's cutoff; panel (c) uses the average cutoff over the previous five years. Negative values indicate reach schools where the student's score falls below the historical cutoff.

Figure B5: Gender differences in score margin by position: alternative cutoff windows



Notes: This figure shows average score margins for the first three and last three choices on students' ROLs by gender using different cutoff windows. Score margin is the difference between a student's standardized Gaokao score and the university's historical admission cutoff, with all scores standardized to 2018 equivalents. Panel (a) uses cutoff over the previous three years; Panel (b) use the previous year's cutoff; panel (c) uses the average cutoff over the previous five years. Negative values indicate universities where the student's score falls below the historical cutoff. Error bars represent 95% confidence intervals.

B.4 Additional results on gender differences in university choices and admissions

Table B2: Gender differences in university applications

Dependent variable	Elite university		Standardized distance		STEM-focused university	
	(1)	(2)	(3)	(4)	(5)	(6)
Female	−0.014*** (0.004)	−0.015*** (0.004)	−0.127*** (0.012)	−0.127*** (0.012)	−0.152*** (0.005)	−0.151*** (0.005)
Gaokao ranking FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Choice-order FE	✓	✓	✓	✓	✓	✓
Demographic controls		✓		✓		✓
Observations	83,701	83,701	83,701	83,701	83,701	83,701
R-squared	0.539	0.539	0.040	0.040	0.057	0.057

Notes: This table reports regression results for the likelihood of applying to elite, distant, and STEM-focused universities by gender, controlling for Gaokao performance. The unit of observation is a student-choice. The dependent variable is an indicator for elite university in columns (1)–(2), standardized distance in columns (3)–(4), and an indicator for STEM-focused university in columns (5)–(6). “Female” is an indicator for female students. Gaokao ranking FE include indicators for each percentile of the Gaokao ranking, plus a dummy for high-school track. Year FE are fixed effects for year of Gaokao. Choice-order FE are fixed effects for position on the ROL. Demographic controls include age, residence status, poverty status, and repeat exam-taker status. Standard errors clustered at the student level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table B3: Gender differences in university admission

Dependent variable	Elite		Distance		STEM-focused	
	(1)	(2)	(3)	(4)	(5)	(6)
Female	−0.011** (0.005)	−0.011** (0.005)	−0.104*** (0.026)	−0.103*** (0.026)	−0.179*** (0.013)	−0.178*** (0.013)
Gaokao ranking FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Demographic controls		✓		✓		✓
Observations	7,484	7,484	6,287	6,287	6,287	6,287
R-squared	0.534	0.535	0.067	0.067	0.069	0.070

Notes: This table reports regression results for characteristics of admitted universities by gender. The unit of observation is a student. The dependent variable is an indicator for elite university (columns 1–2), standardized distance from home (columns 3–4), and an indicator for STEM-focused university (columns 5–6). “Female” is an indicator for female students. Gaokao ranking FE include indicators for each percentile of the Gaokao ranking and high-school track. Year FE are fixed effects for year of Gaokao. Demographic controls include age, residence status, poverty status, and repeat exam-taker status. Columns (1)–(2) include students not admitted to any university (coded as 0 for elite). Columns (3)–(6) have fewer observations for two reasons: students not admitted to any university are excluded, and for 2019–2021 we only obtained detailed admission outcomes for students admitted to elite universities. Standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table B4: Admission cutoffs of admitted universities by gender

Dependent variable	Five-year average cutoff		Three-year average cutoff		Previous year's cutoff	
	(1)	(2)	(3)	(4)	(5)	(6)
Female	0.492 (0.492)	0.508 (0.495)	0.315 (0.495)	0.276 (0.497)	−0.154 (0.561)	−0.193 (0.566)
Gaokao ranking FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Demographic controls		✓		✓		✓
Observations	5,835	5,835	5,835	5,835	5,835	5,835
R-squared	0.910	0.911	0.909	0.910	0.894	0.894

Notes: This table reports regression results examining the relationship between gender and admission cutoffs of admitted universities. The dependent variable is the average cutoff over the previous five years (columns 1–2), the average cutoff over the previous three years (columns 3–4), and the previous year's cutoff (columns 5–6). “Female” is an indicator for female students. Gaokao ranking FE include indicators for each percentile of the Gaokao ranking and high-school track. Year FE are fixed effects for year of Gaokao. Demographic controls include age, residence status, poverty status, and repeat exam-taker status. The sample excludes students not admitted to any university, students admitted to non-elite universities in 2019–2021 (for whom detailed admission data are unavailable), and students admitted to universities with missing historical cutoff data. Standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

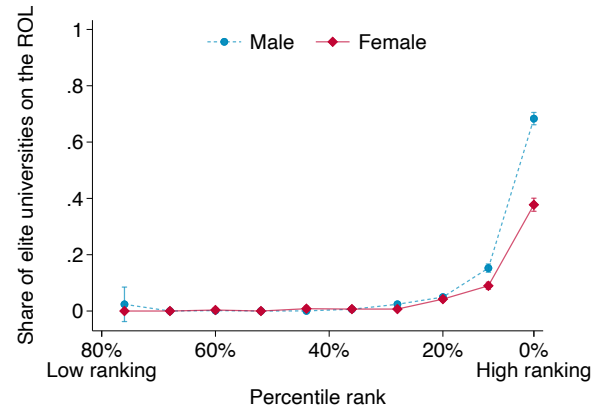
Table B5: Multiple hypothesis testing corrections for the *Female* coefficient

Dependent variable	Coeff.	SE	<i>p</i> -value	<i>p</i> (Holm)	<i>p</i> (W-Y)
Panel (a): University applications (Table B2)					
Elite university	−0.015	(0.004)	<0.001	<0.001	<0.001
Distance from home	−0.127	(0.012)	<0.001	<0.001	<0.001
STEM-focused university	−0.151	(0.005)	<0.001	<0.001	<0.001
Panel (b): University admissions (Table B3)					
Elite university	−0.011	(0.005)	0.023	0.023	0.024
Distance from home	−0.103	(0.026)	<0.001	<0.001	<0.001
STEM-focused university	−0.178	(0.013)	<0.001	<0.001	<0.001

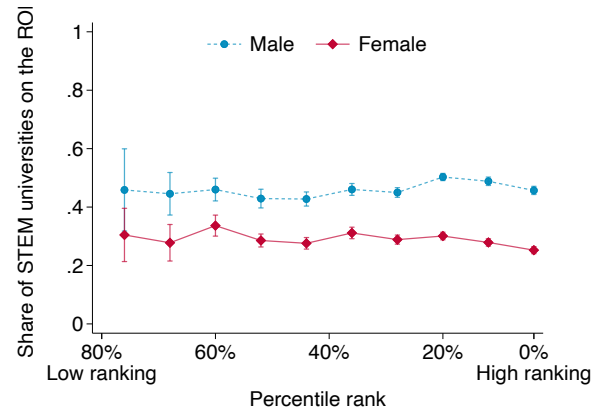
Notes: Each cell reports the coefficient on “Female” from a separate regression of the dependent variable on gender with Gaokao ranking FE (indicators for each percentile of the Gaokao ranking and high-school track), year FE, choice-order FE (Panel A only), and demographic controls (age, residence status, poverty status, and repeat exam-taker status). Holm *p*-values use the Bonferroni–Holm step-down procedure. W-Y *p*-values use the Westfall and Young (1993) free step-down resampling method with 1,000 bootstrap replications. Panel A clusters standard errors at the student level across multiple application choices; Panel B collapses to one observation per student.

B.5 Additional figures and tables

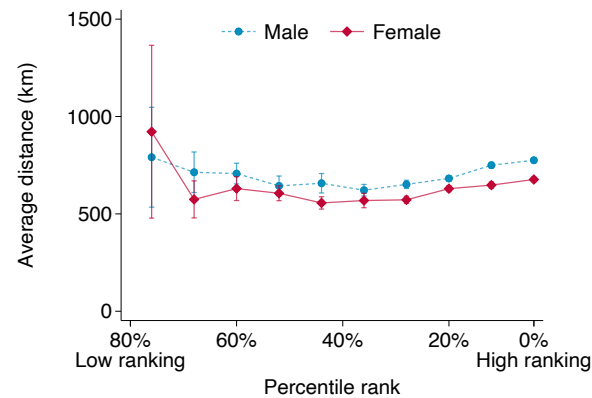
Figure B6: Gender differences in university applications by Gaokao ranking



(a) Elite U application by Gaokao percentile rank



(b) STEM U application by Gaokao percentile rank



(c) Distance of applied U by Gaokao percentile rank

Notes: The figure reports gender differences in university applications by male and female students' Gaokao percentile ranks. Panel (a) shows the share of elite universities in the ROLs. Panel (b) shows the share of STEM universities in the ROLs. Panel (c) shows the average distance of universities on students' ROLs.

Table B6: Gender differences in university applications: Excluding repeat takers and poverty students

Dependent variable	Elite U Excluding poverty (1)	Distance Excluding poverty (2)	STEM U (3)	Elite U Excluding repeat takers (4)	Distance Excluding repeat takers (5)	STEM U (6)	Elite U Excluding both (7)	Distance Excluding both (8)	STEM U (9)
Female	-0.017*** (0.005)	-0.130*** (0.012)	-0.152*** (0.005)	-0.017*** (0.005)	-0.135*** (0.013)	-0.152*** (0.005)	-0.019*** (0.005)	-0.137*** (0.013)	-0.151*** (0.005)
Gaokao ranking FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Choice-Order FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Demographic controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	76,852	76,852	76,852	74,762	74,762	74,762	68,645	68,645	68,645
R-squared	0.542	0.040	0.057	0.539	0.039	0.058	0.542	0.040	0.058

Notes: This table presents regression results analyzing the likelihood of applying to an elite university, standardized distance, and STEM university by gender. Columns (1)-(3) exclude students with poverty status, Columns (4)-(6) exclude repeat exam takers, and Columns (7)-(9) exclude both. The unit of observation is a student choice. The dependent variables are binary indicators for whether the chosen university is elite in Columns (1), (4), (7), the standardized distance of the chosen university in Columns (2), (5), (8), and whether the chosen university is STEM-focused in Columns (3), (6), (9). “Female” is an indicator variable equal to 1 for female students. Gaokao ranking FE include a set of indicator variables for each percentile of the Gaokao ranking, and the high-school track. Year FE are fixed effects for the year of Gaokao. Choice-Order FE are fixed effects for the order of the choice on the ROL. Demographic controls include age, residence status, poverty status, and repeat exam taker status. Standard errors, reported in parentheses, are clustered at the student level to accommodate multiple choices on the ROL per student. *** p<0.01, ** p<0.05, * p<0.1.

C Additional results on economic consequences of gendered university choices

C.1 A simple model of preferences for proximity and university quality

In this section, we develop a simple model to illustrate how stronger preferences for proximity can affect the quality of universities students apply to, where quality is measured by expected post-graduation salary. We present a conceptual framework and provide a numerical example to show why average distance across the ROL better captures proximity preferences than the distance of the admitted university.

A simple model of university choice. Consider a student choosing among a set of universities J , each characterized by quality and distance from home. Let q_j denote the quality of university j (measured by expected salary) and d_j denote the distance from the student's home to university j .

The utility that student i derives from applying to university j is:

$$U_{ij} = \alpha_i q_j - \beta_i d_j,$$

where $\alpha_i > 0$ is the student's preference for quality, and $\beta_i \geq 0$ captures the disutility from distance. We assume $\beta_{\text{female}} > \beta_{\text{male}}$, implying that women have stronger preferences for proximity than men.

Each student i selects a set of K universities $S_i \subseteq J$ to maximize total utility:

$$\max_{S_i} \sum_{j \in S_i} U_{ij}$$

subject to:

$$|S_i| = K$$

Proximity preferences and university quality. Our model demonstrates that as the preference for proximity (β_i) increases, students are more likely to include lower-quality nearby universities in their choice set, reducing the average quality of universities on their ROL. Consider student i 's choice between two universities: university j , which is higher-quality but distant, and university k , which is lower-quality but nearby ($d_k < d_j$). The utility difference is:

$$\Delta U_{ijk} = U_{ij} - U_{ik} = \alpha_i(q_j - q_k) - \beta_i(d_j - d_k)$$

Student i prefers the lower-quality, nearby university k over the higher-quality, distant univer-

sity j if:

$$\Delta U_{ijk} < 0 \implies (q_j - q_k) < \frac{\beta_i}{\alpha_i}(d_j - d_k)$$

This inequality implies that the student chooses the nearby university if the quality difference is less than the product of their relative proximity preference ($\frac{\beta_i}{\alpha_i}$) and the distance difference. As β_i increases, students become willing to accept larger quality sacrifices for proximity. Students with higher β_i (women, given our assumption) are thus more likely to apply to lower-quality nearby universities, resulting in lower average expected salaries across their ROLs.

Numerical example. Consider a simplified scenario with four universities. This example demonstrates how proximity preferences can alter university choices when quality and distance vary together. We examine two cases: one where the student has no preference for proximity ($\beta_i = 0$), and another where this preference is strong ($\beta_i = 1$).

Consider four universities A , B , C , and D , each characterized by quality (expected monthly salary in thousands) and distance (in hundreds of kilometers): A has quality $q_A = 10$ and distance $d_A = 8$; B has $q_B = 8$ and $d_B = 4$; C has $q_C = 6$ and $d_C = 2$; D has $q_D = 4$ and $d_D = 2$.

Suppose the student can include $K = 2$ universities on their ROL. With no preference for proximity ($\beta_i = 0$), utility depends solely on quality:

$$U_{iA} = 10, \quad U_{iB} = 8, \quad U_{iC} = 6, \quad U_{iD} = 4.$$

The student chooses universities A and B , yielding average quality of 9.

With strong preference for proximity ($\beta_i = 1$), utility becomes:

$$U_{iA} = 10 - 8 = 2, \quad U_{iB} = 8 - 4 = 4, \quad U_{iC} = 6 - 2 = 4, \quad U_{iD} = 4 - 2 = 2.$$

The student now chooses universities B and C , yielding average quality of 7. Stronger proximity preferences thus lead to lower-quality university choices.

C.2 Additional figures and tables

Table C1: Gender gap in expected salaries

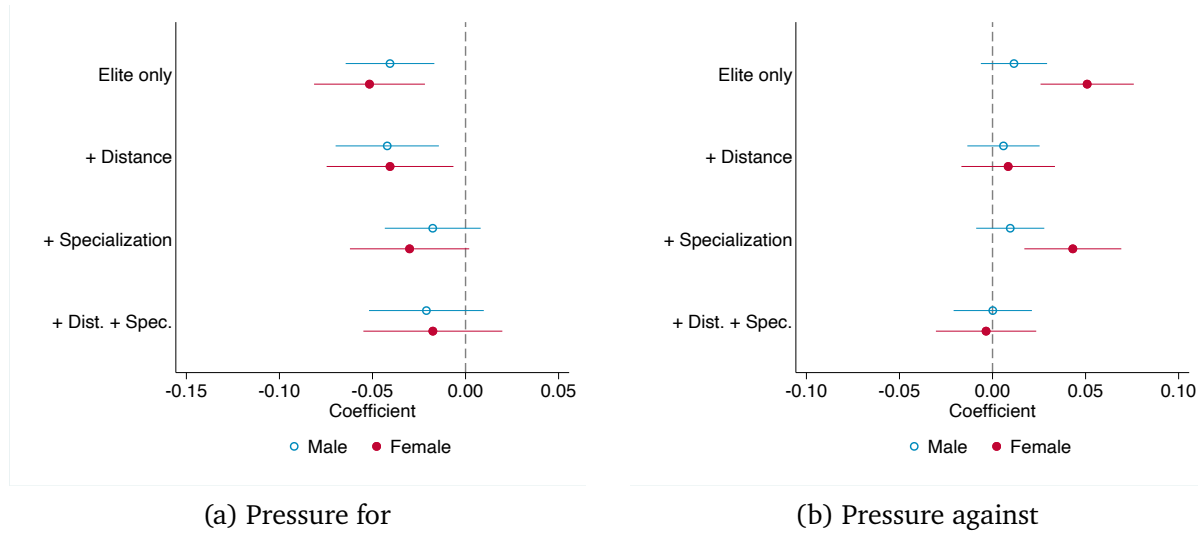
Dependent variable	One-year salary		Five-year salary	
	(1)	(2)	(3)	(4)
Female	−0.072*** (0.022)	−0.023 (0.022)	−0.073*** (0.022)	−0.033 (0.022)
Elite university		0.534*** (0.071)		0.530*** (0.069)
Average distance		0.027*** (0.010)		0.030*** (0.009)
STEM-focused university		0.111*** (0.025)		0.117*** (0.024)
Education-focused university		−0.115*** (0.024)		−0.029 (0.024)
Gaokao ranking FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
Demographic controls	✓	✓	✓	✓
Observations	5,378	5,378	5,378	5,378
R-squared	0.437	0.458	0.453	0.472

Notes: This table reports regression results examining how university characteristics explain the gender gap in expected graduate earnings. The dependent variable is standardized average salary of graduates from the admitted university. One-year salaries are average salaries of 2017 graduates measured in 2018; five-year salaries are average salaries of 2013 graduates measured in 2018. “Female” is an indicator for female students. “Average distance” is standardized average driving distance from home to universities on a student’s ROL. Gaokao ranking FE include indicators for each percentile of the Gaokao ranking and high-school track. Year FE are fixed effects for year of Gaokao. Demographic controls include age, residence status, poverty status, and repeat exam-taker status. Standard errors clustered at the student level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

D Additional results on what explains gendered patterns in university choices

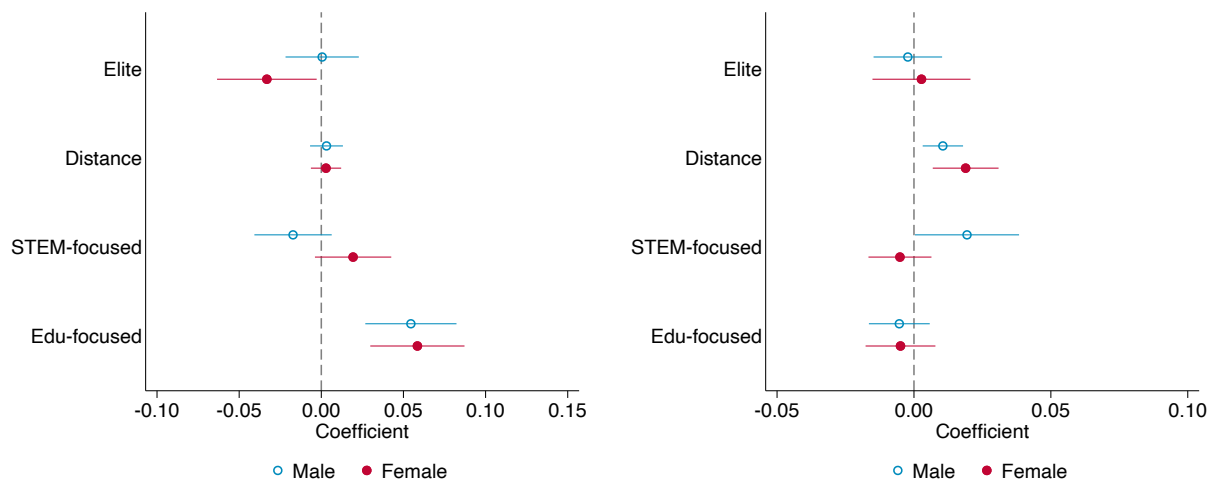
D.1 Additional results on parental pressure and university characteristics

Figure D1: Gendered pressure against elite universities: the role of distance and specialization



Notes: This figure reports the coefficients of elite status on the probability of pressure for (Panel a) and pressure against (Panel b) from a series of linear probability models, separately by gender. Each row adds controls progressively: elite status only, adding distance, adding specialization, and adding both distance and specialization. Parental pressure classifies each student-university pair based on $\text{Pressure Gap}_{ij} = \text{Parental Willingness}_{ij} - \text{Own Willingness}_{ij}$: pressure for ($\text{Gap} > 1$), agree ($|\text{Gap}| \leq 1$), and pressure against ($\text{Gap} < -1$). All specifications include individual fixed effects. Error bars represent 95% confidence intervals based on standard errors clustered at the student level.

Figure D2: How university characteristics predict parental pressure (threshold ± 2)

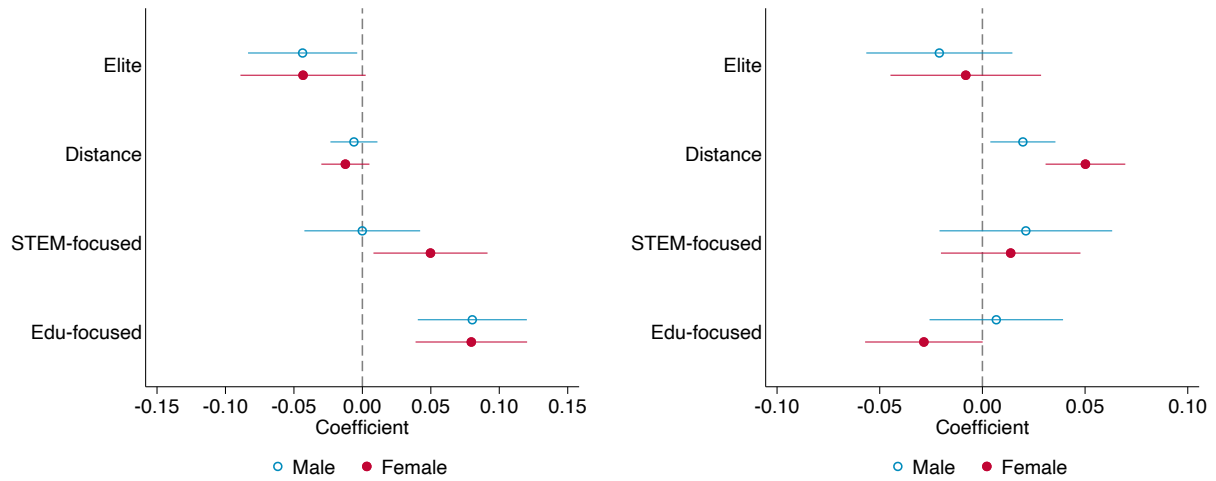


(a) Pressure for

(b) Pressure against

Notes: This figure replicates Figure 7 using a threshold of ± 2 to classify parental pressure: pressure for ($\text{Gap} > 2$), agree ($|\text{Gap}| \leq 2$), and pressure against ($\text{Gap} < -2$). See Figure 7 for the rest of the details.

Figure D3: How university characteristics predict parental pressure (threshold 0)

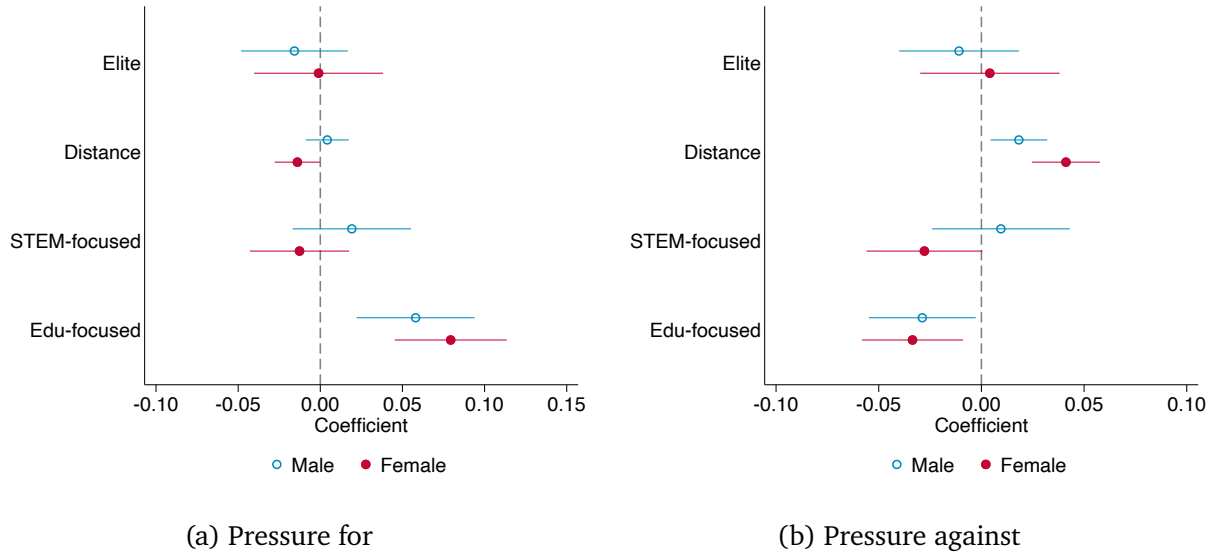


(a) Pressure for

(b) Pressure against

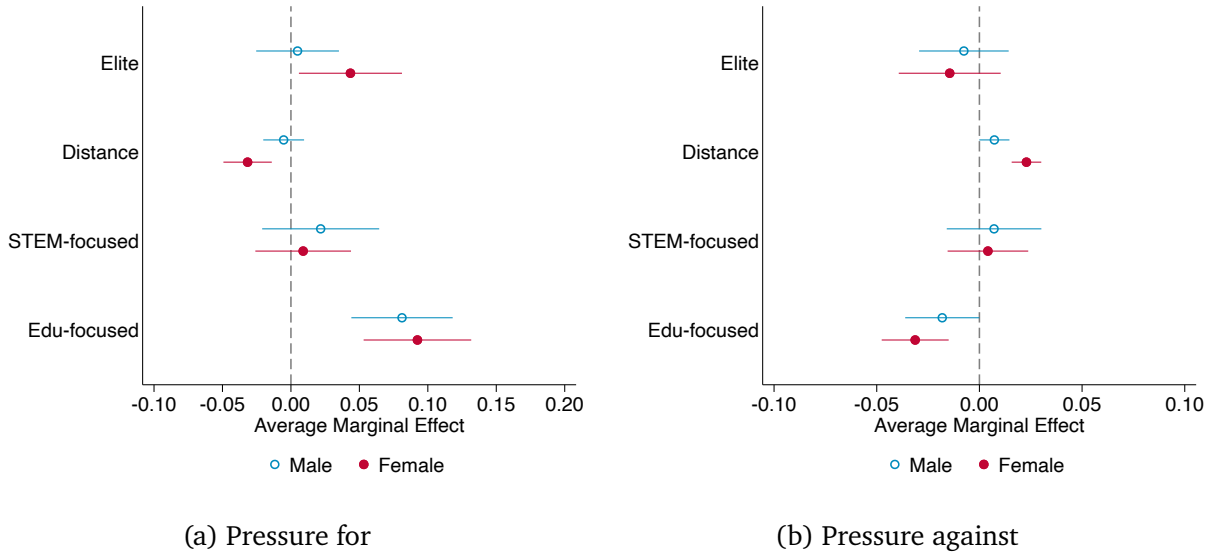
Notes: This figure replicates Figure 7 using a threshold of 0, so that any positive gap is classified as pressure for ($\text{Gap} > 0$), exact agreement as agree ($\text{Gap} = 0$), and any negative gap as pressure against ($\text{Gap} < 0$). See Figure 7 for the rest of the details.

Figure D4: How university characteristics predict parental pressure (demeaned)



Notes: This figure replicates Figure 7 using individually demeaned willingness measures. For each student i , we subtract the student's mean own willingness and mean parental willingness across all universities before computing the pressure gap: $\text{Demeaned Gap}_{ij} = (\text{Parental Willingness}_{ij} - \overline{\text{Parental Willingness}_i}) - (\text{Own Willingness}_{ij} - \overline{\text{Own Willingness}_i})$. Categories are defined as pressure for ($\text{Gap} > 1$), agree ($|\text{Gap}| \leq 1$), and pressure against ($\text{Gap} < -1$). See Figure 7 for the rest of the details.

Figure D5: How university characteristics predict parental pressure (multinomial logit)



Notes: This figure replicates Figure 7 using a multinomial logit model instead of a linear probability model. The dependent variable is a three-category parental pressure measure: “Pressure For”, “Agree”, and “Pressure Against”. Pressure categories are based on $\text{Pressure Gap}_{ij} = \text{Parental Willingness}_{ij} - \text{Own Willingness}_{ij}$: pressure for ($\text{Gap} > 1$), agree ($|\text{Gap}| \leq 1$), and pressure against ($\text{Gap} < -1$). The plotted coefficients are average marginal effects on the probability of pressure for (Panel a) or pressure against (Panel b), estimated separately for male and female students. See Figure 7 for the rest of the details.

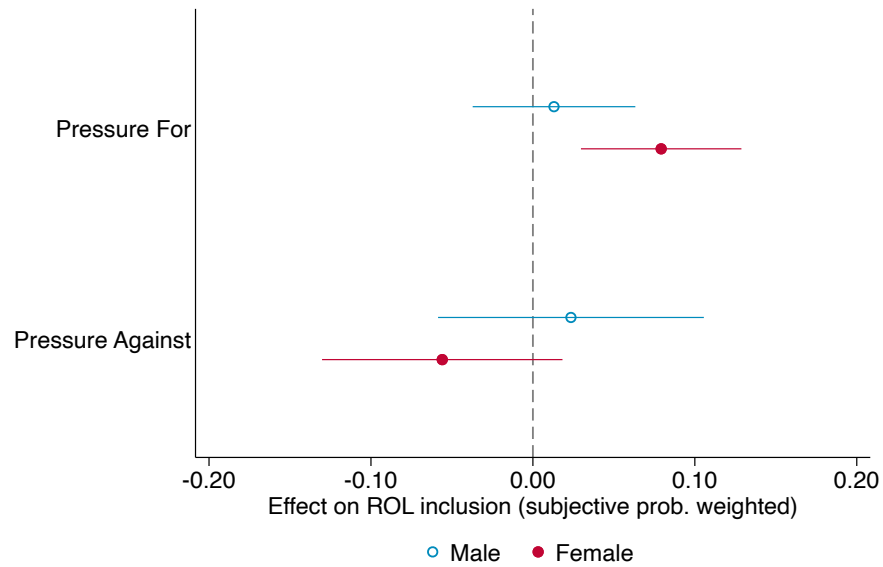
Table D1: Parental pressure and university characteristics

Dependent variable	Pressure for			Pressure against		
	Male (1)	Female (2)	Difference (3)	Male (4)	Female (5)	Difference (6)
Elite	−0.021 (0.016)	−0.018 (0.019)	0.003 (0.025)	0.000 (0.011)	−0.003 (0.014)	−0.004 (0.017)
Distance (std.)	0.003 (0.007)	−0.011* (0.007)	−0.014 (0.010)	0.008 (0.005)	0.036*** (0.008)	0.028*** (0.009)
STEM-focused	0.020 (0.019)	0.015 (0.016)	−0.005 (0.025)	0.010 (0.014)	−0.004 (0.010)	−0.013 (0.017)
Edu-focused	0.074*** (0.017)	0.075*** (0.018)	0.001 (0.025)	−0.008 (0.010)	−0.031*** (0.010)	−0.023* (0.014)
Individual FE	✓	✓	✓	✓	✓	✓
Observations	3,776	3,764	7,540	3,398	3,407	6,805
R-squared	0.398	0.345	0.371	0.285	0.332	0.311

Notes: This table reports coefficients from linear regressions of university characteristics on the probability of being classified as pressure for (columns 1–3) or pressure against (columns 4–6). Columns (1) and (4) are estimated on male students; columns (2) and (5) on female students; columns (3) and (6) report the gender difference from a pooled regression with gender interactions. Pressure categories are based on $\text{Pressure Gap}_{ij} = \text{Parental Willingness}_{ij} - \text{Own Willingness}_{ij}$: pressure for ($\text{Gap} > 1$), agree ($|\text{Gap}| \leq 1$), and pressure against ($\text{Gap} < -1$). The dependent variable equals one for the respective pressure category and zero for agree. Distance is standardized to have mean zero and standard deviation one. Additional controls include university specialization indicators for agriculture, language, and finance, number of actual admissions, and city tier. All specifications include individual fixed effects. Standard errors clustered at the student level are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

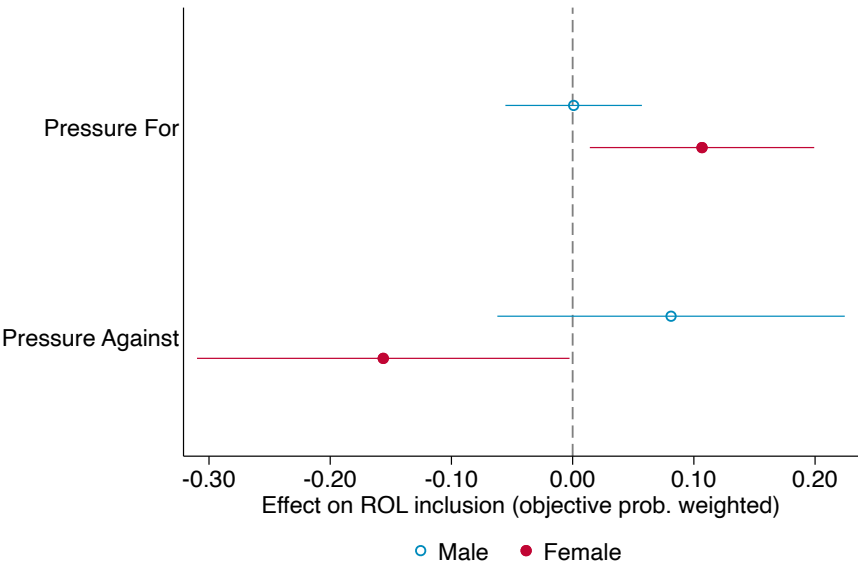
D.2 Additional results on parental pressure and inclusion decisions

Figure D6: Effect of parental pressure on ROL inclusion (subjective probability weighted)



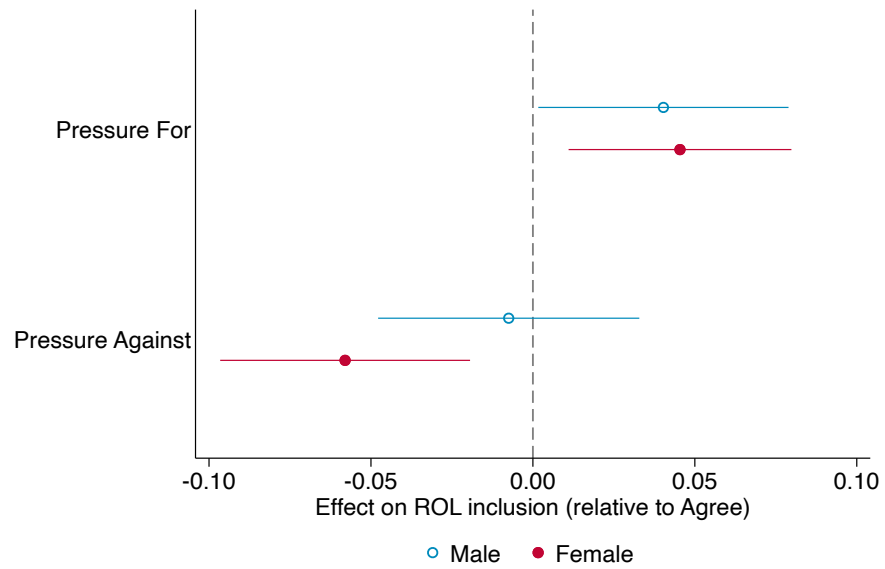
Notes: This figure replicates Figure 8 weighting each student–university observation by the student’s self-reported probability of admission to that university. Observations with higher perceived admission chances receive greater weight. See Figure 8 for the rest of the details.

Figure D7: Effect of parental pressure on ROL inclusion (objective probability weighted)



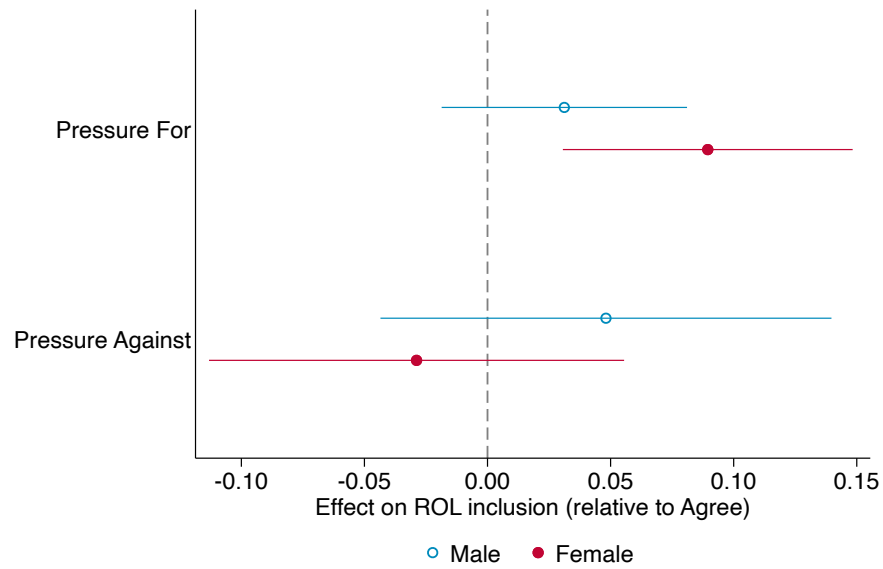
Notes: This figure replicates Figure 8 weighting each student–university observation by the objective admission probability, as described in Appendix B.1. See Figure 8 for the rest of the details.

Figure D8: Effect of parental pressure on ROL inclusion (threshold 0)



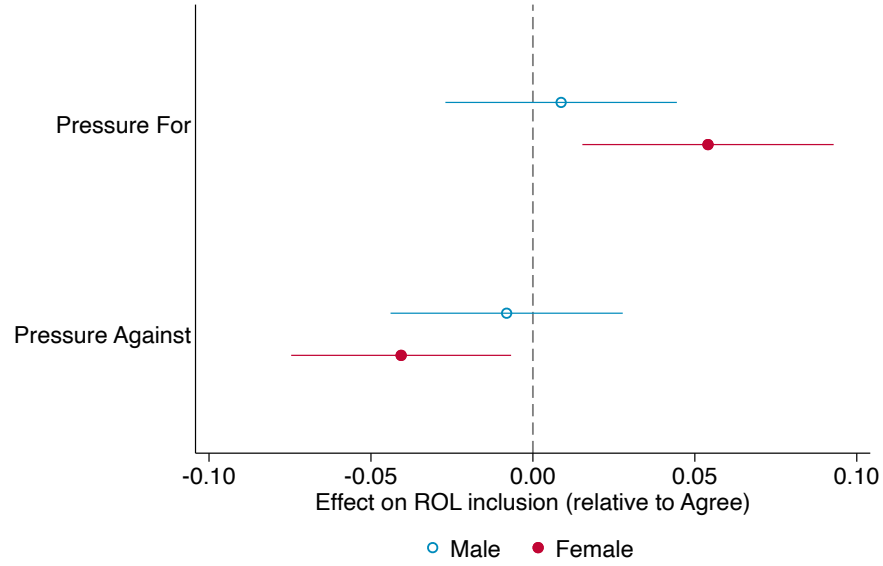
Notes: This figure replicates Figure 8 using a threshold of 0 to classify parental pressure: pressure for ($\text{Gap} > 0$), agree ($\text{Gap} = 0$), and pressure against ($\text{Gap} < 0$). See Figure 8 for the rest of the details.

Figure D9: Effect of parental pressure on ROL inclusion (threshold ± 2)



Notes: This figure replicates Figure 8 using a threshold of ± 2 to classify parental pressure: pressure for ($\text{Gap} > 2$), agree ($|\text{Gap}| \leq 2$), and pressure against ($\text{Gap} < -2$). See Figure 8 for the rest of the details.

Figure D10: Effect of parental pressure on ROL inclusion (demeaned pressure gap)



Notes: This figure replicates Figure 8 using a demeaned pressure gap. For each student i , we compute $\text{Demeaned Gap}_{ij} = (\text{Parental Willingness}_{ij} - \text{Parental Willingness}_i) - (\text{Own Willingness}_{ij} - \text{Own Willingness}_i)$. Pressure categories are defined using the ± 1 threshold applied to the demeaned gap. See Figure 8 for the rest of the details.

Table D2: Differential responsiveness to parental pressure

Dependent variable	Male (1)	Female (2)	All (3)
Pressure for	0.024 (0.022)	0.075*** (0.021)	0.024 (0.022)
Pressure against	0.010 (0.033)	−0.063** (0.029)	0.010 (0.033)
Pressure for × Female			0.051* (0.030)
Pressure against × Female			−0.072 (0.044)
Own willingness (std.)	0.056*** (0.008)	0.061*** (0.008)	0.056*** (0.008)
Own willingness (std.) × Female			0.005 (0.011)
Individual FE	✓	✓	✓
Observations	3,938	3,955	7,893
R-squared	0.175	0.195	0.187

Notes: This table reports coefficients from a linear probability model where the dependent variable is whether student i included university j in their rank-order list. The key explanatory variables are indicators for pressure for and pressure against, with agree as the reference category. Pressure categories are based on $\text{Pressure Gap}_{ij} = \text{Parental Willingness}_{ij} - \text{Own Willingness}_{ij}$: pressure for ($\text{Gap} > 1$), agree ($|\text{Gap}| \leq 1$), and pressure against ($\text{Gap} < -1$). All specifications include individual fixed effects. Standard errors in parentheses are clustered at the student level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

D.3 Details on alternative explanations

This appendix presents detailed results on alternative explanations for gendered patterns in university choices.

Preferences for majors. If women place higher value on majors than men, they might sacrifice university quality to pursue a preferred major. This trade-off arises because the Chinese admission system ranks admitted students by Gaokao scores and assigns high-ranked students to their preferred majors first, while low-ranked students may be assigned to less popular majors (Hu et al., 2024). For this channel to explain our results, women would need to care more about majors than men.

To investigate, we use the frequency of choosing “Accept any major” on a student’s ROL as a measure of preference for university quality versus major. “Accept any major” is an option associated with each listed university. If a student selects this option, they give the university permission to assign them to any major outside their five listed preferences, increasing their admission chances at the cost of potentially receiving a less preferred major. We examine both the share of students choosing “Accept any major” for their first choice and across all choices (Table D3). More than 90% of students choose “Accept any major” for their first choice, and the share is slightly higher for women than men. Additionally, students choose “Accept any major” for more than 80% of universities on their ROLs, again with a higher share among women. This suggests that women are more likely to prioritize university quality over majors than men.

Table D3: Frequency of choosing “Accept any major” by gender

	Female (1)	Male (2)	Difference (3)
Share in first choice	0.951 (0.217)	0.934 (0.248)	0.016*** (0.005)
Share across all choices	0.852 (0.355)	0.816 (0.387)	0.035*** (0.009)
Observations	3,697	3,787	

Notes: This table reports the frequency of selecting “Accept any major” on students’ ROLs. This option allows the university to assign the student to any major if they are not admitted to their five listed majors, increasing admission chances at the cost of potentially receiving a less preferred major. The first row shows the share selecting this option for the first choice; the second row shows the share across all choices. Standard deviations in parentheses for columns (1)–(2); standard errors in parentheses for column (3). *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

As additional evidence, we asked students a multiple-price-list question with two options:

(1) a higher-quality university (as measured by admission cutoff) with a randomly assigned major, or (2) a lower-quality university with autonomy to choose any major. We gradually raised the quality of the first option to elicit the switching point. This measures willingness to sacrifice university quality to secure a preferred major; a larger quality gap at switching indicates stronger major preferences. We find that women and men switch at the same point on average, indicating that women do not care more about majors relative to university quality. We thus find no support for the explanation that women sacrifice university quality to secure preferred majors.

Competitiveness. Male and female students might differ in their attitudes toward competition, and a distaste for competition could lead female students to avoid the more competitive elite universities. To examine this explanation, we follow the experimental design employed by Niederle and Vesterlund (2007) and elicited students' competitiveness in the survey. The students first worked on a Raven's progressive matrices test for a piece rate. Subsequently, we provided them with (hypothetical) opportunities to make selections between piece-rate payments and competitive tournaments. We also elicited their levels of confidence (both overestimation and overplacement) and their risk attitude. Competitiveness is measured by the choice between the piece rate and the tournament after controlling for risk attitude and confidence. The regression results in Appendix Table D4 show that, after controlling for confidence and risk preferences, there is no significant difference between women and men in terms of selecting the more competitive regime. This result suggests no gender difference in competitiveness in our sample, which is consistent with previous findings for China (Zhang, 2019).

Beliefs about returns to different types of universities. Our findings indicate that female students are less likely to apply to elite, distant, and STEM universities than male students of the same qualification. One possible explanation is that female students are misinformed about the returns to these types of universities, which makes them unaware of the economic consequences of their choices.

To address this possible explanation, we elicited students' perceived salaries of graduates from each of the eight universities presented to them. To enhance the accuracy of their estimations, we provided a reference point of the average starting salary of graduates from a well-known university. We further incentivized accurate estimates by offering an extra monetary reward for answers within 100 Chinese yuan of the true monthly salary. Lastly, we vary the type of the university along elite status, distance, and specialization to understand their perceived returns to different university characteristics.

Table D5 Column (1) reports gender differences in perceived expected salaries for universities of different characteristics. Both male and female students recognize the substantial returns to elite education, reporting significantly higher expected salaries for elite than for non-elite uni-

Table D4: Gender differences in competitiveness

Dependent variable	Tournament (1)	Tournament (2)
Female	-0.035 (0.032)	-0.032 (0.034)
Overestimation	0.004 (0.009)	0.003 (0.009)
Overplacement	-0.006 (0.016)	-0.007 (0.016)
Risk preference	✓	✓
Demographic controls		✓
Observations	1,006	1,006
R-squared	0.134	0.141

Notes: This table reports the regression results for the gender difference in competitiveness. The dependent variable “tournament” is equal to 1 when a student chooses the more competitive payment option of a tournament over a piece rate in a hypothetical question in the survey. The regression controls for confidence in the form of “Overestimation” (the difference between the predicted and actual number of correct answers in a quiz of Raven progressive matrices) and “Overplacement” (the difference between actual and predicted rankings in this quiz), and the risk preference. Standard errors, reported in parentheses, are clustered at the individual level. Demographic controls include age, poverty status, residence status, and repeat exam taker status. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

versities. However, there is no gender difference in the perceived return to elite education. The same holds true for the other two dimensions: men and women have similar beliefs regarding the returns to distant versus proximate universities, and similar beliefs regarding STEM universities versus universities with other specializations.

Additionally, it is possible that women expect their own return to attending an elite or STEM university to be low, even if they believe the average return is high (Kiessling et al., 2024). To explore this possibility, we asked students to assess their future salaries after graduation if they were to attend each of the eight universities. Column (2) of Table D5 presents the results on regressing students’ estimated own future salaries on university characteristics by gender. Similar to the estimates of average salaries for past graduates, female students are not significantly more pessimistic about the returns to elite, distant or STEM universities for themselves. Thus, the gendered pattern in university applications is not driven by gender differences in perceived

returns to different types of universities.

Beliefs about ability. University admission outcomes are solely determined by applicants' Gaokao scores. This raises the question of whether Gaokao scores overstate the true ability of female students relative to male students. If female students have lower abilities than their male counterparts with the same Gaokao scores and are aware of it, then opting for non-elite universities could be a sensible move to avoid poor university performance and subsequent career consequences.

However, the evidence suggests the opposite regarding the abilities of male and female students of the same Gaokao scores. In fact, Gaokao scores might understate women's true ability compared to that of men. Using scores in four mock exams, we can compare the average mock exam performance of male and female students with identical Gaokao scores. Typically, Chinese high schools administer four mock exams in the last six months before the Gaokao. The mock exams are designed by the official educational authorities to mimic the Gaokao in both schedule and format, thereby familiarizing students with the Gaokao procedures. We find that, on average, women perform better in these mock exams than male students of the same Gaokao scores (see Table D6), which is consistent with previous findings on gender differences in the Gaokao (Cai et al., 2019) and in other high-stakes testing environments (Azmat et al., 2016; Iriberry and Rey-Biel, 2019; Schlosser et al., 2019; Arenas and Calsamiglia, 2025; Arias et al., 2023; Akyol et al., 2024). Assuming that the average score of four mock exams represents a more accurate measure of ability due to reduced measurement error, it implies that women have higher abilities than men with the same Gaokao score.

Confidence regarding academic performance at university. We next test whether female students are less confident about their future academic performance than equally qualified male students at elite and STEM-focused universities, and whether this motivates them to choose less elite and less STEM-focused universities in their ROLs. We asked students to estimate where they would rank in a class of 100 students if they were admitted to each of the eight universities. On average, students estimate their performance to be slightly better than the median (44 out of 100). We do not observe a gender difference in confidence (44 for males versus 45 for females). We further regress the estimated class rank on university characteristics and interact them with Female; the results are shown in Column (1) of Table D7. The results indicate that female students are not less confident about their performance at elite, STEM-focused, or distant universities. Thus, gender differences in confidence cannot account for the gendered pattern of university choices.

Confidence in admission chances. Female students may be less likely to apply to elite, STEM-focused, or distant universities if they underestimate their chances of admission, despite having chances that are objectively equal to those of males with the same Gaokao score

(Hakimov et al., 2023). To test this alternative explanation, we asked students to estimate the unconditional probability of their score exceeding the admission cutoff at each of the eight universities. On average, male students in our survey estimated their admission chances to the eight universities to be higher than their female counterparts (37% versus 33%), even though the two groups rank similarly in the Gaokao on average. Table D7 Column (2) presents the regression results on the gender difference in perceived admission chances. Although female students estimate a lower probability of admission than male students, this gender difference is uniform across university types: the gap for elite, STEM-focused, and distant universities is not significantly different from that for their counterparts. Thus, the gender difference in perceived admission probabilities cannot explain the observed gendered university choices.

E Survey instruments

E.1 Wave 1 (2022)

1. Your gender: [Single choice]
 - Male
 - Female
2. Household registration (hukou): [Single choice]
 - Rural
 - Non-agricultural (urban)
3. Which of the following ranges does your total Gaokao score fall into? [Single choice]
 - Above the Tier-1 cutoff
 - Below the Tier-1 cutoff but above the Tier-2 cutoff
 - Below the Tier-2 cutoff
4. Suppose there are two universities, A and B. University A's historical average minimum admission score is **10 points** higher than University B's. However, if admitted to University A, you can only be assigned to a major you did not choose, whereas admission to University B allows you to freely choose any major. If both universities would admit you, which would you prefer? [Single choice]
 - University A *[If selected, skip to Question 15]*
 - University B

Repeat the same question with the score gap increasing from 20 to 100 points in increments of 10. Students who choose University A at any point skip to Question 5.

5. How did you arrange the order of universities on your rank-order list? That is, which university did you place as your first choice, which as your second choice, and so on. [Single choice]
 - I placed the university I most want to attend as my first choice, the second most desired as my second choice, and so on; the ordering reflects my true ranking of universities.

- I placed the university most likely to admit me as my first choice, the second most likely as my second choice, and so on; the ordering reflects my probability of being admitted.
 - The universities I chose within the same tier have no difference to me; which university I placed as my first choice was random.
 - Other, please specify: _____
6. When filling out your application, how many years of the universities' historical admission scores/rankings did you reference? [Single choice]
- 5 years or more
 - 4 years
 - 3 years
 - 2 years
 - 1 year
 - I did not reference historical admission scores
7. When filling out your application, did you primarily reference universities' minimum admission scores/cutoff lines, minimum admission rankings, average admission scores, or average admission rankings? [Single choice]
- Minimum admission score
 - Minimum admission ranking
 - Average admission score
 - Average admission ranking
 - Other: _____
8. If the decision were entirely up to you, would you prefer to attend an in-province university, an out-of-province university, or do you have no preference? [Single choice]
- In-province university
 - Out-of-province university
 - No preference
9. From your parents' perspective, would they prefer you to attend an in-province university, an out-of-province university, or do they have no preference? [Single choice]

- In-province university
 - Out-of-province university
 - No preference
10. On a scale of 0 to 10, how much do you agree with the following statements? (0 = completely disagree, 10 = completely agree) [Scale: 0–10 for each]
- “Boys should go out and venture into the world.”
 - “Girls should not go to places too far from home.”
 - “Boys and girls have innate biological differences in mathematical ability.”
 - “Boys are better at science subjects, and girls are better at humanities.”
11. Suppose there are two universities, A and B. Consider two application plans. Plan 1: First choice is University A, second choice is University B. Plan 2: First choice is University B, second choice is University A. If your score does not reach University A’s admission cutoff, is your probability of being admitted to University B higher under Plan 1, higher under Plan 2, or the same? [Single choice]
- Higher under Plan 1
 - Higher under Plan 2
 - The probability is the same under both plans
12. If a student’s score exceeds their first-choice university’s cutoff, but all the majors they listed are full, and the student selected “do not accept major reassignment”—is the probability of the student being admitted to their second-choice university (within the same tier) equal to zero? [Single choice]
- Yes, the probability of being admitted to the second choice is zero
 - No, second-choice admission is unaffected
13. Consider the following two plans. Plan 1: First choice is University A, second choice is blank. Plan 2: First choice is University A, second choice is University B. Which plan gives a higher probability of admission to University A? [Single choice]
- Plan 1
 - Plan 2
 - The probability is the same under both plans

14. Are you generally more inclined to take risks or to avoid risks? Please answer with a number from 0 to 10, where 0 means “I try hard to avoid all risks” and 10 means “I am always ready to take risks.” [Scale: 0–10]
15. What range does your family’s annual household income fall into? [Single choice]
- Below 10,000 yuan
 - 10,000–25,000 yuan
 - 25,000–50,000 yuan
 - 50,000–75,000 yuan
 - 75,000–100,000 yuan
 - 100,000–150,000 yuan
 - 150,000–200,000 yuan
 - 200,000–300,000 yuan
 - 300,000–500,000 yuan
 - Above 500,000 yuan
16. Father’s education level: [Single choice]
- Primary school
 - Middle school
 - High school / technical school / vocational high school
 - Associate degree (*dazhuan*)
 - Bachelor’s degree
 - Master’s degree
 - Doctoral degree
 - Never attended school
17. Mother’s education level: [Single choice]
- [Same options as father’s education level.]*
18. Father’s occupation: [Single choice]
- Farming
 - Self-employed

- Migrant worker
- Public institution employee
- Government employee
- State-owned enterprise employee
- Private enterprise employee
- Other: _____

19. Mother's occupation: [Single choice]

[Same options as father's occupation.]

20. From kindergarten through high school, who had more say in decisions about your education (such as which middle school or high school to attend)? [Single choice]

- Father decided alone
- Mother decided alone
- Mainly father
- Mainly mother
- Parents decided together
- Others decided

21. From kindergarten through high school, who was primarily responsible for supervising and disciplining you? [Single choice]

- Entirely father
- Entirely mother
- Mainly father
- Mainly mother
- Parents equally
- Others

22. Overall, are household decisions mainly made by your father or your mother? [Single choice]

- Mainly father
- Mainly mother

- Both equally
- Other family members

23. Consider the following situation. You can choose between a lottery and a fixed payment.

Option 1: A lottery with a 50% chance of winning 300 yuan and a 50% chance of winning 0 yuan.

Option 2: A guaranteed payment of [X] yuan.

Which do you prefer? [Single choice]

- The 50/50 lottery
- The fixed payment

The guaranteed amount X was determined through an adaptive bisection procedure. Starting at 160 yuan, the amount was adjusted upward or downward based on each response, converging on the student's certainty equivalent over approximately 5 rounds. Possible amounts ranged from 10 to 310 yuan in increments of 10 yuan.

24. Are you an only child? [Single choice]

- Yes *[If selected, skip to end]*
- No

25. *[Shown only if not an only child.]* Please indicate the number of older brothers, older sisters, younger brothers, and younger sisters you have. [Matrix: 0–4 for each]

- Older brothers: 0 / 1 / 2 / 3 / 4
- Older sisters: 0 / 1 / 2 / 3 / 4
- Younger brothers: 0 / 1 / 2 / 3 / 4
- Younger sisters: 0 / 1 / 2 / 3 / 4

E.2 Wave 2 (2023)

1. Your gender: [Single choice]

- Male
- Female

2. Household registration (hukou): [Single choice]

- Rural
 - Non-agricultural (urban)
3. Suppose there are two universities, A and B. University A's historical average minimum admission score is **10 points** higher than University B's. However, if admitted to University A, you can only be assigned to a major you did not choose, whereas admission to University B allows you to freely choose any major. If both universities would admit you, which would you prefer? [Single choice]
- University A *[If selected, skip to Question 14]*
 - University B

Repeat the same question with the score gap increasing from 20 to 100 points in increments of 10. Students who choose University A at any point skip to Question 5.

4. How did you arrange the order of universities on your rank-order list? [Single choice]
[Same options as Wave 1.]
5. When filling out your application, how many years of the universities' historical admission scores/rankings did you reference? [Single choice]
[Same options as Wave 1.]
6. When filling out your application, did you primarily reference universities' minimum admission scores, minimum admission rankings, average admission scores, or average admission rankings? [Single choice]
[Same options as Wave 1.]
7. Your Gaokao score (including bonus points) range: [Single choice]
[Separate score brackets were shown for STEM-track and non-STEM-track students, corresponding to admission tier cutoffs for 2023. STEM-track students: 10 brackets from "589 and above" to "408 and below." Non-STEM-track students: 9 brackets from "579 and above" to "464 and below." The score bracket determined which set of eight universities the student evaluated in Questions a–g.]

Students were presented with eight universities selected to vary in elite status, specialization, and distance from home, and to be among the most popular choices for students in their score range and high-school track. All students within the same score bracket and track evaluated the same eight universities. For each university, students answered the

following eight questions. Here, [University] refers to the specific university name shown to the student. As a salary reference, students were informed that the average monthly starting salary of 2017 graduates from North China University of Water Resources and Electric Power was 4,050 yuan.

- (a) How much do you know about [University]? Please answer with a number from 0 to 10, where 0 means “not familiar at all” and 10 means “very familiar.” [Scale: 0–10]
- (b) If you were to apply to [University], what do you think is the probability that your score exceeds its admission cutoff? Please enter a number from 0 to 100. [Numeric: 0–100]
- (c) If you were admitted to [University], where do you think you would rank among 100 students in your cohort in terms of academic performance over four years of college? Please enter a number from 1 to 100, where 1 means the best. [Numeric: 1–100]
- (d) Please estimate: In 2018, what was the average monthly salary of 2017 bachelor’s degree graduates from [University]? Please enter a number between 3,000 and 10,000. [Numeric: 3,000–10,000]
[Incentivized: students received a bonus payment of 10 yuan if their estimate was within 100 yuan of the true value.]
- (e) Please estimate: If you were admitted to [University] and found employment after graduation, what monthly salary could you earn? Please enter a number between 2,000 and 30,000. [Numeric: 2,000–30,000]
- (f) If you were admitted to [University], how willing would you be to attend? Please answer with a number from 0 to 10, where 0 means “definitely would not want to attend” and 10 means “definitely would want to attend.” [Scale: 0–10]
- (g) If you were admitted to [University], how much would your parents want you to attend? Please answer with a number from 0 to 10, where 0 means “not at all” and 10 means “very much.” [Scale: 0–10]

[Questions a–g were repeated for each of the eight universities presented to the student.]

8. When filling out your application, whom did you consult? [Multiple choice]

- Parents
- Teachers
- Relatives

- Classmates
- Parents' friends
- Previous graduates
- Other: _____
- I did not consult anyone

9. Among those you consulted, who played the most important role in deciding which universities to include on your list? [Single choice]

[Same options as above, excluding "Other."]

10. Overall, whose opinion dominated the final choice of universities and majors? [Single choice]

- Entirely my own decision
- Mainly my own opinion
- Mainly my father's opinion
- Mainly my mother's opinion
- Parents decided together
- Other elders' opinion
- Teacher's opinion
- Other

11. If the decision were entirely up to you, would you prefer to attend an in-province university, an out-of-province university, or do you have no preference? [Single choice]

- In-province university
- Out-of-province university
- No preference

12. From your parents' perspective, would they prefer you to attend an in-province university, an out-of-province university, or do they have no preference? [Single choice]

- In-province university
- Out-of-province university
- No preference

13. If the decision were entirely up to you, what type of university would you prefer? [Single choice]

- Comprehensive
- Normal (teacher training)
- Science and technology
- Finance and economics
- Agriculture and forestry
- Other: _____
- No preference

14. *[Shown only if a specific type was chosen.]* Please briefly explain why you prefer this type of university. [Open-ended]

15. From your parents' perspective, what type of university would they prefer you to attend? [Single choice]

[Same options as Question 13.]

16. On a scale of 0 to 10, how much do you agree with the following statements? (0 = completely disagree, 10 = completely agree) [Scale: 0–10 for each]

- “Boys should go out and venture into the world.”
- “Girls should not go to places too far from home.”
- “Boys and girls have innate biological differences in mathematical ability.”
- “Boys are better at science subjects, and girls are better at humanities.”

17. On a scale of 0 to 10, how much do your **parents** agree with the following statements? (0 = completely disagree, 10 = completely agree) [Scale: 0–10 for each]

- “Boys should go out and venture into the world.”
- “Girls should not go to places too far from home.”
- “Boys and girls have innate biological differences in mathematical ability.”
- “Boys are better at science subjects, and girls are better at humanities.”

18. Suppose there are two universities, A and B. Consider two application plans. Plan 1: First choice is University A, second choice is University B. Plan 2: First choice is University B,

second choice is University A. If your score does not reach University A's admission cutoff, is your probability of being admitted to University B higher under Plan 1, higher under Plan 2, or the same? [Single choice]

- Higher under Plan 1
- Higher under Plan 2
- The probability is the same under both plans

19. If a student's score exceeds their first-choice university's cutoff, but all the majors they listed are full, and the student selected "do not accept major reassignment"—is the probability of the student being admitted to their second-choice university (within the same tier) equal to zero? [Single choice]

- Yes, the probability of being admitted to the second choice is zero
- No, second-choice admission is unaffected

20. Consider the following two plans. Plan 1: First choice is University A, second choice is blank. Plan 2: First choice is University A, second choice is University B. Which plan gives a higher probability of admission to University A? [Single choice]

- Plan 1
- Plan 2
- The probability is the same under both plans

21. Are you generally more inclined to take risks or to avoid risks? Please answer with a number from 0 to 10, where 0 means "I try hard to avoid all risks" and 10 means "I am always ready to take risks." [Scale: 0–10]

22. What range does your family's annual household income fall into? [Single choice]

[Same 10 income brackets as Wave 1.]

23. Father's education level: [Single choice]

[Same 8 education categories as Wave 1.]

24. Mother's education level: [Single choice]

[Same options.]

25. Father's occupation: [Single choice]

[Same 8 occupation categories as Wave 1.]

26. Mother's occupation: [Single choice]

[Same options.]

27. From kindergarten through high school, who had more say in decisions about your education? [Single choice]

[Same 6 options as Wave 1.]

28. From kindergarten through high school, who was primarily responsible for supervising and disciplining you? [Single choice]

[Same options as Wave 1.]

29. Overall, are household decisions mainly made by your father or your mother? [Single choice]

[Same options as Wave 1.]

30. Are you an only child? [Single choice]

- Yes
- No

31. *[Shown only if not an only child.]* Please indicate the number of older brothers, older sisters, younger brothers, and younger sisters you have. [Matrix: 0–4 for each]

32. **Raven's Progressive Matrices.** Please complete the following 9 test questions. You will earn 0.5 yuan for each correct answer. [Single choice for each of 9 items]

[Nine visual pattern-matching items from Raven's Progressive Matrices were presented. Each item displayed an incomplete pattern and asked students to select the missing piece from 6 or 8 options.]

33. How many of the 9 test questions do you think you answered correctly? [Single choice: 0–9]

[Incentivized: 1 yuan bonus for a correct guess.]

34. If we randomly matched you with 3 other students, where do you think your score would rank among the 4 of you? [Single choice]

- 1st (highest)
- 2nd

- 3rd
- 4th (lowest)

[Incentivized: 1 yuan bonus for a correct guess.]

35. Consider the following hypothetical scenario for a second round of 9 questions of the same type and difficulty. You would compete against 3 other students. If you score the highest, you earn 2 yuan per correct answer; otherwise, you earn nothing. How many questions do you predict you would answer correctly? [Single choice: 1–9]
36. For a third round of 9 new questions, please choose your payment scheme: [Single choice]
- Scheme 1: You earn 0.5 yuan for each correct answer (piece rate).
 - Scheme 2: Your score will be compared to three other student's score from Round 2. If your score is the highest, you earn 2 yuan per correct answer; if your score is lower, you earn nothing (tournament).
37. Consider the following situation. You can choose between a lottery and a fixed payment.
- Option 1:* A lottery with a 50% chance of winning 300 yuan and a 50% chance of winning 0 yuan.
- Option 2:* A guaranteed payment of [X] yuan.
- Which do you prefer? [Single choice]
- The 50-50 lottery
 - The fixed payment

Same adaptive bisection procedure as Wave 1, Question 23. Starting at 160 yuan, converging in approximately 5 rounds over amounts ranging from 10 to 310 yuan.

E.3 Wave 3 (2024)

The 2024 survey follows the same structure as the 2023 survey (Wave 2).

Table D5: Gender differences in perceived returns to different university characteristics

Dependent variable	Predicted average graduate salary (1)	Own expected salary if attend (2)
Elite	0.528*** (0.037)	0.270*** (0.032)
Female×Elite	−0.005 (0.057)	0.014 (0.051)
STEM	−0.055 (0.037)	−0.006 (0.030)
Female×STEM	0.013 (0.048)	−0.004 (0.041)
Distance	−0.037** (0.015)	−0.023* (0.013)
Female×Distance	−0.007 (0.023)	−0.010 (0.020)
Individual FE	✓	✓
Observations	7,893	7,893
R-squared	0.614	0.739

Notes: This table reports regression results examining gender differences in students' perceived economic returns to different university characteristics based on the survey data. The unit of observation is a student-university pair from the eight-university survey module. The dependent variable in Column (1) is the student's standardized estimate of the average monthly salary of graduates from the presented university, while the dependent variable in Column (2) is the student's standardized estimate of their own future monthly salary if they were to attend that university. Key independent variables are binary indicators for elite university status and STEM specialization, and the standardized driving distance from home, along with their interaction terms with the female indicator. All specifications include individual fixed effects. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table D6: Differences in performance between mock exams and Gaokao by gender

	Female (1)	Male (2)	Difference (3)
Mock score — Gaokao score	0.021 (0.550)	-0.023 (0.513)	0.044*** (0.015)
Observations	2,605	2,634	

Notes: This table reports the differences between mock exam scores and Gaokao scores by gender, where a positive value means better performance in the mock exam. Both the mock exam scores and the Gaokao scores are standardized by year and high-school academic track. There are four mock exams each year, and “Mock score” refers to the average of the four standardized scores. This table only includes samples of students who took the Gaokao between 2018 and 2022, as mock exam scores are not available for 2023 and 2024. Column (3) shows the difference between female and male students in the difference between the two scores. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table D7: Gender differences in beliefs by university characteristics

Dependent variable	Expected class rank (standardized) (1)	Acceptance rate (perceived) (2)
Elite	0.534*** (0.037)	-0.133*** (0.012)
Female × Elite	0.046 (0.053)	0.004 (0.018)
STEM	-0.084*** (0.031)	0.016** (0.007)
Female × STEM	0.055 (0.043)	0.001 (0.011)
Distance	-0.050*** (0.014)	0.021*** (0.004)
Female × Distance	-0.023 (0.020)	0.001 (0.006)
Individual FE	✓	✓
Observations	8,047	8,047
R-squared	0.653	0.445

Notes: This table reports regression results analyzing gender differences in students' beliefs regarding academic performance and admission probabilities across different university characteristics. The unit of observation is a student-university pair from the eight-university survey module. The dependent variable in Column (1) is the student's standardized expected class ranking if admitted to the university. Column (2) represents the student's perceived unconditional probability of being admitted to the university. Key independent variables are binary indicators for elite university status and STEM specialization, and standardized driving distance from home, along with their interaction terms with the female indicator. All specifications include individual fixed effects. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.