BEYOND THE SCORES: UNDERSTANDING THE GENDER GAP IN ACCESS TO ELITE UNIVERSITIES IN CHINA

Yi Chen
ShanghaiTech

Sikun Dou
Singapore Management

Hongbin Li
Stanford

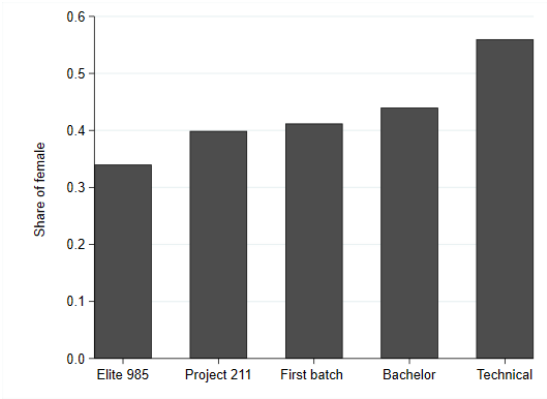
Roujing Wu
Bologna

Wednesday 28th August, 2024

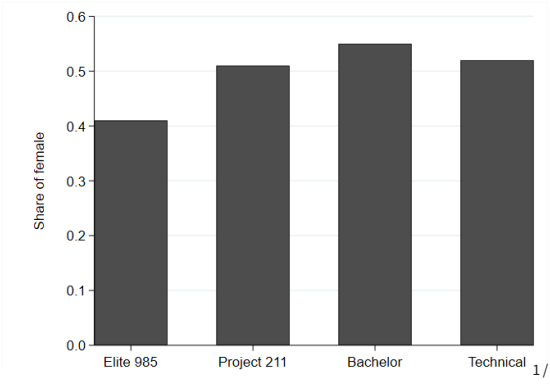
MOTIVATION

- Female share declines as the university hierarchy rises.
- Lower share in elite university over the past 20 years!

1999–2003 (our administrative data)



2022 (from Government Webstie)



RESEARCH QUESTION

- Source: Exam Score + Preference Submission → Women are 3.92% less likely to be admitted by elite university
  - Differences in score (Almost same score distributions) score → 1.72% can be explained by test scores
  - Differences in submission same score → still 2.2% less likely to enroll in elite university
- **Why don't Chinese girls choose more elite universities?**

MAIN FINDINGS

1. Women on average are 2.2% less likely enroll in elite university compared to men; on average 4.2% less likely to “downgrade”
   - Larger gender gap in the science track and in the high-scoring sample.
2. Preferences documented in literature only account for 16% of this gap
   - Major types, Major quartile, Home preference, Retake
3. Two main drivers are university types(45%) and risk attitudes (38%)

LITERATURE CONTRIBUTIONS

This paper

1. provides a more complete picture about the potential causes that lead to the gender gap in accessing elite university
2. connects the gender gap in accessing elite colleges with preferences for STEM-oriented universities
3. advances the understanding of how policies shape gender disparities in higher education
   - China's government-designated elite school list
   - pre-exam V.S. post-exam submission policy

INSTITUTION BACKGROUND

MINIMALIST KNOWLEDGE FOR NATIONAL COLLEGE ENTRANCE EXAM (GAOKAO) IN CHINA

- ONLY way get into elite universities for most students;
- Score is the SOLE criterion for admission
- Three tracks: Science, Humanity and Unified
- Students submit their priority list:
  - Usually, Only the first preference is useful
  - Due to sequential mechanism in 2000s sequential v.s. parallel
    - ★ can list Only one university in the first preference
    - ★ University process applications (only who list it as first preference) in the first round
    - ★ Second preference will be submitted after first round is finished
    - ★ Elite university usually DOES NOT have second round
  - RISKY!!! case Why China adopted it

DATA AND EMPIRICAL STRATEGY

DATA

Administrative data: All students who took Gaokao from 1999 to 2003 in China

- Demographic information; Test score; Admission outcomes

First batch cutoff score data

- Year $\times$ province $\times$ track level

Submission timing policy data

- Year $\times$ province level
- Three types: pre-exam, post-exam-pre-score, post-score Uncertainty Visualize

SAMPLE AND VARIABLES

Sample selection

- Students above first-batch cutoff score
- Only students in academic track

Key outcome variables

- **Elite university**: admitted by any Project 985 universities
- **Downgrade**: not admitted by a first-batch university in the first-preference

summary statistic

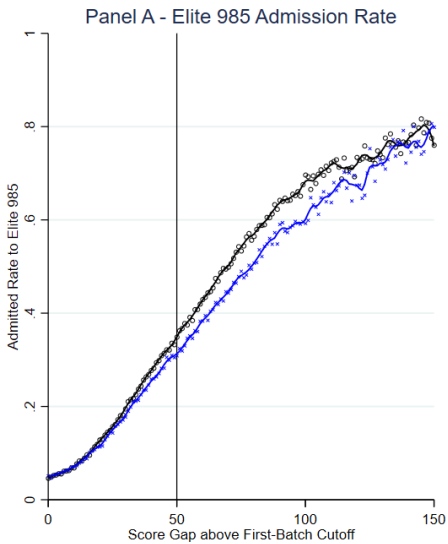
EMPIRICAL SPECIFICATION

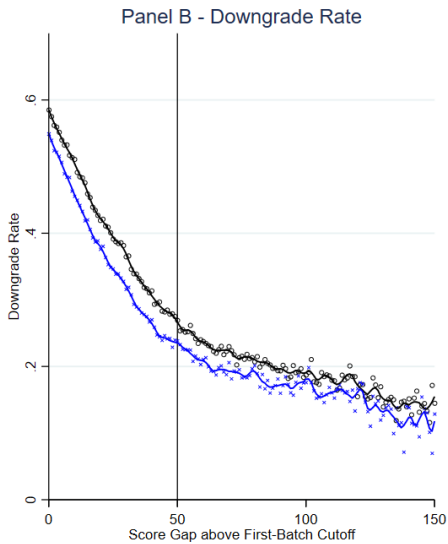
$$Y_{i,s,y,p,t} = \beta_0 + \beta_1 \text{Female} + \beta_2 \mathbf{X}_i + \lambda_{s,y,p,t} + \varepsilon_{i,s,y,p,t}$$

- i–individual; s–score; y–year; p–province; t–track
- $Y_{i,s,y,p,t}$: enroll in **elite university** or experience **downgrade**
- $\mathbf{X}_i$: individual controls (hukou, first-time test taker, ethnicity)
- Fixed effects: $\lambda_{s,y,p,t}$ (score $\times$ year $\times$ province $\times$ track)
  - Students are technically equivalent during the admission process within cell
- Standard errors clustered at province$\times$year$\times$track level.

IS THERE A GENDER DIFFERENCE?

GRAPHICAL EVIDENCE





FEMALES MAKE MORE CONSERVATIVE DECISIONS

Dependent Variable	Elite university (Project 985)				Downgrade			
Track type	Overall	Science	Humanity	Unified track	Overall	Science	Humanity	Unified track
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: All sample								
Female	-0.022***	-0.021***	-0.005***	-0.051***	-0.042***	-0.043***	-0.043***	-0.035***
	(0.003)	(0.002)	(0.002)	(0.013)	(0.003)	(0.003)	(0.003)	(0.009)
Mean of dependent variable	0.238	0.258	0.197	0.234	0.200	0.201	0.210	0.184
Observations	2,139,465	1,513,837	358,903+	266,725	2,127,964	1,504,147	357,092	266,725
R-squared	0.315	0.288	0.316	0.326	0.180	0.179	0.200	0.158
Panel B: High score sample								
Female	-0.048***	-0.045***	-0.000	-0.114***	-0.031***	-0.032***	-0.034***	-0.014
	(0.005)	(0.005)	(0.004)	(0.025)	(0.003)	(0.004)	(0.004)	(0.009)
Mean of dependent variable	0.537	0.440	0.548	0.582	0.200	0.201	0.210	0.184
Observations	550,136	434,044	57,997	58,095	547,122	431,194	57,833	58,095
R-squared	0.234	0.233	0.281	0.164	0.179	0.172	0.242	0.165
Panel C: Low score sample								
Female	-0.013***	-0.011***	-0.005***	-0.035***	-0.046***	-0.047***	-0.045***	-0.041***
	(0.002)	(0.002)	(0.002)	(0.009)	(0.003)	(0.003)	(0.004)	(0.010)
Mean of dependent variable	0.145	0.151	0.136	0.127	0.417	0.420	0.416	0.404
Observations	1,589,303	1,079,790	300,883	208,630	1,580,816	1,072,950	299,236	208,630
R-squared	0.161	0.152	0.201	0.143	0.141	0.135	0.173	0.120
Score×Year×Track×Province FE	✓	✓	✓	✓	✓	✓	✓	✓
Individual Control FE	✓	✓	✓	✓	✓	✓	✓	✓

Heterogeneity | Robustness checks

FEMALES MAKE MORE CONSERVATIVE DECISIONS

Dependent Variable	Elite university (Project 985)				Downgrade			
Track type	Overall	Science	Humanity	Unified track	Overall	Science	Humanity	Unified track
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: All sample								
Female	-0.022***	-0.021***	-0.005***	-0.051***	-0.042***	-0.043***	-0.043***	-0.035***
	(0.003)	(0.002)	(0.002)	(0.013)	(0.003)	(0.003)	(0.003)	(0.009)
Mean of dependent variable	0.238	0.258	0.197	0.234	0.200	0.201	0.210	0.184
Observations	2,139,465	1,513,837	358,903+	266,725	2,127,964	1,504,147	357,092	266,725
R-squared	0.315	0.288	0.316	0.326	0.180	0.179	0.200	0.158
Panel B: High score sample								
Female	-0.048***	-0.045***	-0.000	-0.114***	-0.031***	-0.032***	-0.034***	-0.014
	(0.005)	(0.005)	(0.004)	(0.025)	(0.003)	(0.004)	(0.004)	(0.009)
Mean of dependent variable	0.537	0.440	0.548	0.582	0.200	0.201	0.210	0.184
Observations	550,136	434,044	57,997	58,095	547,122	431,194	57,833	58,095
R-squared	0.234	0.233	0.281	0.164	0.179	0.172	0.242	0.165
Panel C: Low score sample								
Female	-0.013***	-0.011***	-0.005***	-0.035***	-0.046***	-0.047***	-0.045***	-0.041***
	(0.002)	(0.002)	(0.002)	(0.009)	(0.003)	(0.003)	(0.004)	(0.010)
Mean of dependent variable	0.145	0.151	0.136	0.127	0.417	0.420	0.416	0.404
Observations	1,589,303	1,079,790	300,883	208,630	1,580,816	1,072,950	299,236	208,630
R-squared	0.161	0.152	0.201	0.143	0.141	0.135	0.173	0.120
Score×Year×Track×Province FE	✓	✓	✓	✓	✓	✓	✓	✓
Individual Control FE	✓	✓	✓	✓	✓	✓	✓	✓

Heterogeneity | Robustness checks

WHAT EXPLAINS THE GENDER GAP?

PART1: GENDER DIFFERENCE IN PREFERENCE

- Retake
  - Students accept worst could take risk in application
- Home preference
  - Home preference will limit available university pool
- Major preference
  - Elite university may mainly offer some majors
- Strong/weak majors within university
  - Choice between elite university bad major and bad university good major

GENDER DIFFERENCE IN PREFERENCE (EVIDENCE)

Gender difference does exist in all these preferences!

			Major Types			Major Ranking		
Dependent variable	Retake	Local preference	STEM	Business	Humanity	Top Quartile	Middle Quartile	Bottom Quartile
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: All								
Female	-0.0127***	0.0341***	-0.1280***	0.0257***	0.0732***	0.0084***	-0.0140***	0.0055***
	(0.001)	(0.003)	(0.007)	(0.002)	(0.005)	(0.002)	(0.002)	(0.001)
Mean of dependent variable	0.439	0.466	0.562	0.065	0.085	0.283	0.597	0.120
Observations	1647314	2,002,832	2,128,790	2,128,790	2,128,790	2,128,790	2,128,790	2,128,790
R-squared	0.465	0.250	0.332	0.065	0.216	0.095	0.055	0.075
Score×Year×Track×Province FE	✓	✓	✓	✓	✓	✓	✓	✓
Individual Control	✓	✓	✓	✓	✓	✓	✓	✓
Elite university & Downgrade FE	✓	✓	✓	✓	✓	✓	✓	✓

CAN THOSE PREFERENCES EXPLAIN THE GENDER GAP?

Dependent Variable	Elite university					
	(1)	(2)	(3)	(4)	(5)	(6)
Female	-0.0220***	-0.0206***	-0.0295***	-0.0212***	-0.0205***	-0.0184***
	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)
Mean of dependent variable	0.238	0.238	0.238	0.238	0.238	0.238
Observations	2,148,419	1,658,255	2,013 ,377	2,013,375	2,013,375	1,555,112
R-squared	0.315	0.312	0.331	0.347	0.352	0.352
Score×Year×Track×Prov FE	✓	✓	✓	✓	✓	✓
Individual Control	✓	✓	✓	✓	✓	✓
Retake FE		✓				✓
Local preference FE			✓			✓
Major FE				✓	✓	✓
Major Quartiles FE					✓	✓

Yet, those preferences collectively only can explain approximately 16% gap in elite university enrollment.

PART 2: INSTITUTIONAL SETTING OF ELITE UNIVERSITIES

- Government's label of elite universities (Project 985: 39 universities) favors science & technology universities. major share
  - 32 out of 39 can be viewed STEM leading while only 20 out of 39 can be viewed as Humanity leading
- What if government's label favors humanities or does not favor any side?

ALTERNATIVE DEFINITION OF “ELITE”

Need separate definitions for “science elite” and “humanity elite”.

elite university

Wushulian Ranking of Chinese Universities (2003):

- A widely recognized unofficial ranking
- Science/humanity elite: if a university ranks top 39 in the research capacity in science/humanity sub-ranking
- Dual elite: the university ranks top 39 in both ranking

- Do not Control for Preferences: Science Elite:-3.4%(Favor male) → Elite 985:-2.2% (Current Setting) → Dual Elite: +0.4% (Fair)→ Humanities Elite: 4.5%(Favor female)

Track	Overall	Overall	Science	Science	Humanity	Humanity
	(1)	(2)	(3)	(4)	(5)	(6)
R-squared	0.224	0.282	0.215	0.269	0.268	0.337
Panel C: Dual Elite(Fair Setting)						
Female	0.004***	-0.010***	0.005***	-0.010***	-0.007***	-0.011***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
Mean of dependent variable	0.099	0.099	0.094	0.094	0.131	0.131
Observations	2,139,465	2,013,375	1,513,837	1,415,911	358,903	346,842
R-squared	0.308	0.343	0.314	0.350	0.296	0.327
Score×Year×Track×Province FE	✓	✓	✓	✓	✓	✓
Individual Control	✓	✓	✓	✓	✓	✓
Major FE		✓		✓		✓
Major Quartiles FE		✓		✓		✓
Local preference FE		✓		✓		✓

- Control for Preferences: Robust 1% among different tracks → Risk attitudes

PART 3: PRE-EXAM VERSUS POST-EXAM SUBMISSION

- Gaokao: High-stake, High-risk
- If gender difference in risk attitudes plays a role, reducing uncertainties should reduce the gender gap.
- Pre-exam v.s. post-exam submission:
  - Post-score (post-exam) submission removes (reduces) score uncertainties (Wu and Zhong, 2014).
  - A similar reform in Turkey (Arslan et al., 2022)
- The Pros and Cons of "Pre-Exam Submission" vs "Post-Exam Submission" in Gaokao Application

REDUCE UNCERTAINTY DOES HELP WOMEN

Track type	Overall (1)	Overall (2)	Science (3)	Science (4)	Humanity (5)	Humanity (6)
Panel A: Elite 985						
Female	-0.041***	-0.037***	-0.029***	-0.023***	0.007	-0.001
	(0.008)	(0.008)	(0.004)	(0.005)	(0.007)	(0.007)
Female × Post policy	0.025***	0.022***	0.011**	0.006	-0.013*	-0.009
	(0.009)	(0.008)	(0.005)	(0.006)	(0.008)	(0.007)
Mean of dependent variable	0.238	0.238	0.248	0.248	0.197	0.197
Observations	2,149,263	2,023,161	1,522,218	1,424,287	359,381	347,313
R-squared	0.315	0.353	0.316	0.355	0.288	0.332
Score×Year×Track×Province FE	✓	✓	✓	✓	✓	✓
Individual Control	✓	✓	✓	✓	✓	✓
Major FE		✓		✓		✓
Major Quartiles FE		✓		✓		✓
Local preference FE		✓		✓		✓

- -4.1% in pre-exam setting, -1.6% in post-exam setting, -2.2% on average
- No effect for students get sufficient high scores

▸ Robustness Check

NO COSTS OF HIGHER DOWNGRADE RATES

Dependent Variable	Downgrade					
Track	All	All	Science	Science	Humanity	Humanity
	(1)	(2)	(3)	(4)	(5)	(6)
Female	-0.028***	-0.017**	-0.026**	-0.016*	-0.048***	-0.026***
	(0.008)	(0.007)	(0.010)	(0.009)	(0.009)	(0.008)
Female × Post policy	-0.018**	-0.006	-0.022**	-0.007	0.005	0.006
	(0.008)	(0.007)	(0.011)	(0.009)	(0.009)	(0.008)
Mean of dependent variable	0.366	0.366	0.366	0.366	0.375	0.375
Observations	2,137,722	2,011,763	1,512,494	1,414,682	357,564	345,520
R-squared	0.179	0.202	0.178	0.200	0.200	0.219
Score×Year×Track×Province FE	✓	✓	✓	✓	✓	✓
Individual Control	✓	✓	✓	✓	✓	✓
Major FE		✓		✓		✓
Major Quartiles FE		✓		✓		✓
Local preference FE		✓		✓		✓

- Free launch!

CONCLUSION

1. Women are 2.2% less likely to go to an elite university. As rewards, 4.2% less likely to "downgrade.".
   - Translate to 16860 elite seats just for 1999-2003!
2. Gender difference in Preferences $\rightarrow$ only 16% gap can be explained.
3. Disproportional favor STEM universities in elite project $\rightarrow$ Another 46% gender gap.
4. Gender difference in risk attitudes $\rightarrow$ remaining 38% gender gap.
5. An unexpected policy effect: 60% gender gap is reduced in post_exam policy compared to the pre_exam policy.

Thanks for Listening :)

Roujing WU
Ph.D. Student in Economics at the University of Bologna