Gender Differences in College Migration

Zihan Hu
Hyejin Ku
Xinzheng Shi
Yizhi Wang
Ming-ang Zhang*

June 2026

Abstract

Women now surpass men in college attainment in many countries, yet gender earnings gaps persist. We investigate whether women's lower propensity to migrate for college, limiting access to elite universities and stronger labor markets, contributes to this puzzle. Using China's score-based admissions system and nationwide college entrance-exam records, we find girls are 17.9% less likely than equal-scoring boys to enroll outside their home province, with gaps widening with traditional gender norms. Using policy-driven university expansions and pre-policy cross-province admission quotas, we estimate the causal returns to college migration and find them to be similar for both genders. Our estimates imply that closing this gap would shrink the elite-university attendance gap by 21.9% and the expected-wage gap by 11.5%. Notably, the gender gap in college migration is largest in less-developed regions, precisely where returns to out-migration are highest. Parent surveys reveal lower perceived migration returns for daughters as the primary driver.

Keywords: gender gap, college choice, labor mobility, talent allocation

JEL classification: J16, J61, I23

*Zihan Hu: Singapore Management University. zihanhu@smu.edu.sg; Hyejin Ku: University College London, RFBerlin, and CEPR. h.ku@ucl.ac.uk; Xinzheng Shi: Peking University. shixzh@pku.edu.cn; Yizhi Wang: Singapore Management University. yizhi.wang.2025@phdecons.smu.edu.sg; Ming-ang Zhang: Tsinghua University. zhangmingang@tsinghua.edu.cn. We thank Joseph Altonji, Pedro Carneiro, Raj Chetty, Adeline Delavande, Christian Dustmann, John Friedman, Christine Ho, Attila Lindner, Lin Ma, Xin Meng, Jessica Pan, and seminar and conference participants at UCL, CReAM, RFBerlin, IFS, Opportunity Insights, UC3M, Essex, Birmingham, QUB, SMU, AASLE, ES World Congress and CEPR LE Symposium, for helpful discussions and comments. Ruijie Cai provided excellent research assistance. Zihan Hu gratefully acknowledges the Singapore Ministry of Education (MOE) Academic Research Fund (AcRF) Tier 1 grant (Grant No. 24-SOE-SMU-105). Xinzheng Shi gratefully acknowledges the National Social Science Fund of China (25&ZD158) and the National Natural Science Foundation of China (72473076). Ming-ang Zhang gratefully acknowledges the National Natural Science Foundation of China (72203247). All errors are our own.

1. Introduction

Today, women surpass men in tertiary educational attainment in many countries (Goldin, 2006; Goldin, Katz and Kuziemko, 2006; OECD, 2024). Yet even among college graduates, women remain underrepresented at the upper end of the earnings distribution (Bertrand, 2018). The persistence of this earnings gap directs attention beyond the quantity of schooling to its quality, a critical determinant of post-college labor market outcomes.[1] In particular, equally qualified boys and girls may attend universities of differing quality or prestige, because attending a more selective institution often requires leaving home and the propensity to migrate for education may differ by gender. In this paper, we examine gender differences in college migration, quantify their implications for gender gaps in university quality and labor market earnings, and investigate the potential drivers of these differences.

In many settings, high-quality universities are geographically unevenly distributed (Chetty et al., 2020; Fu et al., 2022; Fabre, 2023; Boelmann, 2024; Yang, 2024; Hu and Ma, 2025; Ishimaru, 2025). As a result, students who are willing to move farther from home face a broader effective university choice set and may gain access to higher-quality institutions and stronger peer networks. Moreover, because college location is closely linked to post-graduation work location (Dustmann and Glitz, 2011), attending a college in a more economically vibrant region may also facilitate entry into stronger local labor markets after graduation. Therefore, migration for college, often among the first major voluntary migration decisions young adults make, can have lasting economic consequences through both university quality and location.

The ideal setting for isolating gender differences in college migration involves comparing equally qualified boys and girls who can be admitted to the same set of universities yet make different choices. This allows us to isolate individual migration decisions from other factors, such as gender differences in academic performance or university admission rules that may differentially favor female or male applicants. We leverage China's centralized college admissions system together with administrative data covering the universe of National College Entrance Examination (NCEE), or "Gaokao," takers.

In China, all high school graduates seeking college admission must take the NCEE and submit a ranked list of preferred universities. Admissions are determined solely by exam scores in a Boston mechanism, with no role for interviews, essays, or discretionary review (Li and

[1] See e.g., Kirkebøen, Leuven and Mogstad (2016); Zimmerman (2019); Chetty et al. (2020); Mountjoy and Hickman (2021); Jia and Li (2021); Chetty, Deming and Friedman (2026); Mountjoy (2026).

Qiu, 2025). Students are enrolled in either a science or humanities track in high school, and the NCEE is administered separately by track and province. Students in the same province-track-year who receive the same NCEE score face identical admission opportunities. Conditional on province-track-year-specific scores, observed gender differences in college migration outcomes only reflect gender differences in students' submitted university preference lists.

Our primary data source is administrative records from China's Ministry of Education (MOE) covering the universe of NCEE test takers from 1999 to 2003. The dataset includes over 23 million test takers and contains detailed information on NCEE scores, university and major admitted to (if any), home county, high school and academic track. It also records demographic characteristics, including gender, ethnicity, hukou status (agricultural versus non-agricultural), and birth year and month.[2] The population-level coverage of the data offers several important advantages. It ensures that our analysis is nationally representative, enables a highly granular examination of migration patterns, and allows us to accurately construct key university-level measures, such as province-track-year-specific enrollment quotas.

Conditional on exam performance—measured by percentile rank within province-track-year cells—female students are 4.5 percentage points (p.p.) less likely than their male counterparts to enroll outside their home province, a gap amounting to 17.9% of the sample mean. Because higher-achieving students are more likely to attend out-of-province universities, this gender gap is comparable in magnitude to the effect of an 11-percentage-point decline in percentile rank. The gender gap in college migration widens along the score distribution, indicating a larger gap when students face a less constrained set of eligible universities. It peaks around the 80th–90th percentiles, then narrows somewhat for the very top students.[3] Our results are robust to alternative measures of migration, including the geographic distance (in kilometers) between students' home prefectures (hereafter "cities") and the cities where their universities are located.

We then investigate how the gender gap in college migration associated with traditional gender norms. It is theoretically ambiguous since restrictive gender norms could either limit female mobility ("constraint effect") or motivate girls' migration to areas with greater equality

[2]Hukou is China's household registration system, which ties citizens to a specific location and classifies them as urban or rural, shaping their access to property ownership and public benefits, see Meng (2012). Importantly, hukou status does not restrict students from attending universities outside their hometown.

[3]This pattern likely reflects the spatial concentration of top universities in a few major cities, such as Beijing and Shanghai, which are not the home regions of most high school students. For top-ranking students who are highly likely to gain admission to top universities such as Peking University or Tsinghua University, staying local would entail a substantial downgrade in university quality, resulting in high migration rates for both boys and girls.

("incentive effect"). Empirically, we find a significantly larger gender gap in college migration among students from rural areas and counties where traditional gender norms are more prevalent. Our findings suggest that, on net, traditional local gender norms are more important in constraining female students' mobility than in motivating them to leave.

We then examine the consequences of the gender gap in college migration along two dimensions: university quality and expected post-college wages. We classify universities with Project 985 status as "high-quality" or "prestigious".[4] We compute the expected wage associated with each university by averaging the "best salary offer" received by final-year students, separately by local vs. non-local status and by gender, based on the China College Student Survey (CCSS). [5] To identify a causal effect, we adopt a shift-share instrumental variable (IV) approach (Goldsmith-Pinkham, Sorkin and Swift, 2020; Borusyak, Hull and Jaravel, 2022), where we exploit policy-driven university expansions since 1999 (Che and Zhang, 2018) and pre-policy cross-province admission quotas. Faster expansion in out-of-province universities that historically allocated more seats to a student's home province increases the propensity for college migration for reasons unrelated to the relative quality of, or expected wages associated with, out-of-province versus home-province universities.

Our IV estimates indicate that college migration roughly doubles the likelihood of attending a prestigious university relative to the sample mean and raises expected wages by about 20%. Notably, the causal migration premia in both university quality and expected wages are similar between male and female students. These estimates, together with the observed gender differences in college migration, imply that if girls migrated for college at the same rate as boys with comparable exam scores, the gender gap in university prestige would narrow by 21.9% and in expected wages by 11.5%.

In addition, we find that the estimated returns to college migration—in both university prestige and expected wages—are concentrated in less developed provinces with no Project 985 universities where the gender gap in college migration is also wider. This concerning pattern adds a spatial dimension to the debate on the misallocation of talent (Hsieh et al., 2019): the places where out-migration for college yields the greatest returns are precisely those where the gender gap in college migration is largest.

[4]Project 985 refers to a group of most prestigious universities designated by the Ministry of Education and is widely regarded as China's equivalent of the Ivy League (Zhang, Patton and Kenney, 2013).

[5]Conducted by the China Data Center of Tsinghua University, the CCSS surveys final-year college students across China, collecting information on students, their parental backgrounds, and their best job offers, including locations and monthly salaries. We use the 2011-2013 waves, which cover 65 colleges per year.

Why would equally qualified girls choose college migration at lower rates than boys from the same county, academic track, and with the same exam scores? Our findings that the gender gap is larger among students from rural areas and in counties with more traditional gender norms point to the importance of the local environment in shaping adolescents' educational choices. To unpack "traditional gender norm" into concrete concerns and trade-offs, we conduct a new survey targeting parents of upper-level (grade 11 or 12) high school students. The survey was administered in May 2025, one month before the 2025 NCEE, so most respondents should have given considerable thought to their children's college applications. The survey elicited parents' preferences for their child's college location and their beliefs about the benefits and costs of migrating for college.[6]

Although the survey was conducted around 25 years after the period covered in our NCEE data—parents of daughters remain significantly less likely than parents of sons to support college migration. The most influential factor behind this gap is perceived returns to college migration, accounting for approximately half of the gender difference in parental preferences and echoing research on the role of perceived returns in educational choices (Jensen, 2010; Zafar, 2011; Wiswall and Zafar, 2015). Parents of girls are significantly more likely to believe that the benefits of migration are smaller for females than for males, possibly reflecting an underestimation of the pecuniary returns for girls or parents placing less weight on university prestige and labor market earnings for daughters.[7] Parents of daughters also express significantly stronger preferences for local spouse matching and greater safety concerns about migration, accounting for 13.3% and 7.8% of the gender gap, respectively. Notably, while parental expectations for old-age companionship influence migration preferences, we find no gender differences in these expectations.

Another potential explanation for the gender gap in migration is gender differences in family socioeconomic status (SES). If female students disproportionately come from lower-SES households, resource constraints could limit their ability to migrate. Our analysis based on CCSS data, however, reveals the opposite: female college students are more likely to come

[6]Our survey focuses on parental preferences rather than students' own preferences due to the administrative burden of surveying minors through the survey company. Although parents cannot unilaterally dictate this decision, they exert substantial influence. Based on the CCSS data, we find that 41.6% of students list their parents as the most important decision-makers in finalizing the university preference list, while only 39.8% list themselves.

[7]In studying determinants of college enrollment in Mexico, Attanasio and Kaufmann (2017) show that the relative weight of labor market versus marriage market returns is larger for boys than for girls. Separately, Delavande and Zafar (2019) show, in the context of Pakistan, that students emphasize alignment of a school's teachings with their own ideology over expected earnings.

from families with higher incomes and more educated parents than their male counterparts from the same home city. [8] Family SES therefore cannot explain the gender gap in college migration.

We also investigate whether gender differences in decision-making under uncertainty, particularly risk aversion and self-confidence (see Eckel and Grossman (2008) and Croson and Gneezy (2009) for comprehensive reviews), contribute to the gender differences in college migration.[9] First, if female students are more risk-averse, they may prefer to apply to less selective universities with a higher likelihood of admission, thereby avoiding a sharp drop in university quality.[10] Second, if female students are less confident when forecasting their eventual scores and ranking, they may also be more inclined to apply to less selective universities. We find evidence consistent with women adopting a more conservative application strategy: conditional on the same rank percentile, women are more likely than men to be admitted to their first-choice university and to gain admission to any university.

However, we find little evidence that gender differences in risk attitudes and self-confidence translate into a gender gap in college migration. We leverage variation in the timing of university applications across provinces and over time. We classify university application systems into three levels of uncertainty based on the information available to applicants at the time of application: low-information regimes, where applications are submitted before the exam; medium-information regimes, submitted after the exam but before scores are released; and high-information regimes, submitted after scores are known. If risk aversion and self-confidence under uncertainty were key drivers, the gender gap should narrow in high-information regimes, where uncertainty is largely resolved. Exploiting within-province changes in application systems, we find no evidence of a reduced gender gap in medium- or high-information regimes. These results suggest that, while gender differences in risk aversion and self-confidence may influence application behavior more generally, they are unlikely to be the primary drivers of the

[8] This pattern partly reflects the fact that low-SES and rural households tend to exhibit stronger son preference and are less likely to send daughters to high school, a prerequisite for taking the NCEE. So the girls who appear in the NCEE data are more likely to be drawn from higher-SES families than the boys.

[9] A recent study by Cortés et al. (2023) shows that gender differences in risk aversion and overoptimism regarding future job prospects lead men and women to accept job offers at different points in the search process, contributing to gender gaps in accepted wages.

[10] Under the Boston mechanism, applicants who list a school as their first choice are considered first. Once a school fills its seats, rejected applicants move to their next choice and compete only for the remaining seats (Chen and Kesten, 2017). This creates a strategic trade-off: although University A may be only slightly better than University B, applying to A first can be risky because rejection may push the applicant down to a much worse outcome, such as University F. By contrast, listing the safer but still highly desirable University B as the first choice may substantially reduce the downside risk of admission to a much lower-ranked school.

gender migration gap.

We add to the large literature seeking to understand the remaining gender gap in earnings (Goldin, 2014; Blau and Kahn, 2017).[11] In particular, this study is closely related to recent work linking gender differences in commuting distance and job-search radius to gender wage gaps (Petrongolo and Ronchi, 2020; Le Barbanchon, Rathelot and Roulet, 2021; Caldwell and Danieli, 2024). Whereas this literature focuses on geographic constraints faced by women in the labor market, we highlight the transition from high school to college as an earlier stage at which gender disparities in geographic mobility emerge. We are also related to Borker (2021), who shows that women in Delhi are willing to forgo better universities in exchange for a safer commuting route. We extend this insight to cross-city college migration, showing that safety concerns matter beyond daily commuting. We further highlight other drivers of gender differences in migration preferences, including perceived returns to college migration.

Our findings also speak to the literature on how gender norms shape women's educational and career decisions (Fortin, 2005; Bertrand, 2011; Jayachandran, 2015; Bertrand et al., 2021). Existing work shows that women may avoid out-earning their husbands (Bertrand, Kamenica and Pan, 2015) and that single women sometimes conceal career ambitions because of marriage-market concerns (Bursztyn, Fujiwara and Pallais, 2017). We show that traditional gender norms can constrain geographic mobility even among highly qualified young women, effectively tying them to their home regions. As a result, women attend lower-quality universities and face reduced economic opportunities relative to men with comparable pre-college academic performance.

Finally, we contribute to the literature on internal migration (Blanchard and Katz, 1992; Molloy, Smith and Wozniak, 2011; Jia et al., 2023). Existing studies focus on migration driven by labor market opportunities and document higher migration rates among college-educated workers than among non-college-educated workers (Bound and Holzer, 2000; Wozniak, 2010; Malamud and Wozniak, 2012; Diamond, 2016). In contrast, we focus on migration for college education itself. In related work, Hu and Ma (2025) develop and estimate a dynamic spatial

[11] Prominent explanations include gender differences in sorting into firms (Card, Cardoso and Kline, 2016; Goldin et al., 2017; Barth, Kerr and Olivetti, 2021), occupations (Pan, 2015; Goldin and Katz, 2016; Wasserman, 2023), and fields of study (Turner and Bowen, 1999; Zafar, 2013; Wiswall and Zafar, 2018; Sloane, Hurst and Black, 2021; Wiswall and Zafar, 2021; Altonji et al., 2025), wage bargaining (Biasi and Sarsons, 2022; Roussille, 2024), work hours (Bertrand, Goldin and Katz, 2010; Azmat and Ferrer, 2017; Cortés and Pan, 2019), preferences for schedule flexibility (Mas and Pallais, 2017), and discrimination (Goldin and Rouse, 2000; Sarsons, 2017; Benson, Li and Shue, 2026). A recent survey by Olivetti, Pan and Petrongolo (2024) illustrates the many ways that gender differences in preferences and constraints may manifest in the gender gap in observed labor market outcomes.

model to examine how the uneven distribution of educational resources and college migration contribute to spatial inequality in China. By contrast, we focus on the *gender dimension* of college migration, and its implications for gender disparities in institutional quality and labor market outcomes.

2. Background

2.1. Quality and Spatial Distribution of Universities

During our sample period, there were around 1,000 universities in China, with around 60% being four-year universities and the rest as three-year universities, and the quality of these universities varies substantially.[12] At the top of the hierarchy sit 34 Project 985 Universities—widely viewed as China's Ivy League—followed by Project 211 Universities, another 57 institutions.[13] Project 985 and 211 Universities together received roughly 70% of all higher education research funding in China.[14]

Elite universities in China are highly unevenly distributed across provinces (Yang, 2024). Figure 1 shows the spatial distribution of Project 985 Universities across provinces as of 2003. Beijing and Shanghai alone hosted 6 and 3 Project 985 universities respectively, together accounting for 26.47% of all those elite universities despite representing a tiny fraction of China's population and geographic area. In sharp contrast, 13 out of 31 provinces had no Project 985 universities. This uneven distribution implies that, for high-achieving students from provinces with a worse endowment of high-quality universities, migration out of their home province is often necessary to attend better universities.

2.2. Exam-based Admission System

China's college admission system is organized around the NCEE. It can be understood as a centralized, score-based allocation based on the Boston mechanism, where students submit a ranked list of preferred universities (Chen and Kesten, 2017; Li and Qiu, 2025). The main advantage of this setting is that students in the same province, track, and year who receive the

[12] Source: *Basic Information on General Higher Education Institutions in 2000*, Ministry of Education in China.
[13] An additional five universities were elevated to Project 985 status after our data period concluded; we restrict our definition to the original 34 institutions throughout our empirical analysis.
[14] Source: *Compilation of Science and Technology Statistics of Higher Education Institutions*, Ministry of Education in China.

same NCEE score face identical admission opportunities, with no role for interviews, essays, or discretionary review. Conditional on province-track-year-specific scores, any observed gender differences in college migration outcomes only reflect gender differences in the university preference lists submitted by the students.

The NCEE is taken at the end of high school. Students are enrolled in either a science or humanities track in high school, and the NCEE is administered separately by track. Each of the 31 provinces administers its own version of the exam, making scores comparable within but not across province-tracks. Each university announces a pre-determined quota for each province-track, which is public information available to all students. Therefore, students compete only with peers in the same province-track, and students with the same NCEE score face identical admission opportunities.[15] Importantly, universities do not conduct interviews or use university-specific aptitude tests in the regular admission process during our data period.

Universities are assigned to different admission tiers. Admissions are conducted sequentially by tier. Higher-tier colleges admit students first, and lower-tier colleges begin their admission process only after admissions in higher tiers have been completed. The first tier (Tier 1) refers to the selective undergraduate programs, typically including Project 985 universities as well as other selective universities. The second tier (Tier 2) includes regular four-year undergraduate institutions that are less selective than Tier 1 but still grant bachelor's degrees. Below these are specialist or vocational colleges, which usually offer three-year diploma programs.

Students submit preference lists within admission tiers.[16] Within each tier, they rank their preferred colleges, typically 6–8 institutions. For each listed college, they also rank preferred majors, usually 3–5 majors.[17] The information available to students at the time of application differs across provinces. In some provinces, students submit preferences before taking the NCEE; in others, they submit preferences after the exam but before knowing their scores; and in still others, they submit preferences after learning their scores. Since the number of students within each score bracket will also be announced, students knowing their scores also understand

[15] While the vast majority of students rely solely on their exam performance, a small fraction receive bonus points for specific characteristics, such as ethnic minority status or exceptional athletic talent. To account for this, we use students' total score, including bonus points, rather than the raw examination score in our empirical analysis, ensuring comparability within the same admission choice set.

[16] Students may submit preference lists for multiple tiers, but lower-tier options become relevant only if a student is not admitted in a higher tier.

[17] Students must additionally indicate whether they are willing to accept reassignment to other majors within the same college if they are not admitted to any of their listed majors. Accepting major reassignment gives the university flexibility to place the student in an unlisted major with remaining seats; refusing reassignment can be costly, because a student who is admitted to a college but rejected by all listed majors may lose the admission opportunity at that college.

their relative ranking within the corresponding province-track.

During our sample period, admission operates under the Boston mechanism, a sequential matching system that allocates students to colleges round by round: each college first considers only students who ranked it as their top choice, ranks those applicants by total score, and admits them up to its quota for the relevant province-track. Students who fail to clear the resulting cutoff are passed to the next round, where their second-choice colleges apply the same procedure subject to remaining capacity, and so on.[18]

This Boston mechanism creates a strategic trade-off. For example, although University A may be only slightly better than University B, listing A as the first choice can be risky because rejection may push the applicant down to a much worse outcome, such as University F. By contrast, listing the safer but still highly desirable University B as the first choice may substantially reduce the downside risk of admission to a much lower-ranked school.[19]

Once admitted, students are allocated to majors within the college based on their major preferences and score: higher-scoring students are placed in their preferred majors first, with each major filling up to its pre-announced seats. A student who clears no listed major's cutoff and has declined major reassignment may be rejected by the college and passed to the next college on the preference list. Each student can receive only one admission offer. Once a student is admitted to a college, the admission process terminates for that student.

3. The NCEE Administrative Data

Our primary data is the individual-level administrative records for the universe of Chinese students (excluding those in Tibet) who took the NCEE between 1999 and 2003. The dataset is administered by the Ministry of Education (MOE) of China. It contains detailed information on around 23.35 million test takers, such as their gender, birth year, hukou (agricultural vs. non-agricultural), exam-taking year, academic track (humanities vs. science), NCEE raw test score, bonus point, and their home county. Since universities admit students based on the sum of the NCEE raw score and bonus points, we also use this sum to characterize students' ranking and university eligibility. Critically for our analysis, the dataset also includes final college and

[18] Over the past decade, many Chinese provinces have transitioned from the Boston mechanism to the parallel mechanism, whose implications are examined in Chen and Kesten (2017). During our study period, however, the Boston mechanism remained the dominant system. One exception is Hunan Province, which began piloting the parallel mechanism in 2001. Our results are highly robust to excluding Hunan Province.

[19] See e.g., Abdulkadiroğlu and Sönmez (2003) and Pathak and Sönmez (2008) for discussions on the manipulation vulnerability of the Boston mechanism.

major admission outcomes for each student.

Since our analysis focuses on migration for college, we focus on students who were admitted to any college (61.3% of all exam takers). We exclude students with missing information on crucial variables such as academic track and home county (3.46% of the observations). After these refinements, our final analytical sample consists of around 13.5 million students.[20]

Our main outcome variable—college migration—is defined as a dummy variable equal to one if the admitted college is located in a different province from the student's home province. We use provinces as the relevant geographic boundary for several reasons. First, provinces represent significant administrative, economic, and cultural units in China, with provincial boundaries often demarcating distinct labor markets, policy regimes, and social networks (Young, 2000). Second, the NCEE is administered provincially with scores comparable only within each province-track. Third, related to the second point, universities allocate a predetermined number of seats to students from each province. This quota system means that students compete only against other applicants from their home province. Provincial boundaries are therefore not merely geographic divisions but are structurally embedded in the admission process itself, making them the natural unit for analyzing college migration decisions. In addition to this extensive margin of migration, we also compute the straight-line geographic distance between the city where the university is located and the student's hometown.[21] This intensive margin of migration measurement captures migration intensity, includes within-province moves, and reveals both whether and how far students travel for college education.

Table 1 provides summary statistics for key variables, separately for boys and girls. Several patterns are worth noting. There are significantly more boys than girls in our sample, with girls accounting for only 44% of the sample, yielding a male-female ratio of 1.27 (= 56/44). This gap reflects deep-seated demographic and social patterns in China. First, there is a skewed sex ratio at birth of approximately 1.08 for the relevant birth cohorts (1980–1985), driven by widespread son preference in China (Qian, 2008; Almond, Li and Zhang, 2019). Second, girls, especially those from rural households, are less likely to attend high school (Zeng et al., 2014; Connelly and Zheng, 2003), a prerequisite for taking the NCEE exam.

As shown in Table 1, boys are 6 p.p. more likely to choose a university outside their home province and migrate 19% farther on average. However, this gender gap in college mi-

[20] Appendix Figure A1 shows the distribution of NCEE scores by gender for all exam-takers (solid line) and also for those admitted to any college (dashed line). As expected, the distribution of admitted individuals is shifted to the right for both boys and girls.

[21] The distance is coded as zero if the university is located in the student's home city.

gration cannot simply be attributed to differences in migration preferences or environmental constraints. Boys are simultaneously more likely to be in the science track and have slightly higher NCEE scores and rankings. Comparing students with identical scores from the same province-track is therefore essential. Notably, the share of agricultural hukou is higher among boys relative to girls. This pattern is an artifact of our focusing on high school graduates taking the NCEE rather than the general population. As mentioned above, low-SES and rural households tend to exhibit stronger son preference and are less likely to send daughters to high school. Therefore, females in the NCEE data are more likely to come from higher SES families than males. We further examine this issue in Section 6.2.

4. Gender Differences in College Migration

4.1. Main Results

4.1.1. Graphical Evidence

We begin our analysis by documenting the raw gender differences in college migration across NCEE score percentiles. We first rank students by their total NCEE scores within each province-track-year and then divide them into 100 percentile groups, with higher percentiles indicating higher rankings. For each percentile, we compute migration shares separately for boys and girls. Panel (a) of Figure 2 shows migration out of the home province for college, our main outcome of interest. Circles denote girls and triangles denote boys. For both groups, higher-ranked students are substantially more likely to migrate out of province, with rates rising from about 15% among the lowest-ranked students to over 70% among the highest-ranked. This pattern suggests that broader access to high-quality universities increases the propensity to leave the home province.

Furthermore, boys are more likely to go out of province for college education compared to girls at the same NCEE percentile. The gap is persistent across most percentile ranks (except the bottom 10%). As shown in Panel (b) of Figure 2, which calculates raw gender differences across percentile ranks, the gender gap in college migration widens along the score distribution, indicating a larger gap when students face a less constrained set of eligible universities. It peaks around the 80th–90th percentiles, then narrows somewhat for the very top students. This likely reflects the spatial concentration of elite universities in a few major cities, such as Beijing and

Shanghai, which are not the home regions of most high school students. For top-ranking students who are very likely to be admitted to the most elite universities such as Peking University or Tsinghua University, staying local would mean a substantial downgrade in university quality, making migration highly likely for both boys and girls. In contrast, for non-top students, the quality penalty from staying local is smaller, allowing other factors to play a larger role in their decision-making.

Panels (c) and (d) of Figure 2 show similar patterns using an alternative measure of college migration: the distance (in kilometers) between the university city and the home city. The pattern of gender difference in migration distance is similar to that based on out of province migration. Taking students at the 40th, 60th, 80th, and 100th percentiles as examples, boys, on average, migrate approximately 50, 72, 103, and 23 kilometers further from home than girls, respectively, corresponding to about 3.0% to 22.3% greater distances for boys.

4.1.2. Regression Analysis

We investigate this gender difference in college migration further by estimating the following equation controlling for students' hometowns and NCEE percentile ranks:

$$Migration_i = c_0 + \beta Female_i + \psi_{cst(i)} + \phi_{r(i)} + X_i'\lambda + \varepsilon_i \tag{1}$$

where $Migration_i$ is a measure of college migration for a student i while $Female_i$ indicates whether student i is female. Once we observe student i, we know his or her home county c, track s, and the year t when they took the NCEE. We therefore include county-track-year fixed effects, $\psi_{cst(i)}$, to account for common shocks affecting college choices and migration decisions of individuals in the same home county-track-year.[22] $r(i)$ is the percentile rank of the individual i within the corresponding province-track-year, the level at which students are ranked in the college admission process based on their NCEE scores. We include a full set of percentile rank dummies, $\phi_{r(i)}$, which non-parametrically control for the NCEE performance. We also include a set of demographic controls, X_i, including hukou type, ethnic minority indicator, and age at the time of taking the NCEE. Our coefficient of interest is β, which shows the gender gap in college migration conditional on the same home county-track-year and the same performance on the college entrance exam. Standard errors are clustered by home county.

[22] County is an administrative unit below prefecture/city, which in turn is below province. Hence, the inclusion of county-track-year fixed effects allows comparison of individuals with the same origin-community characteristics and academic track (i.e., science vs. humanities).

Table 2 presents the main results regarding the gender gap in migration for college education. The dependent variable is an indicator for attending university outside the hometown province, with the female dummy being the primary variable of interest. All regressions include county-track-year fixed effects, ensuring that the comparison is made between boys and girls who take the NCEE in the same county, academic track, and year. The estimates in column (1) show that female students are 5.1 p.p. less likely to choose a college outside their province than male students, and this difference is statistically significant at the 1% level.

In column (2), we further include each student's NCEE percentile rank. The coefficient on percentile rank (0.004) indicates that a 1-percentage-point increase in a student's ranking increases the probability of migrating out of the province by 0.4 p.p. The coefficient on the female indicator (-0.045) is comparable in magnitude to the effect of an 11-percentage-point decline in NCEE ranking ($= \frac{-0.045}{0.004}$). In column (3), we replace the linear control for percentile rank with percentile-rank fixed effects to allow for a more flexible control of the students' academic ranking, and the results remain similar.

In column (4), our preferred specification, we further add student demographic characteristics, including dummies for agricultural hukou, ethnic minority, and age at the time of the NCEE exam. The estimates show that girls are, on average, 4.5 p.p. less likely to migrate out of their province for college education compared to boys from the same county-track-year who are eligible for a comparable set of university admissions. This remaining gender gap, after including the full set of controls, is statistically significant at the 1% level and represents 17.9% of the dependent-variable mean ($= \frac{-0.045}{0.251}$).

To explore heterogeneity in the gender gap across the score distribution, we estimate equation (1) separately for each percentile. The estimates are plotted in Figure A2. The figure reveals a pattern consistent with the raw data in Figure 2: the gender gap widens with NCEE scores, except among students in the top 10% of the distribution, where the pattern reverses.

We also conduct analysis separately by academic track with results reported in Appendix Table A1. The gender gap in college migration is larger among science-track students, at 4.8 p.p. relative to a mean of 0.265, but remains sizable among humanities-track students, at 2.9 p.p. relative to a mean of 0.215. These results indicate that the gender gap is not confined to science-track students, although it is more pronounced among them.

4.2. Robustness

This section shows that the gender gap in college migration is robust to alternative migration measures and sample restrictions. We begin by examining alternative definitions of migration in Table 3. Column (1) replaces the out-of-province binary outcome with the university–hometown distance (in kilometers), capturing the intensive margin that pools within- and out-of-province moves. We find that boys, on average, enroll at institutions located about 12.5% further from home ($= \frac{-45.155}{360.424}$). Column (2) examines within-province migration, measured by moves to the provincial capital among students who enroll within their home province and are not originally from the provincial capital. Girls are 4.2 p.p. (or 8.8% of the dependent-variable mean) less likely than academically comparable boys from the same county-track-year to migrate to the provincial capital, confirming the gap exists even for shorter-distance moves, though its magnitude is smaller than the primary out-of-province gap.

Columns (3) and (4) use indicator variables for migration to high-opportunity destinations: (i) any top-30 city by total GDP and (ii) any top-10 city by total GDP (based on year 2000 rankings). For each column, students originally from the corresponding set of cities are excluded. The estimates show that female students remain less likely than male students to migrate to economically developed cities. Specifically, females are 2.5 p.p. less likely to move to top-30 GDP cities and 1.4 p.p. less likely to move to top-10 GDP cities, corresponding to 6.4% and 7.8% of the respective means.

These magnitudes are smaller compared with the main results based on out-of-province migration representing 17.9% of the dependent mean. The narrower gender gap for migration to economically developed cities likely reflects larger benefits from studying there (elite universities, stronger labor markets, broader networks), which attenuate gender-specific constraints. As the benefits of migration become more substantial, gender-specific constraints that influence migration decisions will play a relatively smaller role, resulting in a smaller gender gap in these contexts. This pattern parallels our finding of a smaller gender gap among the very top students (see Figure A2), for whom the expected returns to migration are especially large.

Next, we check the robustness related to sample construction and institutional features, and report the results in Appendix Table A2. Column (1) replicates the baseline results for comparison. In column (2), we restrict the sample to students admitted to their first-choice university, whose admissions more directly reflect intended choices rather than fallback placements. In

column (3), we exclude students admitted through special categories (such as sports or arts), whose admissions are governed by different rules. Column (4) excludes students whose high school province differs from their NCEE province, to avoid conflating college migration with pre-college migration. Finally, in column (5), we exclude students who received bonus points (e.g., for ethnic minority status or exceptional athletic talents), as males and females may differ in their likelihood of receiving such points. Across all these exercises, the estimated gender gap in college migration remains large and statistically significant.

4.3. Heterogeneity by Family and Hometown Characteristics

Having shown that female students are, on average, less likely to migrate for college than their male counterparts, we now examine how this gender gap varies by family and hometown characteristics. In particular, we focus on heterogeneity driven by hukou status (agricultural vs. non-agricultural) and prevailing gender norms in home counties.

We begin by investigating heterogeneity according to individuals' hukou status. Because educational attainment and income are generally lower in rural areas, students with an agricultural hukou are more likely to come from families with lower SES. In column (1) of Table 4, we present the interaction between the female indicator and an agricultural hukou indicator. The coefficient on the interaction term is similar in magnitude to the main effect of Female, implying that the gender gap in college migration is nearly twice as large among rural students as among urban students.

Next, we investigate the role of local gender norms in students' home counties. The effect of more traditional gender norms on the gender gap is theoretically ambiguous. On the one hand, restrictive gender norms could limit female mobility (the "constraint effect") by depressing girls' aspirations and lowering the importance of their career success in the eyes of their parents. On the other hand, a local environment with a stronger bias against females may incentivize girls' migration to more gender-equal areas (the "incentive effect").

We consider three different county-level gender norm measurements. The first is the sex ratio at birth. Based on the *Tabulation on the 2000 China Population Census (By County)*, we calculate the male-to-female sex ratio at birth for children aged 0–4.[23] A higher sex ratio indicates a stronger preference for sons and more conservative local gender norms (Edlund,

[23] For sex ratios at birth, ideally, we would use data for 0-year-olds in the 2000 Census. However, the county-level tabulation only reports the number of boys and girls aged 0-4 instead of a single-year age 0.

1999; Choi and Hwang, 2015). The second and third gender norm measurements are derived from the China General Social Survey (CGSS). Similar to the General Social Survey in the U.S., CGSS collects information on social behaviors, beliefs, ideology, and lifestyles. Using the 2010 CGSS data, which cover 12,000 households in 122 counties, we compute county-level averages of agreement with the following two statements separately: (1) "Men should focus on career, while women should focus on family," and (2) "Marrying well is more important for women than doing well." The responses are rated on a scale from 1 (strongly disagree) to 5 (strongly agree), where higher averages indicate stronger traditional gender norms in the local county. For comparability and ease of interpretation, we standardize all three county-level measures to have mean zero and unit standard deviation.

We interact the female indicator with each measure of gender norms as key independent variables, with county-track-year fixed effects absorbing the county-level gender norm measurements. The results are presented in columns (2)-(4) of Table 4.[24] The coefficients on the interaction terms are significantly negative across specifications, suggesting that the gender migration gap is wider in counties with more traditional gender role attitudes. In columns (2)-(4), a one-standard-deviation increase in these social norm measures amplifies the gender gap in college migration by 6.7% ($= \frac{0.003}{0.045}$), 17.9%($= \frac{0.007}{0.039}$), and 15.4% ($= \frac{0.006}{0.039}$) of the baseline gap, respectively. We interpret these coefficients as net effects of two opposing mechanisms: females' incentive to migrate toward more gender-egalitarian regions versus the constraints imposed by local conservatism. The consistently negative coefficients suggest that the constraints imposed by traditional gender norms dominate the "escape" motive.

5. Consequences of the Gender Gap in College Migration

Geographic immobility mechanically restricts the choice set of universities, lowering the quality of institutions students attend. This effect can be exacerbated in disadvantaged regions where high-quality institutions are scarce. Beyond university quality, students who remain local, especially those from less developed provinces, are less likely to access superior economic opportunities and labor market networks in more developed cities. Both university quality and labor market access channels can lead to lower expected wages.

[24] It is worth noting that, when matched with the CGSS data, the number of observations in columns (3)–(4) declines because the matched dataset includes only 122 counties. However, the coefficient on the Female indicator remains similar in magnitude across samples.

To investigate the consequences, we first obtain causal estimates of the college migration premium in terms of both university quality and expected wages. Combining these estimated premiums with the gender gap in migration documented earlier, we quantify the contribution of the gender migration gap to the gender gaps in university quality and wages through counterfactual analyses, simulating scenarios in which female students exhibit the same geographic mobility as their male counterparts.

5.1. College Migration Premium

Because students sort into college migration based on individual-specific expected returns, comparing the realized outcomes of migrants with those of non-migrants does not recover the causal returns to migration (Borjas, 1987; Dahl, 2002). Therefore, we adopt a shift-share IV approach (Goldsmith-Pinkham, Sorkin and Swift, 2020; Borusyak, Hull and Jaravel, 2022), exploiting supply-side variation in university admission quotas. Specifically, we combine policy-driven university expansions initiated in 1999 with pre-policy variation in province-specific quotas at different universities. For example, an aggregate quota expansion at a Beijing-based university that historically allocates a large share of its admissions quota to Sichuan province will disproportionately increase migration opportunities for Sichuan students relative to students from other provinces.

The university expansions come from a nationwide policy initiated by the Chinese State Council in 1999. The policy was centrally mandated and generated substantial variation in the magnitude of expansion across universities (Che and Zhang, 2018). As discussed in Section 2, universities also set different admission quotas for students from different provinces. Because each university's propensity to admit students from a given province is highly persistent over time, we use the 1999 cross-province quota distribution—the first year of our sample—as the "shares" and exclude 1999 from the IV analysis.

Formalizing the idea above, consider the following instrument capturing the pull factor:

$$Out_{pst} = \frac{\sum_{j \notin J_p} \frac{Quota^{p}_{j,s,1999}}{Quota_{j,s,1999}} \times Quota_{j,s,t}}{N_{pst}} \tag{2}$$

where $Quota^{p}_{j,s,1999}$ refers to the admission quota at university j allocated to province p and track s in 1999 and $Quota_{j,s,1999}$ is the total quota available at university j for track s in 1999. The ratio captures the tendency of university j seats going to province p's students. We combine

this initial share with university j's seats for track s in year t, $Quota_{j,s,t}$. We construct instrumental variable Out_{pst} by summing the projected admission quota over the set of all universities outside the students' own province p, where J_p denotes the set of universities within province p.[25] We further normalize it by the number of exam takers, N_{pst}, in the corresponding province-track-year.

Out_{pst} defined above captures the pull factor that draws students away from their home province by expanding opportunities in destination provinces. Migration decisions are fundamentally determined by both pull and push factors. Analogously, we also construct In_{pst} as in equation (3), based on projected quotas at universities within the province, capturing the retention force that affects students' propensity to leave:

$$In_{pst} = \frac{\sum_{j \in J_p} \frac{Quota^{p}_{j,s,1999}}{Quota_{j,s,1999}} \times Quota_{j,s,t}}{N_{pst}} \tag{3}$$

Based on these two instruments, we estimate the following first-stage equation:

$$Migration_i = c_1 + \gamma_1 In_{pst(i)} + \gamma_2 Out_{pst(i)} + \theta_{cs(i)} + \delta_{t(i)} + \phi_{r(i)} + X_i'\lambda + u_i \tag{4}$$

where the dependent variable $Migration_i$ is an indicator for whether student i migrates out of their home province for college, and $In_{pst(i)}$ and $Out_{pst(i)}$ serve as the instruments. Once we observe student i, we know his or her track s, home county c within province p, and the year t when they took the NCEE. We condition on county-track fixed effects ($\theta_{cs(i)}$) and year fixed effects ($\delta_{t(i)}$).[26] We further control for percentile rank fixed effects, $\phi_{r(i)}$, and individual demographic characteristics (X_i), including hukou status, ethnic minority indicator, and age at the time of taking the NCEE.

Panel A of Table 5 presents the first-stage results. Columns (1) and (2) are based on the full sample of admitted students, estimated separately by gender. Taking column (1) as an example, among female students, a one-unit increase in the out-of-province quota per student increases the probability of out-migration by 26.7 p.p., acting as a robust pull factor. Conversely, a

[25] Appendix Figure B1 demonstrates a high correlation between projected and actual quotas at the university-province-track-year level.

[26] Note that because $In_{pst(i)}$ and $Out_{pst(i)}$ vary at the province-track-year level, we cannot include county-track-year fixed effects as in equation (1), since counties are nested within provinces and county-track-year fixed effects will absorb all usable variation at the province-track-year level. Our results in Table 2 remain robust if we replace county-track-year fixed effects with county-track fixed effects and year fixed effects, indicating that this choice is not consequential.

one-unit increase in the in-province quota per student decreases the out-migration probability by 16.1 p.p., highlighting a retention effect. Both estimates are statistically significant at the 1% level. This push-pull pattern holds for both genders, although male students in column (2) exhibit slightly larger responses to both types of quota expansion than their female counterparts in column (1).

Due to the LATE nature of our IV estimations, it is important to understand whether the compliers come from segments of the NCEE score-percentile distribution where the gender gap in migration is salient. Appendix Figure B2 plots the first-stage estimates across percentiles by gender. Both pull (out-of-province quota) and retention (in-province quota) effects show strong predictive power across score distribution. Crucially, the instruments have the strongest effects at upper-middle and top percentiles (60th–90th), where gender gaps in migration are largest. Since quota expansions induce compliance among high-achieving students who drive gender differences in migration, our IV estimates are highly relevant for counterfactual analyses.

Our analysis focuses on two dimensions of student outcomes: university quality and post-college labor market outcomes. To measure university quality or prestige, we require a metric that is both objective and widely accepted. Due to the lack of reliable, comprehensive university rankings during our sample period (early 2000s), we employ "Project 985" status as our primary proxy for institutional quality, since it is the most widely recognized indicator of university prestige in China, carrying a reputation and labor market signal analogous to the Ivy League in the United States (Zhang, Patton and Kenney, 2013).

Measuring labor market outcomes is more challenging since our NCEE data does not track students after their graduation from universities. To address this, we construct expected wages using 2011-2013 waves of the China College Student Survey (CCSS) data, which records the highest salary offer received prior to graduation for each final-year student.[27] We compute the average monthly starting salary (in Chinese yuan) for each university, further stratified by gender and student origin (local vs. non-local to the university's province), to account for potential labor market advantages of local students stemming from hukou status or local networks (Schmutte, 2015). We then match these university-gender-origin-specific averages to the NCEE admission records, restricting the sample to universities covered by the CCSS. This allows us to impute expected post-college wages for about 1.1 million students. The first-stage

[27] The CCSS survey is conducted by the China Data Center of Tsinghua University (see Jia and Li (2021) for details). During 2011-2013, CCSS surveys 300 graduating senior students per university from 65 universities each year. It provides detailed information on their family background and labor market outcomes.

results for this matched sample are reported in Columns (3) and (4) of Panel A of Table 5, and are consistent with those for the full sample in columns (1) and (2). Essentially, this measure represents a gender- and origin-specific labor market value of each university, such as university quality and location advantage.

Panel B of Table 5 reports the second stage estimates of our 2SLS regressions for the 2000–2003 sample (the corresponding OLS estimates are shown in Appendix Table B1). Columns (1) and (2) show that migrating female students are 10.2 p.p. more likely to attend a Project 985 university, while the corresponding increase for male students is 12.3 p.p. Hence, college migration roughly doubles the likelihood of elite university attendance relative to the sample mean. Columns (3) and (4) show migrating out of province yields a substantial return in terms of labor market outcomes: expected monthly wage increases by 701.8 yuan for females and 599.8 yuan for males, corresponding to a 24.9% and 19.4% increase compared to the respective sample means. Overall, these results show that the returns to migration are similar between genders. If anything, the point estimates suggest that girls may have slightly higher migration premiums in terms of expected wages.

The estimates above underscore the critical role of geographic mobility in accessing prestigious universities and for better post-college labor market outcomes, given the heavily unevenly distributed high-quality educational resources in China. Table 6 separates provinces into those with and without Project 985 universities (the corresponding first-stage results are reported in Appendix Table B2). It is important to note that the presence or absence of Project 985 universities is not only an indicator of the availability of elite university seats, but also serves as a broad predictor of general educational resources and economic development opportunities.[28] The results show that migration premiums in both dimensions are concentrated among students from provinces without any Project 985 universities, whereas the benefit of leaving a resource-rich home province is substantially smaller. Importantly, the gender gap in college migration is also larger in provinces without Project 985 universities (see Appendix Table B3). This reveals a concerning pattern of spatial-gender mismatch: the places where migration can deliver the most substantial economic and educational benefit are precisely the locations where female students face greater geographic mobility constraints relative to their male counterparts.

[28] For example, as of 2003, Beijing and Shanghai together hosted 9 Project 985 universities, accounting for 26.47% of all Project 985 universities. By contrast, provinces without Project 985 universities, such as Ningxia and Guizhou, tend to be less developed and offer fewer high-quality educational and employment opportunities.

5.2. Counterfactual Exercise

Given the substantial gains from out-of-province college attendance, the gender gap in college migration may contribute to gender disparities in university quality and subsequent labor market outcomes. We conduct a counterfactual analysis to quantify how much the gender gaps in university quality and expected post-graduation wages would shrink if female students had the same migration probability as male students.

Specifically, for each outcome, we multiply the gender gap in college migration by the corresponding migration premium for females—0.102 for attending a prestigious university and 701.8 yuan for expected wages—and then divide this product by the observed gender gap in that outcome. We use the coefficient on female (0.045) from column (4) of Table 2 as the gender gap in migration. As shown by the means of dependent variables in Table 5, boys are 2.1 p.p. more likely to attend Project 985 universities and have expected wages 274.2 yuan higher than girls. Applying these parameters, we find that equalizing the probability of female migration to that of males would narrow the gender gap in Project 985 university attendance by 21.9% and the gap in expected wages by 11.5%.

6. Potential Drivers of the Gender Gap in College Migration

Having documented the gender gap in college migration and its consequences, we now explore potential drivers of this gap, focusing on the following candidates: (1) gender differences in perceived benefits and costs of college migration; (2) gender differences in family background; and (3) gender differences in decision-making under uncertainty.

6.1. Perceived Benefits and Costs of College Migration

The robust gender gap in college migration documented in Section 4 is based on the cohorts taking the NCEE between 1999 and 2003. Given China's rapid economic development and shifting cultural attitudes over the past two decades, a natural question is whether this gender gap in migration preferences persists today. Moreover, building on our finding that the gap is larger in counties with more traditional gender norms, we aim to unpack the broad concept of gender norms into concrete concerns and trade-offs underlying gender differences in migration decisions. To address both questions, we design and conduct an original survey.

6.1.1. Survey Design

The survey was administered online in May 2025 through a professional survey company. It specifically targets parents with children currently in the final two years of high school in China (grade 11 or 12), whose children will take the NCEE approximately one year later or one month after the survey. This timing is strategically chosen to capture families at a critical decision-making period, since the NCEE takes place in early June each year. By May of their final or penultimate high school year, families are deeply engaged in NCEE preparation, with university selection often among the most discussed household topics. Respondents are therefore more likely to offer well-considered views on college choices.

Ideally, we would survey both parents and students to understand the preferences of both parties toward college migration and to examine potential intergenerational differences in attitudes toward studying outside one's home province. However, we face a practical constraint. Since students in our target population are not yet adults, the survey company informed us that surveying minors requires approval from both their parents and schools. Given the heavy administrative burden associated with such procedures, we chose to survey parents only. Although parents cannot fully dictate their children's college choices, they exert considerable influence in China, where respect for parental authority remains culturally salient. The CCSS data support this point: 41.6% of students list their parents as the most important decision-makers in finalizing the university preference list, while only 39.8% list themselves.

The survey covers Guangdong, Hubei, and Shaanxi provinces with target sample sizes proportional to their respective populations of high school graduates.[29] We distributed 3,995 questionnaires and collected 2,992 valid responses, with a response rate of 74.9%.[30] This high response rate is likely attributable to the salience and relevance of the topic for those parents. After excluding observations with invalid city information, we obtain a final sample of 2,940 parents, with 1,447 respondents from Guangdong, 980 from Hubei, and 513 from Shaanxi.

The survey collects demographic and academic information about the child (in grade 11 or 12), including the child's gender, number of sisters and brothers, city of high school, academic

[29] We chose these provinces for two main reasons. First, they represent the eastern, central, and western regions of China with distinct levels of GDP per capita, thereby capturing geographic and socioeconomic diversity. Second, all three provinces host several high-quality universities, including Project 985 institutions, which ensures that parents face a meaningful choice between encouraging children to remain in the home province versus migrating out of the province for higher education. The gender gap in college migration among students in these three provinces is 4.3 p.p. in the NCEE data, quite similar to the national estimate of 4.5 p.p. reported in Table 2.

[30] Only respondents who took more than two minutes to complete the questionnaire are treated as valid responses.

track (science or humanities), and parents' assessment of the university level the child is expected to attain (ranging from Project 985 universities to vocational colleges). We also collect information on parental background, including the gender of the survey respondent (whether the father or the mother was surveyed), hukou type, education level, and monthly income (in bracketed categories). Appendix Table C2 presents corresponding descriptive statistics.

We first measure parental preferences for college migration based on the following question: *"If your child could attend universities of the same quality in the home province, in another province (excluding Beijing and Shanghai), or in Beijing or Shanghai, how would you like your child to choose? Please rank the three options from most preferred to least preferred."* We define "migration preference" as 1 if the respondent ranks either of the out-of-province options above the home-province option. The summary statistics in Panel B of Appendix Table C2 show that girls' parents are 5 p.p. less likely to prefer college migration than boys' parents.

We ask about specific concerns and trade-offs regarding college migration only after parents indicate their general location preference.[31] We ask parents to rate their degree of concern regarding safety, the ability to provide timely assistance, and financial costs associated with attending university in another province. We also elicit concerns related to finding a spouse if their child migrates, as well as whether they have a strong preference for their child to marry someone from the home province. In addition, we solicit expectations regarding future elderly care, distinguishing between financial support and time spent accompanying them. The descriptive statistics of corresponding variables are summarized in Panel B of Appendix Table C2, and the exact wording of each question is attached in Appendix C.1.

One important factor in migration decisions is the perceived returns to migration. We ask parents whether they believe attending university outside the home province yields greater benefits for boys, for girls, or similar benefits for both. This question directly elicits parents' subjective assessments of whether out-of-province college attendance generates gender-specific returns. Based on the gender of the respondent's child, we transform this categorical response into a binary indicator of whether parents believe migration benefits are larger for their child's own gender than for the opposite gender. As shown in Panel B of Appendix Table C2, one quarter of boys' parents believe migration benefits are larger for boys, while only 7.4% of girls' parents believe migration benefits are larger for girls.

Appendix Table C1 examines how parental concerns and perceived benefits correlate with

[31] This is to prevent their university location preference rankings from being influenced by the salience of specific concerns raised in the subsequent questions.

their preference for their child to attend college outside the home province. Column (1), pooling parents of boys and girls, shows that several factors are strongly associated with migration preferences. Marriage-related considerations are important: parents who are concerned about their child finding a suitable spouse outside the province, especially those who prefer their child's future spouse to come from the home province, are significantly less likely to prefer out-of-province colleges. Elderly-care considerations also matter, but mainly through the demand for physical proximity: expectations of time or companionship support are negatively associated with migration preference, whereas expectations of financial support are not. Practical concerns about safety and the ability to provide timely assistance are also negatively associated with migration preference, while cost concerns are insignificant in the pooled sample. In contrast, parents who believe that out-of-province study yields greater benefits for their child's gender are significantly more likely to support migration. Columns (2) and (3) report estimates separately for parents of boys and parents of girls. Overall, the relationship between migration preferences and the above factors is broadly similar across child gender.[32]

6.1.2. Parental Attitudes to College Migration by Child Gender

In Table 7, we examine the differences in parental attitudes to college migration by child gender. All regressions include the following set of controls: expected university-level fixed effects, income-bracket fixed effects, home city-by-academic-track fixed effects, and demographic characteristics.[33] Column (1) shows that parents of girls are less likely to prefer an out-of-province university than parents of boys. The estimated gender gap in migration preference is 5.4 p.p. Interestingly, this magnitude is similar to the gender gap of 4.5 p.p. documented in Table 2 using administrative data approximately 25 years ago.[34] This suggests a persistent gender gap in college migration preferences in China, despite the significant economic development and rising female educational attainment.

Next we examine whether there are systematic gender differences in the perceived costs and benefits of college migration. Regarding general marriage prospects, column (2) shows that

[32] One exception is cost concern, which significantly reduces migration preferences among parents of boys but not among parents of girls. This pattern suggests that weaker support for daughters' migration is unlikely to be driven by greater cost sensitivity or lower willingness to invest financially in daughters. Instead, gender gaps in migration preferences appear to reflect non-financial factors.

[33] Demographic characteristics include whether the respondent is the mother, whether the respondent has agricultural hukou, the respondent's years of education, and whether the child is an only child.

[34] When considered relative to the mean, the gap in Table 2 represents 17.9% of the mean migration rate. In this survey analysis, the corresponding ratio is 13.7%($= \frac{0.054}{0.394}$), only slightly lower.

parents of sons are more concerned than parents of daughters that studying outside the province would make it difficult for their child to find a suitable spouse. Although migration to a better city or university may expand the marriage-market choice set, it may also weaken access to local family networks that facilitate matching. This concern may be especially salient for sons, given that men face greater competition in China's marriage market due to the imbalanced sex ratio (Wei and Zhang, 2011). Parents of sons may also worry that their children would face disadvantages in destination marriage markets if local families prefer sons-in-law from the same province. Consistent with this possibility, column (3) shows that parents of daughters are more likely to prefer that their child's future spouse come from the home province. This pattern may reflect concerns that daughters are more likely to relocate after marriage if they marry someone from another province; a local spouse may therefore help keep daughters geographically closer to their parents in the long run.[35]

These two marriage-related measures therefore have different implications for the gender gap in migration preferences. Parents' greater concern about sons' general marriage prospects works against the observed gender gap in migration preferences and cannot explain weaker support for daughters' migration. By contrast, parents' stronger preference for daughters to marry someone from the home province helps explain why parents are less willing to support daughters' out-of-province college attendance.

Turning to columns (4) and (5) regarding elderly care, the results indicate that parents of daughters have lower expectations for old-age financial support and report similar levels of expectation for time and companionship support compared to parents of sons. Given that financial concerns show no significant association with migration preference (Appendix Table C1) and that expectations for companionship are similar across genders, it appears that elderly care considerations play only a limited role in explaining the gender gap in college migration.

Columns (6)-(8) focus on other parental concerns related to their child's college migration. Understandably, parents of daughters are significantly more likely to express safety concerns than parents of sons, a factor that likely contributes to the gender gap in college migration.[36] This difference in safety concern could reflect both objective differences in vulnerability—particularly in unfamiliar environments where they lack established social networks and local knowledge—and potentially gendered cultural norms about appropriate protection of daughters

[35] See Bau (2021) for a discussion of how matrilocal and patrilocal traditions affect parental investments in children's education.

[36] This echoes Borker (2021)'s finding that safety concerns play an important role in university choice within Delhi.

versus sons. Meanwhile, there are no clear gender differences in concerns about the financial cost of migration or the ability to provide timely assistance.

To quantify the relative importance of these mechanisms, we combine the estimated coefficients from Table 7 and Appendix Table C1. Assuming these migration concerns affect migration preferences according to the full-sample estimates in Column (1) of Appendix Table C1, eliminating the gender differences in those concerns would reduce the overall gender gap in migration preference by: (1) 13.3% due to local spouse preferences; (2) 7.8% due to safety concerns; and (3) 3.1% due to expectations for elderly-care time companionship.[37] If we instead use the girls-only sample coefficients from Column (3) of Appendix Table C1 to calibrate these shifts, the reduction in the gender gap is even more pronounced: 17.0% for local spouse preferences, 10.2% for safety concerns, and 2.5% for elderly-care expectations for companionship.

Column (9) of Table 7 examines gender differences in the perceived benefits of migration. Parents of daughters are significantly more likely to believe that out-of-province college yields lower benefits for girls compared to boys. Following the same decomposition approach as for the concern variables, we find that eliminating gender differences in perceived migration benefits could reduce the gender gap in migration preference by 30.4% using the full sample estimation results or 52.0% using the girls' sample estimation results. This represents a substantially larger contribution than any of the individual concern variables examined earlier. This highlights perceived migration benefits as a critical mechanism driving gender gaps in college location preferences, consistent with the literature highlighting the importance of perceived benefits in driving education choices (Jensen, 2010; Zafar, 2011; Wiswall and Zafar, 2015).

The gender differences in perceived benefits may reflect three channels. First, parents may expect the returns to migration to diverge over the life cycle. Although our causal estimates in Section 5 show similar immediate returns for men and women in terms of university quality and initial wages, parents may believe that daughters benefit less in the long run if they expect women to face shorter or less continuous career trajectories due to slower promotion, earlier retirement, or career interruptions (Goldin, 2014; Bertrand, 2018; Cortés and Pan, 2023). Second, parents may place different weights on pecuniary and non-pecuniary outcomes by child

[37] The 13.3% figure is obtained by multiplying the gender difference in local spouse preference, 0.068, by its association with migration preference, -0.106, and dividing by the gender difference in migration preference, -0.054. Similarly, the implied contributions are $\frac{0.064\times(-0.066)}{-0.054} = 7.8\%$ for safety concerns and $\frac{0.021\times(-0.080)}{-0.054} = 3.1\%$ for expectations for elderly-care time companionship.

gender: even when material returns are similar, parents of daughters may give greater weight to safety, marriage prospects, or proximity to family. Third, parents may simply underestimate daughters' returns to migration, representing an information friction that could potentially be reduced by providing information on actual returns (Jensen, 2010).

6.2. Family Background

Although tuition varies little across regions for universities of the same tier, attending college away from home entails additional transportation and living costs. Students from wealthier households may be better able to afford migration. In this subsection, we examine whether gender differences in family background can explain the gender gap in college migration.

The relationship between gender and family SES among NCEE test takers is theoretically ambiguous. Son preference may make boys more likely to come from affluent families, either through sex selection at birth or through increased parental savings after having a son (Wei and Zhang, 2011). Conversely, if son preference is stronger in less developed areas, boys may be overrepresented in lower-SES households. In addition, during our study period, girls, especially those from rural households, are less likely to attend high school (Zeng et al., 2014; Connelly and Zheng, 2003). As a result, girls who do reach high school and take the NCEE may come from higher-SES families. These competing forces make the relationship an empirical question.

Using CCSS data on household income and parental background, columns (1) to (5) of Appendix Table D1 show that female students come from families with higher income, better-educated parents, and parents who are less likely to hold agricultural hukou.[38] However, girls may be more likely to have siblings due to son preference—through differential stopping behavior after the birth of a boy versus a girl—which could dilute family resources. Column (6) shows that female students do have slightly more siblings on average. Nevertheless, column (7) indicates that family income per child remains higher for female than for male students, suggesting that girls are not more likely to come from disadvantaged households even after accounting for resource dilution. These patterns are consistent with our survey evidence, which shows that cost is not an important factor in college migration decisions for parents of girls.

[38]These regressions include home city-track-year FE and NCEE score ventile-rank FE.

6.3. *Decision-making under Uncertainty*

Another potential explanation for the gender gap in college migration is gender differences in decision-making under uncertainty (see Eckel and Grossman (2008) and Croson and Gneezy (2009) for comprehensive reviews). During our data period, students in most provinces were required to submit their rank-ordered lists of universities before knowing the score or even before taking the exam. Moreover, as discussed in Section 2, the Boston mechanism for college assignment implies that failing to gain admission to one's first-choice university can result in a sharp drop in university quality rather than a smooth transition to a next-best alternative. Such a high-stakes environment with a significant amount of uncertainty may interact with gendered behavioral traits in two ways. First, if female students are more risk-averse, they may prefer to apply to less selective universities with a higher likelihood of admission, in order to avoid a sharp drop in university quality or not being admitted to any university. Second, if female students are less confident when forecasting their eventual scores and ranking (Niederle and Vesterlund, 2007), they may be more inclined to apply to less selective universities.

Being more risk-taking or overconfident may increase the likelihood of listing an out-of-reach university as one's first choice, thereby reducing the probability of first-choice admission. Consistent with this interpretation, Panel (a) of Appendix Figure D1 shows that girls have a higher first-choice admission rate, suggesting a more conservative first choice in the university preference list. Panel (c) further shows that girls also have a higher probability of admission to any college throughout the percentile ranks. If both boys and girls were equally likely to include appropriate "safety schools"—institutions that almost guarantee admission given their academic performance—lower down on their preference lists, differences in first-choice admission would not necessarily generate differences in overall admission to any college. Therefore, the gender difference in admission to any college suggests that boys are less likely to include appropriate safety schools as fallback options.[39] This indicates that the more conservative application strategy of girls, relative to boys, extends beyond just the first choice, which echoes the findings of Saygin (2016) and Delaney and Devereux (2021) in the contexts of Turkey and Ireland, respectively.

However, whether gender differences in application strategy, driven by risk attitudes or

[39] Students submit preference lists in a centralized system, where they can list up to 6–8 universities for each admission tier. Because adding universities is costless, the failure to include an appropriate safety school in the lower part of the list reflects differential application behavior rather than a constraint on the number of universities that can be listed.

confidence, translate into a gender gap in college migration is a distinct question. In fact, as we show below, we find little evidence that they do. First, across percentile ranks, gender differences in first-choice admission, shown in Panel (b) of Appendix Figure D1, differ markedly from the migration gap in Figure 2. While the migration gap is negligible for low-ranking students and most salient around the 80th–90th percentiles, the gap in first-choice admission is most pronounced for those in the bottom (below 20th percentile) and the middle (40th–60th percentiles) of the distribution, fluctuating around zero for high-ranking students. Similarly, across percentiles, the pattern of gender gap in admission to any college, shown in Panel (d) of Appendix Figure D1, also does not align with the migration gap. These patterns suggest that, although gendered behavioral traits may shape application strategies and admission outcomes, they are unlikely to be an important driver of the observed gender differences in college migration.

Second, we examine more directly the role of gendered preferences in explaining the gender gap in college migration. To do so, we exploit province-by-year variation in the information available to applicants at the time of college application. We classify application regimes into three categories: low-information regimes, in which applications are submitted before the NCEE; medium-information regimes, in which applications are submitted after the examination but before scores are released; and high-information regimes, in which applications are submitted after exam scores are known. Between 1999 and 2003, 13 provinces altered their admission policies, allowing us to compare the gender gap within the same province under different information regimes over time.

Moving from low- to high-information regimes substantially increases the information available when applications are submitted, reducing the uncertainty inherent in the application process.[40] Moreover, because applicants in high-information regimes know both their exam scores and rankings, there is limited scope for gender differences in self-confidence to influence application behavior.[41] Consequently, if gender differences in risk aversion or self-confidence are important drivers of the migration gap, the gap should narrow under high information regimes, where much of the uncertainty surrounding admissions outcomes has been resolved.

Column (1) of Appendix Table D2 replicates the baseline specification using the sample of the 13 provinces with changes in application timing. Estimates are consistent with those from

[40] Uncertainty is not completely eliminated even in high-information regimes, since admission cutoffs for individual universities remain unknown ex ante.

[41] In high-information regimes, provinces typically publish the number of students scoring above each score threshold (e.g., 600, 610, 620), allowing applicants to infer their relative standing.

the main sample. In columns (2) and (3), we interact the *Female* indicator with two dummy variables: *High Information* (equal to 1 if the province-year is in a high-information regime, 0 otherwise) and *Low Information* (equal to 1 if the province-year is in a low-information regime, 0 otherwise), with medium-information regimes serving as the omitted category. We include a full set of *Female* × *Province* dummies to account for time-invariant provincial differences in the gender gap. Column (3) further includes *Female* × *Year* dummies to control flexibly for the time-specific gender gap. If overconfidence or differential risk aversion are important drivers of the migration gender gap, we would expect a positive and significant coefficient on *Female* × *High Information*, indicating a smaller gender gap, and a negative and significant coefficient for *Female* × *Low Information*. We observe neither pattern in columns (2) and (3).

Taken together, these findings indicate that, while gender differences in risk aversion and self-confidence may shape application behavior more generally, causing female students to pursue more conservative application strategies, they are unlikely to account for the observed gender gap in college migration.

7. Conclusion

In this paper, we study an important yet underexplored source of gender disparities in labor market outcomes that emerges early in the life cycle: gender differences in college migration. Using administrative data covering the universe of national college entrance exam takers in China between 1999 and 2003, we document a substantial gender gap in college migration. Conditional on exam performance—the only criterion for university admission—female students are 4.5 percentage points less likely than otherwise comparable male students to enroll in an out-of-province university. This gap is particularly pronounced in areas with more traditional gender-role attitudes.

To assess the consequences of these differences, we exploit policy-driven university expansions together with pre-policy cross-province admission quotas to estimate the causal returns to college migration. Our IV estimates indicate that college migration roughly doubles the likelihood of attending a prestigious university and increases expected wages by about 20%, with similar returns for male and female students. Consequently, if girls migrated for college at the same rate as equally qualified boys, the gender gap in university prestige would narrow by

21.9% and the gap in expected wages by 11.5%. We also document a concerning pattern that the places where out-migration for college yields the greatest returns are precisely those where the gender gap in college migration is largest, adding a spatial dimension to the debate on the misallocation of talent (Hsieh et al., 2019).

Our findings demonstrate that gendered spatial choices can contribute to disparities in labor market outcomes between men and women, even among college-educated individuals. Universities differ greatly in their labor market returns due to both institutional quality and location, yet social norms and family expectations can lead equally qualified female and male students to make different college migration choices, with lasting consequences. This perspective complements existing explanations of the gender wage gap that emphasize gender differences in commuting distances and job-search radii (Petrongolo and Ronchi, 2020; Le Barbanchon, Rathelot and Roulet, 2021; Caldwell and Danieli, 2024). We show that gendered access to spatial opportunity may begin to emerge much earlier, at the college choice stage.

References

Abdulkadiroğlu, Atila, and Tayfun Sönmez. 2003. "School Choice: A Mechanism Design Approach." *American Economic Review*, 93(3): 729–747.

Almond, Douglas, Hongbin Li, and Shuang Zhang. 2019. "Land reform and sex selection in China." *Journal of Political Economy*, 127(2): 560–585.

Altonji, Joseph G., John Eric Humphries, Yagmur Yuksel, and Ling Zhong. 2025. "Decomposing trends in the gender gap for highly educated workers." National Bureau of Economic Research.

Attanasio, Orazio P., and Katja M. Kaufmann. 2017. "Education choices and returns on the labor and marriage markets: Evidence from data on subjective expectations." *Journal of Economic Behavior & Organization*, 140: 35–55.

Azmat, Ghazala, and Rosa Ferrer. 2017. "Gender gaps in performance: Evidence from young lawyers." *Journal of Political Economy*, 125(5): 1306–1355.

Barth, Erling, Sari Pekkala Kerr, and Claudia Olivetti. 2021. "The dynamics of gender earnings differentials: Evidence from establishment data." *European Economic Review*, 134: 103713.

Bau, Natalie. 2021. "Can policy change culture? Government pension plans and traditional kinship practices." *American Economic Review*, 111(6): 1880–1917.

Benson, Alan, Danielle Li, and Kelly Shue. 2026. ""Potential" and the gender promotion gap." *American Economic Review*, 116(2): 375–417.

Bertrand, Marianne. 2011. "New perspectives on gender." In *Handbook of Labor Economics*. Vol. 4, 1543–1590. Elsevier.

Bertrand, Marianne. 2018. "Coase lecture – The glass ceiling." *Economica*, 85(338): 205–231.

Bertrand, Marianne, Claudia Goldin, and Lawrence F. Katz. 2010. "Dynamics of the gender gap for young professionals in the financial and corporate sectors." *American Economic Journal: Applied Economics*, 2(3): 228–255.

Bertrand, Marianne, Emir Kamenica, and Jessica Pan. 2015. "Gender identity and relative income within households." *The Quarterly Journal of Economics*, 130(2): 571–614.

Bertrand, Marianne, Patricia Cortes, Claudia Olivetti, and Jessica Pan. 2021. "Social norms, labour market opportunities, and the marriage gap between skilled and unskilled women." *The Review of Economic Studies*, 88(4): 1936–1978.

Biasi, Barbara, and Heather Sarsons. 2022. "Flexible wages, bargaining, and the gender gap." *The Quarterly Journal of Economics*, 137(1): 215–266.

Blanchard, Olivier Jean, and Lawrence F. Katz. 1992. "Regional evolutions." *Brookings papers on economic activity*, 1992(1): 1–75.

Blau, Francine D., and Lawrence M. Kahn. 2017. "The gender wage gap: Extent, trends, and explanations." *Journal of Economic Literature*, 55(3): 789–865.

Boelmann, Barbara. 2024. "Women's missing mobility and the gender gap in higher education: Evidence from Germany's university expansion." ECONtribute Discussion Paper.

Borjas, George J. 1987. "Self-selection and the earnings of immigrants." *American Economic Review*, 77(4): 531–553.

Borker, Girija. 2021. "Safety first: Perceived risk of street harassment and educational choices of women." World Bank Policy Research Working Paper 9731.

Borusyak, Kirill, Peter Hull, and Xavier Jaravel. 2022. "Quasi-experimental shift-share research designs." *The Review of Economic Studies*, 89(1): 181–213.

Bound, John, and Harry J. Holzer. 2000. "Demand shifts, population adjustments, and labor market outcomes during the 1980s." *Journal of Labor Economics*, 18(1): 20–54.

Bursztyn, Leonardo, Thomas Fujiwara, and Amanda Pallais. 2017. "'Acting wife': Marriage market incentives and labor market investments." *American Economic Review*, 107(11): 3288–3319.

Caldwell, Sydnee, and Oren Danieli. 2024. "Outside options in the labour market." *Review of Economic Studies*, 91(6): 3286–3315.

Card, David, Ana Rute Cardoso, and Patrick Kline. 2016. "Bargaining, sorting, and the gender wage gap: Quantifying the impact of firms on the relative pay of women." *The Quarterly Journal of Economics*, 131(2): 633–686.

Chen, Yan, and Onur Kesten. 2017. "Chinese college admissions and school choice reforms: A theoretical analysis." *Journal of Political Economy*, 125(1): 99–139.

Chetty, Raj, David J. Deming, and John N. Friedman. 2026. "Diversifying society's leaders? The determinants and causal effects of admission to highly selective private colleges." *The Quarterly Journal of Economics*, 141(1): 51–145.

Chetty, Raj, John N. Friedman, Emmanuel Saez, Nicholas Turner, and Danny Yagan. 2020. "Income segregation and intergenerational mobility across colleges in the United States." *The Quarterly Journal of Economics*, 135(3): 1567–1633.

Che, Yi, and Lei Zhang. 2018. "Human capital, technology adoption and firm performance: Impacts of China's higher education expansion in the late 1990s." *The Economic Journal*, 128(614): 2282–2320.

Choi, Eleanor Jawon, and Jisoo Hwang. 2015. "Child gender and parental inputs: No more son preference in Korea?" *American Economic Review*, 105(5): 638–643.

Connelly, Rachel, and Zhenzhen Zheng. 2003. "Determinants of school enrollment and completion of 10 to 18 year olds in China." *Economics of Education Review*, 22(4): 379–388.

Cortés, Patricia, and Jessica Pan. 2019. "When time binds: Substitutes for household production, returns to working long hours, and the skilled gender wage gap." *Journal of Labor Economics*, 37(2): 351–398.

Cortés, Patricia, and Jessica Pan. 2023. "Children and the remaining gender gaps in the labor market." *Journal of Economic Literature*, 61(4): 1359–1409.

Cortés, Patricia, Jessica Pan, Laura Pilossoph, Ernesto Reuben, and Basit Zafar. 2023. "Gender differences in job search and the earnings gap: Evidence from the field and lab." *The Quarterly Journal of Economics*, 138(4): 2069–2126.

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

Dahl, Gordon B. 2002. "Mobility and the return to education: Testing a Roy model with multiple markets." *Econometrica*, 70(6): 2367–2420.

Delaney, Judith M, and Paul J Devereux. 2021. "Gender differences in college applications: Aspiration and risk management." *Economics of Education Review*, 80: 102077.

Delavande, Adeline, and Basit Zafar. 2019. "University choice: The role of expected earnings, nonpecuniary outcomes, and financial constraints." *Journal of Political Economy*, 127(5): 2343–2393.

Diamond, Rebecca. 2016. "The determinants and welfare implications of US workers' diverging location choices by skill: 1980–2000." *American Economic Review*, 106(3): 479–524.

Dustmann, Christian, and Albrecht Glitz. 2011. "Migration and education." In *Handbook of the Economics of Education*. Vol. 4, 327–439. Elsevier.

Eckel, Catherine C., and Philip J. Grossman. 2008. "Men, women and risk aversion: Experimental evidence." *Handbook of Experimental Economics Results*, 1: 1061–1073.

Edlund, Lena. 1999. "Son preference, sex ratios, and marriage patterns." *Journal of Political Economy*, 107(6): 1275–1304.

Fabre, Anaıs. 2023. "The geography of higher education and spatial inequalities." Working paper.

Fortin, Nicole M. 2005. "Gender role attitudes and the labour-market outcomes of women across OECD countries." *Oxford Review of Economic Policy*, 21(3): 416–438.

Fu, Chao, Junjie Guo, Adam J. Smith, and Alan Sorensen. 2022. "Students' heterogeneous preferences and the uneven spatial distribution of colleges." *Journal of Monetary Economics*, 129: 49–64.

Goldin, Claudia. 2006. "The quiet revolution that transformed women's employment, education, and family." *American Economic Review*, 96(2): 1–21.

Goldin, Claudia. 2014. "A grand gender convergence: Its last chapter." *American Economic Review*, 104(4): 1091–1119.

Goldin, Claudia, and Cecilia Rouse. 2000. "Orchestrating impartiality: The impact of "blind" auditions on female musicians." *American Economic Review*, 90(4): 715–741.

Goldin, Claudia, and Lawrence F. Katz. 2016. "A most egalitarian profession: Pharmacy and the evolution of a family-friendly occupation." *Journal of Labor Economics*, 34(3): 705–746.

Goldin, Claudia, Lawrence F. Katz, and Ilyana Kuziemko. 2006. "The homecoming of American college women: The reversal of the college gender gap." *Journal of Economic Perspectives*, 20(4): 133–156.

Goldin, Claudia, Sari Pekkala Kerr, Claudia Olivetti, and Erling Barth. 2017. "The expanding gender earnings gap: Evidence from the LEHD-2000 census." *American Economic Review*, 107(5): 110–114.

Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift. 2020. "Bartik instruments: What, when, why, and how." *American Economic Review*, 110(8): 2586–2624.

Hsieh, Chang-Tai, Erik Hurst, Charles I. Jones, and Peter J. Klenow. 2019. "The allocation of talent and U.S. economic growth." *Econometrica*, 87(5): 1439–1474.

Hu, Naiyuan, and Lin Ma. 2025. "Educational migration in China." Working paper.

Ishimaru, Shoya. 2025. "Geographic mobility of youth and spatial gaps in local college and labor market opportunities." *Journal of Labor Economics*, 43(4): 1251–1294.

Jayachandran, Seema. 2015. "The roots of gender inequality in developing countries." *Annual Review of Economics*, 7(1): 63–88.

Jensen, Robert. 2010. "The (perceived) returns to education and the demand for schooling." *The Quarterly Journal of Economics*, 125(2): 515–548.

Jia, Ning, Raven Molloy, Christopher Smith, and Abigail Wozniak. 2023. "The economics of internal migration: Advances and policy questions." *Journal of Economic Literature*, 61(1): 144–180.

Jia, Ruixue, and Hongbin Li. 2021. "Just above the exam cutoff score: Elite college admission and wages in China." *Journal of Public Economics*, 196: 104371.

Kirkebøen, Lars J., Edwin Leuven, and Magne Mogstad. 2016. "Field of study, earnings, and self-selection." *The Quarterly Journal of Economics*, 131(3): 1057–1111.

Le Barbanchon, Thomas, Roland Rathelot, and Alexandra Roulet. 2021. "Gender differences in job search: Trading off commute against wage." *Quarterly Journal of Economics*, 136(1): 381–426.

Li, Hongbin, and Xinyao Qiu. 2025. "Heuristics in self-evaluation: Evidence from the centralized college admission system in China." *Review of Economics and Statistics*, 1–10.

Malamud, Ofer, and Abigail K. Wozniak. 2012. "The impact of college education on geographic mobility." *Journal of Human Resources*, 47(4): 913–950.

Mas, Alexandre, and Amanda Pallais. 2017. "Valuing alternative work arrangements." *American Economic Review*, 107(12): 3722–3759.

Meng, Xin. 2012. "Labor market outcomes and reforms in China." *Journal of Economic Perspectives*, 26(4): 75–102.

Molloy, Raven, Christopher L. Smith, and Abigail Wozniak. 2011. "Internal migration in the United States." *Journal of Economic Perspectives*, 25(3): 173–196.

Mountjoy, Jack. 2026. "Marginal returns to public universities." *The Quarterly Journal of Economics*, 141(1): 429–497.

Mountjoy, Jack, and Brent R. Hickman. 2021. "The returns to college (s): Relative value-added and match effects in higher education." National Bureau of Economic Research.

Niederle, Muriel, and Lise Vesterlund. 2007. "Do women shy away from competition? Do men compete too much?" *The Quarterly Journal of Economics*, 122(3): 1067–1101.

OECD. 2024. *Education at a Glance 2024: OECD Indicators.* OECD Publishing, Paris, https://doi.org/10.1787/c00cad36-en.

Olivetti, Claudia, Jessica Pan, and Barbara Petrongolo. 2024. "The evolution of gender in the labor market." In *Handbook of Labor Economics*. Vol. 5, 619–677. Elsevier.

Pan, Jessica. 2015. "Gender segregation in occupations: The role of tipping and social interactions." *Journal of Labor Economics*, 33(2): 365–408.

Pathak, Parag A, and Tayfun Sönmez. 2008. "Leveling the playing field: Sincere and sophisticated players in the Boston mechanism." *American Economic Review*, 98(4): 1636–1652.

Petrongolo, Barbara, and Maddalena Ronchi. 2020. "Gender gaps and the structure of local labor markets." *Labour Economics*, 64: 101819.

Qian, Nancy. 2008. "Missing women and the price of tea in China: The effect of sex-specific earnings on sex imbalance." *The Quarterly Journal of Economics*, 123(3): 1251–1285.

Roussille, Nina. 2024. "The role of the ask gap in gender pay inequality." *The Quarterly Journal of Economics*, 139(3): 1557–1610.

Sarsons, Heather. 2017. "Interpreting signals in the labor market: Evidence from medical referrals." Working paper.

Saygin, Perihan Ozge. 2016. "Gender differences in preferences for taking risk in college applications." *Economics of Education Review*, 52: 120–133.

Schmutte, Ian M. 2015. "Job referral networks and the determination of earnings in local labor markets." *Journal of Labor Economics*, 33(1): 1–32.

Sloane, Carolyn M., Erik G. Hurst, and Dan A. Black. 2021. "College majors, occupations, and the gender wage gap." *Journal of Economic Perspectives*, 35(4): 223–248.

Turner, Sarah E., and William G. Bowen. 1999. "Choice of major: The changing (unchanging) gender gap." *ILR Review*, 52(2): 289–313.

Wasserman, Melanie. 2023. "Hours constraints, occupational choice, and gender: Evidence from medical residents." *The Review of Economic Studies*, 90(3): 1535–1568.

Wei, Shang-Jin, and Xiaobo Zhang. 2011. "The competitive saving motive: Evidence from rising sex ratios and savings rates in China." *Journal of Political Economy*, 119(3): 511–564.

Wiswall, Matthew, and Basit Zafar. 2015. "Determinants of college major choice: Identification using an information experiment." *The Review of Economic Studies*, 82(2): 791–824.

Wiswall, Matthew, and Basit Zafar. 2018. "Preference for the workplace, investment in human capital, and gender." *The Quarterly Journal of Economics*, 133(1): 457–507.

Wiswall, Matthew, and Basit Zafar. 2021. "Human capital investments and expectations about career and family." *Journal of Political Economy*, 129(5): 1361–1424.

Wozniak, Abigail. 2010. "Are college graduates more responsive to distant labor market opportunities?" *Journal of Human Resources*, 45(4): 944–970.

Yang, Y. Angela. 2024. "On the spatial allocation of college seats: Human capital production and the distribution of skilled labor." Available at SSRN: https://ssrn.com/abstract=4966963.

Young, Alwyn. 2000. "The razor's edge: Distortions and incremental reform in the People's Republic of China." *The Quarterly Journal of Economics*, 115(4): 1091–1135.

Zafar, Basit. 2011. "How do college students form expectations?" *Journal of Labor Economics*, 29(2): 301–348.

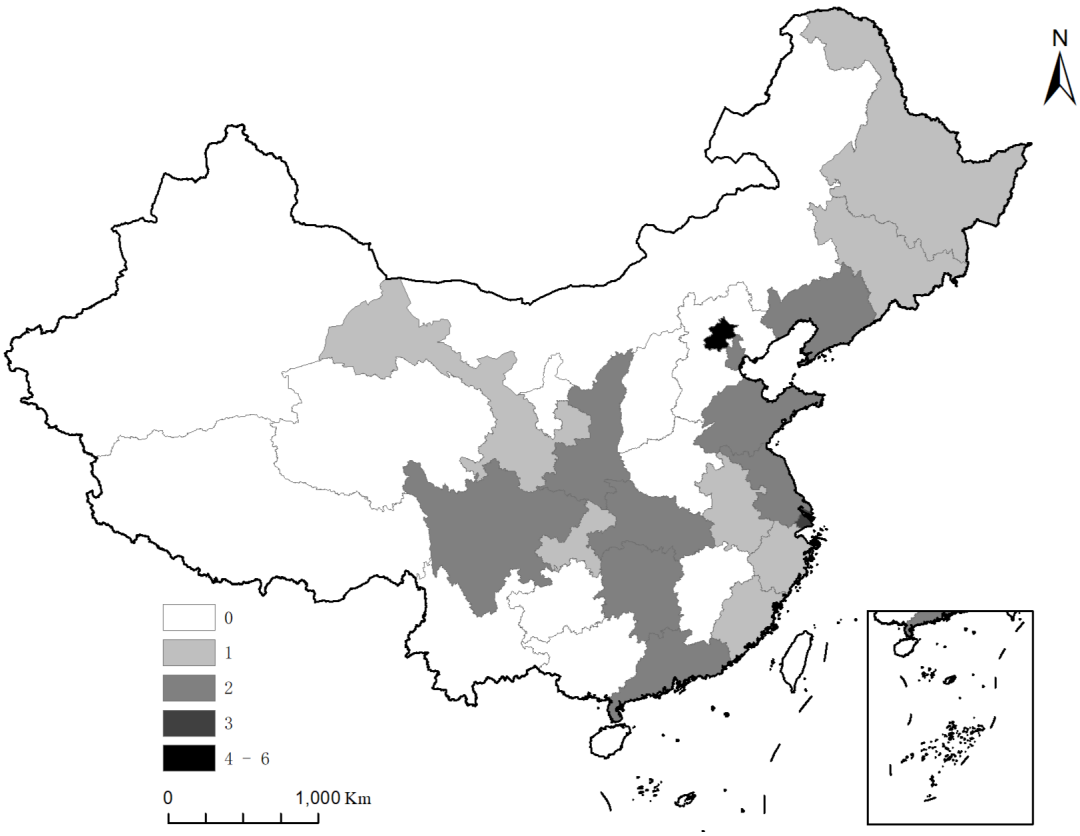
Zafar, Basit. 2013. "College major choice and the gender gap." *Journal of Human Resources*, 48(3): 545–595.

Zeng, Junxia, Xiaopeng Pang, Linxiu Zhang, Alexis Medina, and Scott Rozelle. 2014. "Gender inequality in education in China: A meta-regression analysis." *Contemporary Economic Policy*, 32(2): 474–491.

Zhang, Han, Donald Patton, and Martin Kenney. 2013. "Building global-class universities: Assessing the impact of the 985 Project." *Research Policy*, 42(3): 765–775.

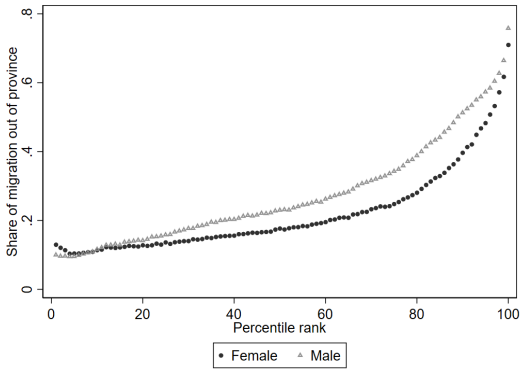
Zimmerman, Seth D. 2019. "Elite colleges and upward mobility to top jobs and top incomes." *American Economic Review*, 109(1): 1–47.

Figure 1: Spatial Distribution of Elite Universities

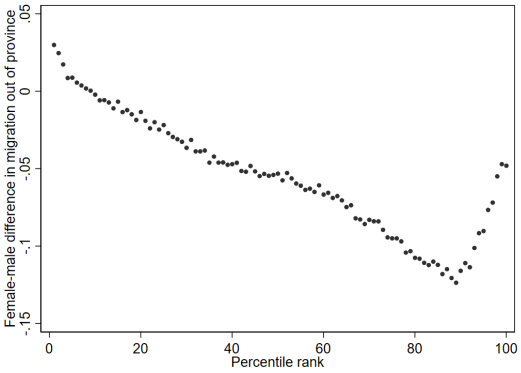


Notes: This figure plots the provincial count of Project 985 (elite) universities in China as of 2003. Darker shades of grey indicate a higher count. Provinces shaded in solid white indicate zero Project 985 universities.

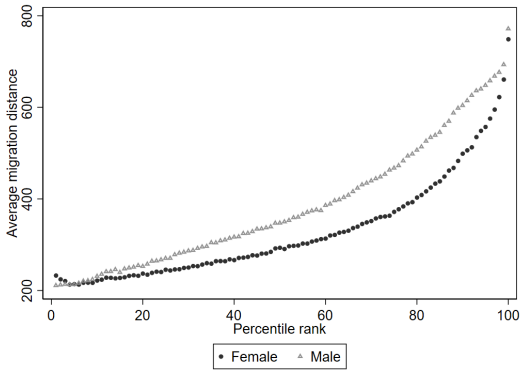
Figure 2: Gender Differences in College Migration by Exam Percentile Rank



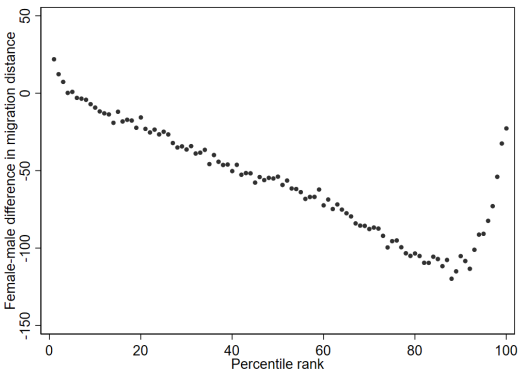
(a) Share of attending college out of province



(b) Female-male difference in share of attending college out of province



(c) Average distance between home and college cities



(d) Female-male difference in average distance between home and college cities

Notes: This figure plots college migration outcomes and gender differences across percentile ranks. The sample consists of all students admitted to college. Panels (a) and (c) report the average outcomes by gender, with grey hollow triangles denoting male students and black solid circles denoting female students. Panel (a) plots the share of students migrating out of their home province, while Panel (c) plots the average distance (in kilometers) between the home city and the college city. Panels (b) and (d) report the gender difference, calculated as the female mean minus the male mean, for the outcomes in Panels (a) and (c), respectively. In Panels (b) and (d), black solid circles represent the difference in means between female and male students within each bin. In all panels, the x-axis represents the student's percentile rank, divided into 100 integer bins.

Table 1: Summary Statistics

	(1)	(2)	(3)	(4)	(5)	(6)
	Full sample		Female		Male	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Female	0.44	0.50	1.00	0.00	0.00	0.00
Out of province	0.25	0.43	0.22	0.41	0.28	0.45
Migration distance	360.42	497.38	325.18	472.66	387.87	514.14
College in provincial capital	0.61	0.49	0.59	0.49	0.62	0.49
College in top 30 GDP cities	0.49	0.50	0.50	0.50	0.49	0.50
College in top 10 GDP cities	0.26	0.44	0.26	0.44	0.26	0.44
Science (vs. Humanities)	0.72	0.45	0.61	0.49	0.80	0.40
Percentile rank	50.00	28.87	49.33	28.55	50.53	29.10
Agricultural hukou	0.49	0.50	0.44	0.50	0.54	0.50
Age	18.71	1.14	18.58	1.08	18.81	1.17
Ethnic minority	0.07	0.25	0.07	0.26	0.07	0.25
985 university	0.05	0.23	0.04	0.20	0.06	0.24
Observations	13,524,632		5,922,069		7,602,563	

Notes: This table reports summary statistics for the individual-level variables in the baseline sample. Columns (1)–(2) report the means and standard deviations for the full sample, while columns (3)–(4) and columns (5)–(6) report statistics for the female and male subsamples, respectively. *Female* is an indicator equal to 1 if the student is female. *Out of province* is an indicator equal to 1 if the student's university is located in a different province from their hometown. *Migration distance* represents the straight-line distance (in kilometers) between the hometown city and the university city. *College in provincial capital* is an indicator equal to 1 if the student's university is located in a provincial capital city or one of the four direct-administered municipalities (Beijing, Shanghai, Tianjin, and Chongqing). *College in top 30 GDP cities* and *College in top 10 GDP cities* are indicators equal to 1 if the university is located in one of the top 30 or top 10 cities, respectively, ranked by total GDP in 2000. *Science (vs. Humanities)* is an indicator equal to 1 for students in the science track, and 0 for those in the humanities track. *Percentile rank* denotes the student's percentile ranking within their province-track-year cohort based on the NCEE score. *Agricultural hukou* is an indicator for holding an agricultural hukou. *Age* represents the student's age at the time of taking the NCEE. *Ethnic minority* is an indicator equal to 1 if the student belongs to a non-Han ethnic group. *985 university* is an indicator equal to 1 if the university is part of Project 985 (a group of elite research universities). The number of observations is reported in the last row.

Table 2: Gender Gap in College Migration: Baseline Results

	(1)	(2)	(3)	(4)
Dependent variable:	Out of province			
Female	-0.051***	-0.045***	-0.043***	-0.045***
	(0.001)	(0.001)	(0.001)	(0.001)
Percentile rank		0.004***		
		(0.000)		
County-track-year FE	Yes	Yes	Yes	Yes
Percentile rank FE			Yes	Yes
Demographic characteristics				Yes
Dependent mean	0.251	0.251	0.251	0.251
Observations	13,524,632	13,524,632	13,524,632	13,524,632

Notes: This table reports gender differences in out-of-province college migration for the sample of admitted students. The dependent variable equals 1 if a student's admitted college is located in a different province from their home province, and 0 otherwise. The independent variable of interest is an indicator for being female. Column (1) controls for county-track-year fixed effects. Column (2) adds a continuous control for the student's percentile rank (calculated within each province-track-year cohort). Column (3) replaces the continuous rank control with fixed effects for each of the 100 integer percentile bins. Column (4) adds student demographic characteristics, including agricultural hukou dummy, ethnic minority dummy, and age dummies. Robust standard errors, clustered at the home county level, are reported in parentheses. *** Significant at the 1 percent level; ** significant at the 5 percent level; * significant at the 10 percent level.

Table 3: Gender Gap Using Alternative College Migration Measures

Dependent variable:	(1) Migration distance	(2) To provincial capital	(3) To top 30 GDP cities	(4) To top 10 GDP cities
Female	-45.155***	-0.042***	-0.025***	-0.014***
	(0.669)	(0.001)	(0.001)	(0.001)
County-track-year FE	Yes	Yes	Yes	Yes
Percentile rank FE	Yes	Yes	Yes	Yes
Demographic characteristics	Yes	Yes	Yes	Yes
Dependent mean	360.424	0.479	0.391	0.180
Observations	13,524,632	7,755,958	10,027,874	11,843,492
Sample	Baseline	Students within province and hometown not in provincial capital	Exclude students from top 30 GDP cities	Exclude students from top 10 GDP cities

Notes: This table reports gender differences in college migration using alternative migration measures. Column (1) uses the baseline sample, with migration distance, measured in kilometers between the home city and the college city, as the dependent variable. Column (2) restricts the sample to students attending university within their home province, excluding those from provincial capitals or direct-administered municipalities; the dependent variable is an indicator equal to 1 if the student attends a university in the provincial capital. Columns (3) and (4) exclude students from the top 30 and top 10 GDP cities, respectively, ranked by total GDP in 2000. The dependent variables are indicators for attending a university in one of the top 30 GDP cities in column (3) and one of the top 10 GDP cities in column (4). The independent variable of interest in all columns is an indicator for being female. Demographic characteristics include agricultural hukou dummy, ethnic minority dummy, and age dummies. Robust standard errors, clustered at the home county level, are reported in parentheses. *** Significant at the 1 percent level; ** significant at the 5 percent level; * significant at the 10 percent level.

Table 4: Heterogeneity by Hukou Status and Local Gender Norms

	(1)	(2)	(3)	(4)
Dependent variable:		Out of province		
Female × Agricultural hukou	-0.028***			
	(0.001)			
Female × Gender norm		-0.003***	-0.007***	-0.006***
		(0.001)	(0.002)	(0.002)
Female	-0.031***	-0.045***	-0.039***	-0.039***
	(0.001)	(0.001)	(0.002)	(0.002)
Gender norm measure		Sex ratio at birth	Gender role	Marriage importance
County-track-year FE	Yes	Yes	Yes	Yes
Percentile rank FE	Yes	Yes	Yes	Yes
Demographic characteristics	Yes	Yes	Yes	Yes
Dependent mean	0.251	0.251	0.182	0.182
Observations	13,524,632	13,524,632	1,137,525	1,137,525
Sample	Baseline	Baseline	CGSS matched	CGSS matched

Notes: This table reports estimates of the heterogeneous effects by hukou status and local gender norms on out-of-province college migration. The dependent variable equals 1 if a student's admitted college is located in a different province from their home province, and 0 otherwise. Column (2) measures gender norms using the male-to-female sex ratio for the population aged 0–4 based on the 2000 China Population Census. Columns (3) and (4) measure norms using the county-level average agreement with the statement "Men should focus on career, while women should focus on family" (column 3) and "Marrying well is more important for women than doing well" (column 4), based on data from the 2010 China General Social Survey (CGSS). Responses in the CGSS are rated on a scale from 1 (strongly disagree) to 5 (strongly agree). All gender norm measures are standardized to have a mean of zero and a standard deviation of one across counties. Demographic characteristics include agricultural hukou dummy, ethnic minority dummy, and age dummies. Robust standard errors, clustered at the home county level, are reported in parentheses. *** Significant at the 1 percent level; ** significant at the 5 percent level; * significant at the 10 percent level.

Table 5: 2SLS Estimates of the Premium to College Migration: Elite University Admission and Expected Wages

	(1)	(2)	(3)	(4)
	Baseline sample		NCEE-CCSS matched	
Sample	Female	Male	Female	Male
Panel A: First Stage				
Dependent variable:	Out of province			
Out-of-province quota	0.267***	0.333***	0.250***	0.333***
	(0.043)	(0.037)	(0.068)	(0.062)
In-province quota	-0.161***	-0.260***	-0.211***	-0.388***
	(0.019)	(0.018)	(0.029)	(0.026)
Dependent mean	0.222	0.282	0.315	0.384
Kleibergen-Paap F-stat	55.467	121.664	28.043	112.716
Panel B: Second Stage				
Dependent variable:	985 university admission		Expected wages	
Out of province	0.102***	0.123***	701.847***	599.760***
	(0.029)	(0.021)	(223.802)	(75.989)
Dependent mean	0.041	0.062	2813.735	3087.909
Year FE	Yes	Yes	Yes	Yes
County-track FE	Yes	Yes	Yes	Yes
Percentile rank FE	Yes	Yes	Yes	Yes
Demographic characteristics	Yes	Yes	Yes	Yes
Observations	5,275,842	6,711,450	432,055	696,935

Notes: This table reports the 2SLS estimates of the premium to college migration. Panel A reports the first-stage estimates, and Panel B reports the second-stage estimates. The sample is restricted to students admitted to college between 2000 and 2003. Columns (1) and (2) use the full 2000-2003 sample, while columns (3) and (4) are restricted to the NCEE-CCSS matched sample. Columns (1) and (3) report estimates for the female subsample, and columns (2) and (4) for the male subsample. In Panel A, the dependent variable equals 1 if a student's admitted college is located in a different province from their home province, and 0 otherwise. The instruments are the predicted out-of-province and in-province admission quotas, constructed as in Equations 2 and 3. In Panel B, the dependent variable is an indicator for admission to a Project 985 university in columns (1) and (2), and expected monthly wages (in RMB) in columns (3) and (4). Expected wages are calculated from the CCSS data as the average monthly wage of the best job offer for graduates from each university (specific to the student's gender and origin), then matched to the NCEE data. Demographic characteristics include agricultural hukou dummy, ethnic minority dummy, and age dummies. Robust standard errors, clustered at the home county level, are reported in parentheses. *** Significant at the 1 percent level; ** significant at the 5 percent level; * significant at the 10 percent level.

Table 6: Migration Premium by Local Availability of Elite Universities

	(1)	(2)	(3)	(4)
	Female		Male	
Sample	Home prov. with 985	Home prov. without 985	Home prov. with 985	Home prov. without 985
Panel A				
Dependent variable:	985 university admission			
Out of province	-0.040	0.283***	-0.013	0.213***
	(0.047)	(0.028)	(0.045)	(0.016)
Dependent mean	0.048	0.025	0.071	0.041
Observations	3,670,737	1,605,105	4,690,719	2,020,731
Panel B				
Dependent variable:	Expected wages			
Out of province	337.262	1,166.889***	424.404***	733.691***
	(430.698)	(170.922)	(101.741)	(76.842)
Dependent mean	2866.999	2683.497	3139.477	2972.225
Observations	306,700	125,355	482,090	214,845
Year FE	Yes	Yes	Yes	Yes
County-track FE	Yes	Yes	Yes	Yes
Percentile rank FE	Yes	Yes	Yes	Yes
Demographic characteristics	Yes	Yes	Yes	Yes

Notes: This table reports second-stage estimates from the instrumental variable regressions, stratified by local availability of Project 985 universities. The sample is restricted to students admitted to college between 2000 and 2003. The independent variable, out-of-province college migration, is instrumented using the shift-share predicted in-province and out-of-province admission quotas. Corresponding first-stage estimates are reported in Appendix Table B2. The dependent variable is an indicator for admission to a Project 985 university in Panel A, and expected monthly wages (in RMB) in Panel B. Expected wages are calculated from the CCSS data as the average monthly wage of the best job offer for graduates from each university (specific to the student's gender and origin), then matched to the NCEE data. Panel A uses the full 2000-2003 sample, while Panel B uses the NCEE-CCSS matched sample. Columns (1) and (2) report estimates for the female subsample, separated by whether the student's home province has at least one Project 985 university. Columns (3) and (4) report estimates for the male subsample, using the same stratification. Demographic characteristics include agricultural hukou dummy, ethnic minority dummy, and age dummies. Robust standard errors, clustered at the home county level, are reported in parentheses. *** Significant at the 1 percent level; ** significant at the 5 percent level; * significant at the 10 percent level.

Table 7: Gender Differences in Parental Preference, Concerns, and Perceived Benefits Regarding College Migration

Dependent variable:	(1) Migration preference	(2) Finding a spouse	(3) Local spouse	(4) Elderly care (financial)	(5) Elderly care (time)	(6) Safety	(7) Cost	(8) Timely assistance	(9) Perceived benefit
Female	-0.054**	-0.111***	0.068**	-0.028*	0.021	0.064***	0.002	0.013	-0.169***
	(0.023)	(0.020)	(0.030)	(0.016)	(0.025)	(0.010)	(0.020)	(0.008)	(0.017)
Expected university-level FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Income-bracket FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
City-track FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Demographic characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,940	2,940	2,940	2,940	2,940	2,940	2,940	2,940	2,940
Dependent mean	0.394	0.482	0.328	0.098	0.540	0.856	0.567	0.910	0.177

Notes: This table reports estimates on gender differences in parental preference for out-of-province college attendance, as well as in perceived concerns and benefits. *Migration preference* equals 1 if the parent ranks an out-of-province university (including universities outside the home province but excluding Beijing and Shanghai, or universities in Beijing or Shanghai) above an in-province university, assuming equal quality across options. *Finding a spouse* equals 1 if the respondent reports being "slightly concerned" or more concerned about whether the child would be able to find a suitable spouse if studying outside the home province. *Local spouse* equals 1 if the respondent agrees, either to some extent or strongly, that the child's future spouse should be from the home province. *Elderly care (financial)* equals 1 if the respondent indicates that their primary expectation regarding the child's elderly-care responsibility is to provide financial support, rather than to spend time accompanying parents or to have no such expectation. *Elderly care (time)* equals 1 if the respondent's primary expectation is for the child to spend time accompanying parents. *Safety*, *Cost*, and *Timely assistance* equal 1 if the respondent reports being "slightly concerned" or more concerned about city safety, the financial costs of out-of-province study, or the family's ability to provide timely assistance or networking support, respectively. *Perceived benefit* equals 1 if the respondent believes that studying outside the province yields greater benefits for the child's own gender than for the opposite gender. The *Female* indicator equals 1 if the child is female. Expected university-level FE are based on parents' assessment of the child's future university level. Income-bracket FE are based on income brackets. City-track FE are based on the child's home city interacted with academic track (science vs. humanities). Demographic characteristics include whether the respondent is the mother, whether the respondent has an agricultural hukou, the respondent's years of education, and whether the child is an only child. Robust standard errors clustered by city are reported in parentheses. *** Significant at the 1 percent level; ** significant at the 5 percent level; * significant at the 10 percent level.

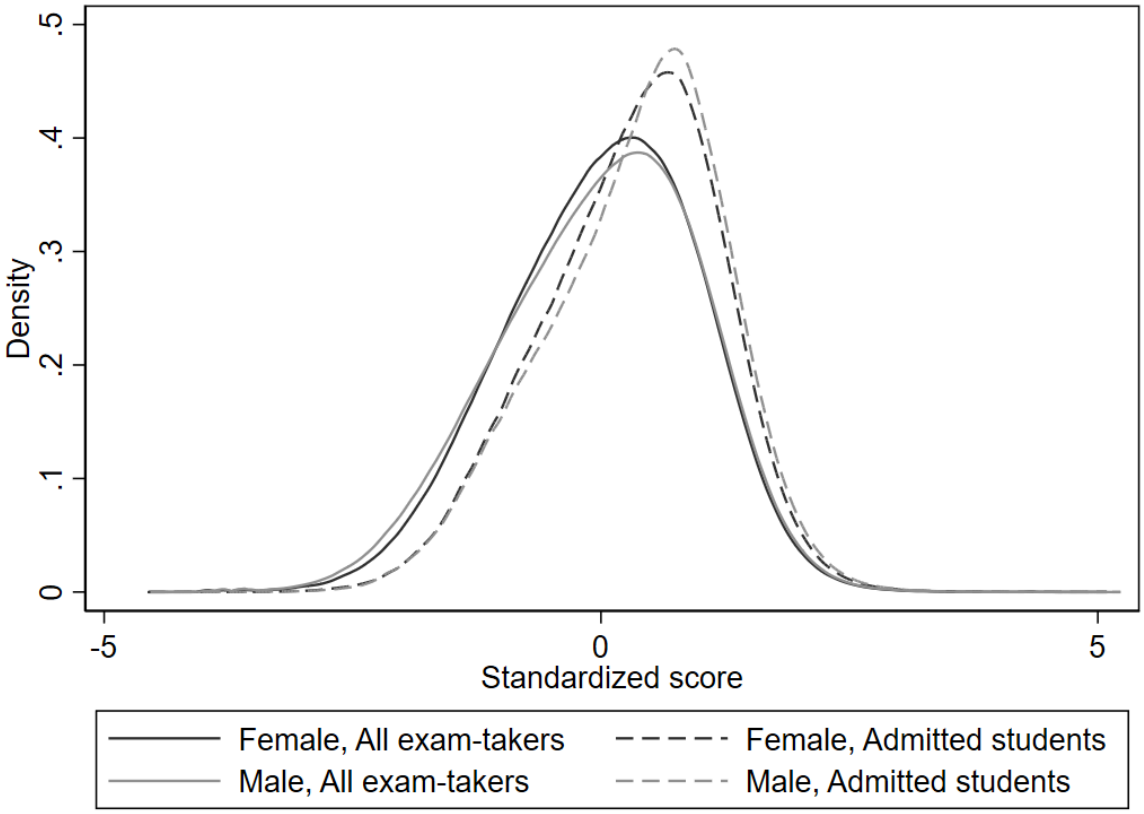
45

Online Appendix for

Gender Differences in College Migration

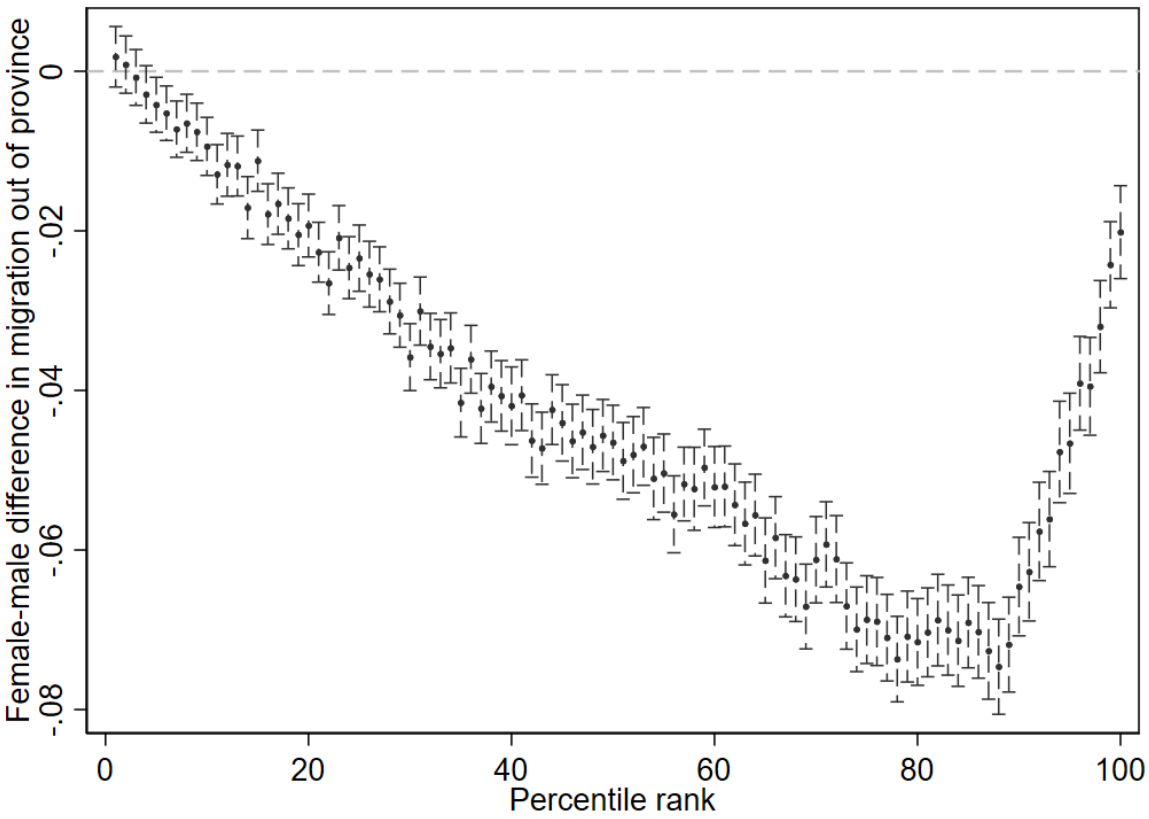
A. Summary and Robustness

Figure A1: Distribution of Standardized NCEE Scores by Gender and Admission Status



Notes: This figure plots the distributions of standardized NCEE scores by gender and admission status. The solid lines represent the sample of all students who took the NCEE, while the dashed lines represent the sample of admitted students. Black lines denote female students, and grey lines denote male students. Scores are standardized to have a mean of zero and a standard deviation of one within each province-track-year cohort, based on the sample of all exam-takers.

Figure A2: Estimated Gender Gap in College Migration across Percentile Ranks



Notes: This figure plots the estimated gender gaps in out-of-province college migration probability across percentile ranks. Each marker shows the coefficient on the female indicator from a separate regression estimated using the subsample of students within that percentile bin. The dependent variable equals 1 if a student's admitted college is located in a different province from their home province, and 0 otherwise. All regressions control for county-track-year fixed effects and demographic characteristics, including agricultural hukou dummy, ethnic minority dummy, and age dummies. The x-axis represents the student's percentile rank within each province-track-year cohort. Vertical dashed lines indicate 95 percent confidence intervals, based on robust standard errors clustered at the home county level.

Table A1: Gender Gap in College Migration by Academic Track

	(1)	(2)
Dependent variable:	Out of province	
Female	-0.048***	-0.029***
	(0.001)	(0.001)
County-year FE	Yes	Yes
Percentile rank FE	Yes	Yes
Demographic characteristics	Yes	Yes
Dependent mean	0.265	0.215
Observations	9,712,588	3,812,044
Sample	Science track	Humanities track

Notes: This table reports gender differences in out-of-province college migration by academic track. The dependent variable equals 1 if a student's admitted college is located in a different province from their home province, and 0 otherwise. The independent variable of interest is an indicator for being female. Column (1) reports estimates for students in the science track, while column (2) reports estimates for students in the humanities track. Demographic characteristics include agricultural hukou dummy, ethnic minority dummy, and age dummies. Robust standard errors, clustered at the home county level, are reported in parentheses. *** Significant at the 1 percent level; ** significant at the 5 percent level; * significant at the 10 percent level.

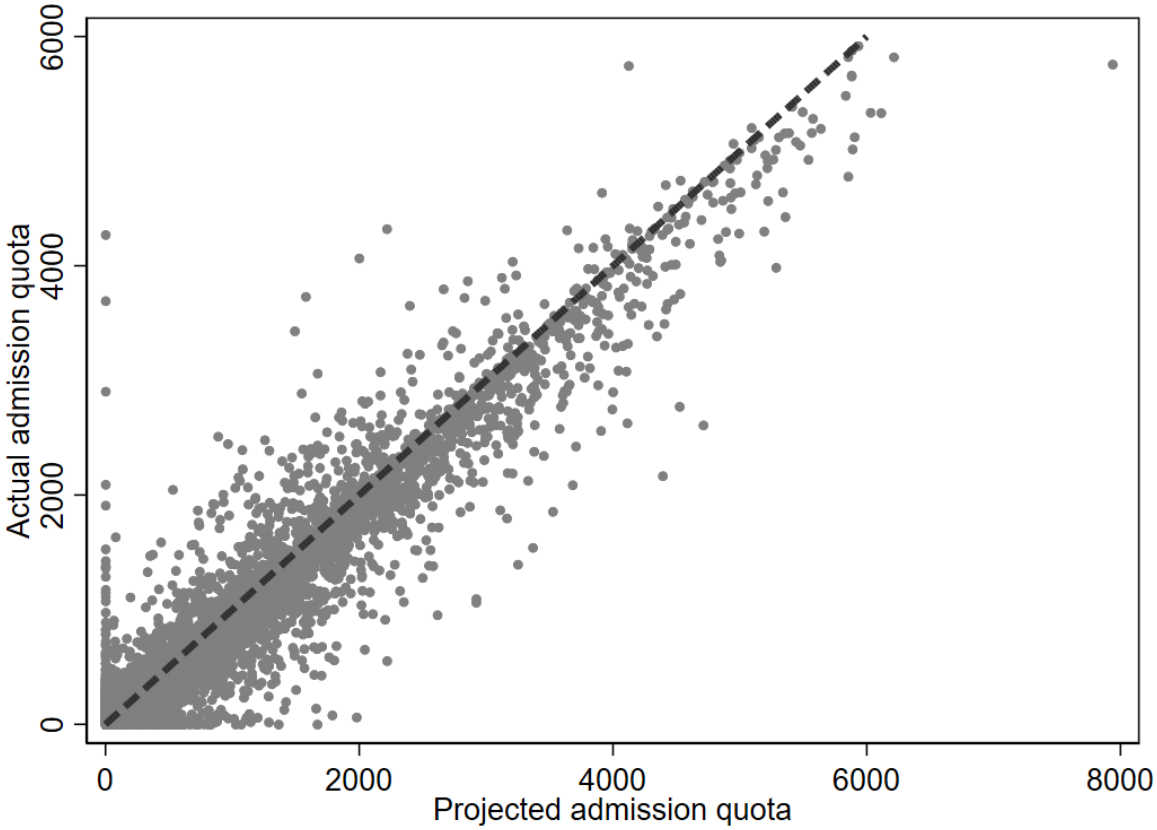
Table A2: Robustness of the Gender Gap in College Migration

	(1)	(2)	(3)	(4)	(5)
Dependent variable:			Out of province		
Female	-0.045***	-0.050***	-0.044***	-0.045***	-0.045***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
County-track-year FE	Yes	Yes	Yes	Yes	Yes
Percentile rank FE	Yes	Yes	Yes	Yes	Yes
Demographic characteristics	Yes	Yes	Yes	Yes	Yes
Dependent mean	0.251	0.264	0.259	0.250	0.247
Observations	13,524,632	8,753,911	13,030,153	13,474,805	12,239,724
Sample	Baseline	First-choice admission	Normal exam type	Exclude HS migrants	Without bonus scores

Notes: This table reports robustness checks for the baseline results. The dependent variable equals 1 if a student's admitted college is located in a different province from their home province, and 0 otherwise. Column (1) reports the baseline estimates. Column (2) restricts the sample to students admitted to their first-choice university. Column (3) restricts the sample to students classified as the "normal" examination type, excluding those in special categories, such as sports and arts. Column (4) excludes high school (pre-college) migrants, defined as students whose NCEE province differs from their high school province. Column (5) excludes students with bonus scores. Demographic characteristics include agricultural hukou dummy, ethnic minority dummy, and age dummies. Robust standard errors, clustered at the home county level, are reported in parentheses. *** Significant at the 1 percent level; ** significant at the 5 percent level; * significant at the 10 percent level.

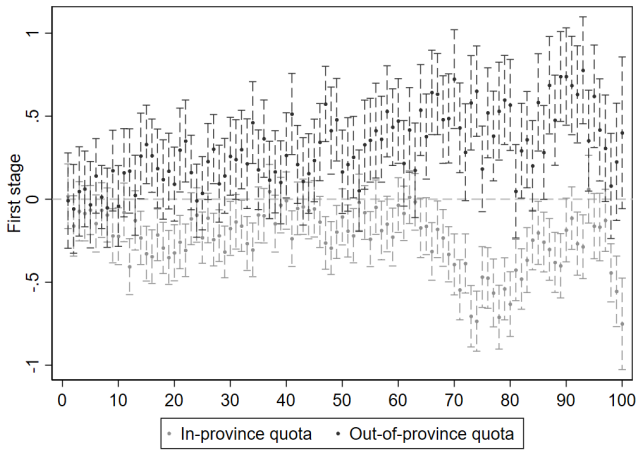
B. Additional Results on Consequences and Counterfactual Analysis

Figure B1: Correlation Between Actual and Projected Admission Quotas

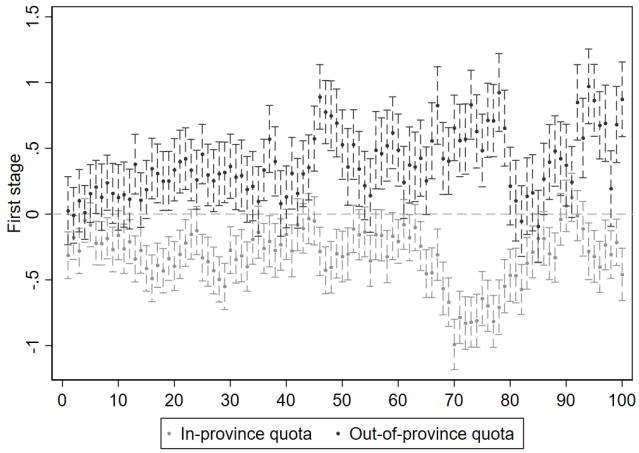


Notes: This figure plots the actual admission quota against the shift-share projected admission quota. Each dot represents an observation at the university-province-track-year level for the years 2000–2003. Observations with an actual admission quota exceeding 6,000 are excluded to remove extreme outliers. The dashed line represents the 45-degree line.

Figure B2: First-Stage Estimates by Percentile Rank and Gender



(a) First stage estimates by percentile rank for females



(b) First stage estimates by percentile rank for males

Notes: This figure plots the first-stage estimates of the instrumental variables across percentile ranks, separately by gender. The sample is restricted to students admitted to college between 2000 and 2003. Panel (a) reports the estimates for the female subsample, and Panel (b) reports the estimates for the male subsample. Each marker represents the coefficient on the corresponding instrument from a separate regression estimated using the subsample of students within that percentile bin. The two instruments are the shift-share predicted in-province and out-of-province admission quotas, shown by grey and black dots, respectively. The x-axis represents the student's percentile rank, divided into 100 integer bins. Vertical dashed lines indicate 95 percent confidence intervals, based on robust standard errors clustered at the home county level. All regressions control for year fixed effects, county-track fixed effects, and student demographic characteristics, including agricultural hukou dummy, ethnic minority dummy, and age dummies.

Table B1: OLS Estimates of College Migration on Student Outcomes

	(1)	(2)	(3)	(4)
	Baseline sample		NCEE-CCSS matched	
Sample	Female	Male	Female	Male
Dependent variable:	985 university admission	985 university admission	Expected wages	Expected wages
Out of province	0.021***	0.017***	376.791***	150.827***
	(0.001)	(0.001)	(16.075)	(10.911)
Year FE	Yes	Yes	Yes	Yes
County-track FE	Yes	Yes	Yes	Yes
Percentile rank FE	Yes	Yes	Yes	Yes
Demographic characteristics	Yes	Yes	Yes	Yes
Dependent mean	0.041	0.062	2813.735	3087.909
Observations	5,275,842	6,711,450	432,055	696,935

Notes: This table reports OLS estimates of the relationship between out-of-province college migration and student outcomes. The sample is restricted to students admitted to college between 2000 and 2003. Columns (1) and (3) report estimates for the female subsample, and columns (2) and (4) report estimates for the male subsample. The dependent variable is an indicator for admission to a Project 985 university in columns (1) and (2), and the expected monthly wages (in RMB) in columns (3) and (4). Expected wages are calculated from the CCSS data as the average monthly wage of the best job offer for graduates from each university (specific to the student's gender and origin), then matched to the NCEE data. Demographic characteristics include agricultural hukou dummy, ethnic minority dummy, and age dummies. Robust standard errors, clustered at the home county level, are reported in parentheses. *** Significant at the 1 percent level; ** significant at the 5 percent level; * significant at the 10 percent level.

Table B2: First-Stage Estimates by Local Availability of Elite Universities

Dependent variable:	(1)	(2)	(3)	(4)
	Out of province			
	Female		Male	
Sample	Home prov. with 985	Home prov. without 985	Home prov. with 985	Home prov. without 985
Panel A: Baseline sample				
Out-of-province quota	0.249***	0.454***	0.278***	0.602***
	(0.044)	(0.105)	(0.045)	(0.053)
In-province quota	-0.040**	-0.304***	-0.144***	-0.474***
	(0.018)	(0.049)	(0.020)	(0.039)
Dependent mean	0.190	0.294	0.250	0.356
Observations	3,670,737	1,605,105	4,690,719	2,020,731
Kleibergen-Paap F-stat	15.831	52.783	31.694	141.190
Panel B: NCEE-CCSS matched sample				
Out-of-province quota	0.222***	0.474***	0.371***	0.511***
	(0.084)	(0.113)	(0.086)	(0.086)
In-province quota	-0.144***	-0.328***	-0.384***	-0.485***
	(0.033)	(0.056)	(0.034)	(0.043)
Dependent mean	0.267	0.433	0.350	0.460
Observations	306,700	125,355	482,090	214,845
Kleibergen-Paap F-stat	10.318	27.174	63.478	90.276
Year FE	Yes	Yes	Yes	Yes
County-track FE	Yes	Yes	Yes	Yes
Percentile rank FE	Yes	Yes	Yes	Yes
Demographic characteristics	Yes	Yes	Yes	Yes

Notes: This table reports the first-stage estimates corresponding to the regressions in Table 6. The dependent variable equals 1 if a student's admitted college is located in a different province from their home province, and 0 otherwise. The instruments are the shift-share predicted out-of-province and in-province admission quotas. Sample restrictions, column specifications, control variables, and standard error clustering are identical to those described in the notes to Table 6. *** Significant at the 1 percent level; ** significant at the 5 percent level; * significant at the 10 percent level.

Table B3: Gender Gap in College Migration by Local Availability of Elite Universities

	(1)	(2)
Dependent variable:	Out of province	
Female	-0.045***	-0.042***
	(0.001)	(0.001)
Female × Home prov. without 985		-0.010***
		(0.001)
County-track-year FE	Yes	Yes
Percentile rank FE	Yes	Yes
Demographic characteristics	Yes	Yes
Dependent mean	0.251	0.251
Observations	13,524,632	13,524,632

Notes: This table reports the baseline gender gap in college migration and its heterogeneity by the availability of Project 985 universities in the student's home province. The dependent variable equals 1 if a student's admitted college is located in a different province from their home province, and 0 otherwise. Column (1) reports the baseline estimates. Column (2) adds an interaction between the female indicator and a dummy variable indicating that the student's home province lacks a Project 985 university. Demographic characteristics include agricultural hukou dummy, ethnic minority dummy, and age dummies. Robust standard errors, clustered at the home county level, are reported in parentheses. *** Significant at the 1 percent level; ** significant at the 5 percent level; * significant at the 10 percent level.

C. Self-designed Survey

Table C1: Association between Migration Preference and Parental Concerns and Perceived Benefits

	(1)	(2)	(3)
Sample	All	Boys	Girls
Dependent variable:	Migration preference		
Finding a spouse	-0.038**	-0.034	-0.043
	(0.016)	(0.028)	(0.045)
Local spouse	-0.106***	-0.086**	-0.135***
	(0.027)	(0.032)	(0.033)
Elderly care (financial)	-0.028	-0.043	0.004
	(0.033)	(0.042)	(0.051)
Elderly care (time)	-0.080***	-0.090***	-0.063***
	(0.022)	(0.032)	(0.020)
Safety	-0.066***	-0.064**	-0.086**
	(0.021)	(0.030)	(0.034)
Cost	-0.025	-0.054*	0.021
	(0.019)	(0.028)	(0.026)
Timely assistance	-0.059**	-0.054	-0.046
	(0.028)	(0.032)	(0.054)
Perceived benefit	0.097***	0.079**	0.166***
	(0.027)	(0.033)	(0.050)
Expected university-level FE	Yes	Yes	Yes
Income-bracket FE	Yes	Yes	Yes
City-track FE	Yes	Yes	Yes
Demographic characteristics	Yes	Yes	Yes
Observations	2,940	1,667	1,273

Notes: This table reports estimates on how parental preference regarding college migration is associated with their concerns and perceived benefits. The definition of all variables can be found in Appendix C.1. Demographic characteristics include whether the respondent is the child's mother, the respondent's hukou status (agricultural vs. non-agricultural), years of education, and whether the child has any siblings. Robust standard errors, clustered by city, are reported in parentheses. *** Significant at the 1 percent level; ** significant at the 5 percent level; * significant at the 10 percent level.

Table C2: Summary Statistics (Survey)

	(1)	(2)	(3)	(4)	(5)	(6)
	All		Boys		Girls	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Panel A: Child						
Expected university level						
985 university	0.163	0.370	0.174	0.379	0.149	0.356
211 university	0.206	0.404	0.199	0.399	0.214	0.411
Tier 1 university	0.429	0.495	0.425	0.495	0.434	0.496
Tier 2 university	0.154	0.361	0.155	0.362	0.152	0.360
Vocational college	0.046	0.209	0.044	0.205	0.048	0.214
No college	0.002	0.045	0.002	0.049	0.002	0.040
Science track	0.662	0.473	0.786	0.410	0.499	0.500
Only child	0.467	0.499	0.456	0.498	0.482	0.500
Panel B: Parent						
Mother	0.589	0.492	0.400	0.490	0.838	0.368
Agricultural hukou	0.475	0.499	0.471	0.499	0.480	0.500
Years of education	14.917	2.697	14.955	2.621	14.867	2.793
Monthly income level (RMB)						
Less than 2000	0.011	0.105	0.010	0.098	0.013	0.115
2000-5000	0.132	0.338	0.113	0.316	0.156	0.363
5000-10000	0.482	0.500	0.478	0.500	0.489	0.500
10000-20000	0.303	0.460	0.328	0.470	0.271	0.445
More than 20000	0.071	0.258	0.072	0.259	0.071	0.256
Migration preference	0.394	0.489	0.416	0.493	0.365	0.482
Concerns about studying out of home province						
Finding a spouse	0.482	0.500	0.541	0.498	0.404	0.491
Safety	0.856	0.351	0.818	0.386	0.907	0.290
Cost	0.567	0.496	0.569	0.495	0.564	0.496
Timely assistance	0.910	0.286	0.900	0.300	0.922	0.268
Local spouse	0.328	0.470	0.303	0.460	0.361	0.480
Elderly care (financial)	0.098	0.297	0.103	0.303	0.092	0.289
Elderly care (time)	0.540	0.498	0.554	0.497	0.523	0.500
Perceived benefit	0.177	0.382	0.256	0.436	0.074	0.262
Observations	2,940		1,667		1,273	

Notes: *Expected university level* is based on parental expectations regarding the university attendance of their child. *985 university* refers to a group of the most prestigious universities designated under Project 985; *211 university* refers to a broader group of highly selective universities designated under Project 211, excluding 985 universities; and *Tier 1 university* refers to remaining first-tier universities, excluding both 985 and 211 universities. The *Science track* indicator takes the value of 1 if the student is in the science track. The *Only child* indicator takes the value of 1 if the respondent has only one child. *Mother* is a dummy variable indicating whether the respondent is the mother of the child. *Agricultural hukou* takes the value of 1 if the respondent has an agricultural hukou. *Years of education* and *Monthly income level* refer to the respondent's education years and monthly income, respectively. For the definitions of the remaining variables regarding migration preference and subjective attitudes, please refer to Appendix C.1.

C.1. Survey Questions on College Choice

Question 1 (Migration preference):

If your child could attend universities of the same quality in the home province, in another province (excluding Beijing and Shanghai), or in Beijing or Shanghai, how would you like your child to choose? Please rank the three options from most preferred to least preferred."

(1) University in the home province

(2) University in another province (excluding Beijing and Shanghai)

(3) Universities in Beijing or Shanghai

> **Variable Construction: Migration preference** is coded as 1 if the parent ranks option (1) below any of the other alternatives (i.e., not ranked first), and 0 otherwise.

Question 2 (Parental concerns):

If your child were to attend university in another province, how concerned would you be about the following factors?

	Very Concerned	Somewhat Concerned	Slightly Concerned	Not Concerned
(a) Safety of the city	☐	☐	☐	☐
(b) Ability to provide timely assistance or networking support to your child	☐	☐	☐	☐
(c) Costs of studying in another province (e.g., travel expenses)	☐	☐	☐	☐
(d) Difficulty in finding a suitable spouse	☐	☐	☐	☐

> **Variable Construction:** For each sub-item, we construct an indicator variable—**Safety**, **Timely assistance**, **Cost**, and **Finding a spouse**—that equals 1 if the parent reports any level of concern (i.e., answers "Very concerned," "Somewhat concerned," or "Slightly concerned"), and 0 otherwise.

Question 3 (Local spouse):

Do you prefer that your child's spouse will be from the same home province?

(1) Strongly Disagree

(2) Disagree to some extent

(3) Neutral

(4) Agree to some extent

(5) Strongly Agree

Variable Construction: Local spouse is an indicator variable equal to 1 if the parent responds "Agree to some extent" or "Strongly Agree" (options 4 and 5), and 0 otherwise.

Question 4 (Elderly care):

What is your primary expectation regarding your child's role in your old-age support?

(1) Financial support
(2) Time and companionship
(3) No expectations; they do not need to take responsibility for my elderly care

Variable Construction: Elderly care (financial) is an indicator variable equal to 1 if the parent chooses option (1), and 0 otherwise. **Elderly care (time)** is an indicator variable equal to 1 if the parent chooses option (2), and 0 otherwise.

Question 5 (Perceived benefit):

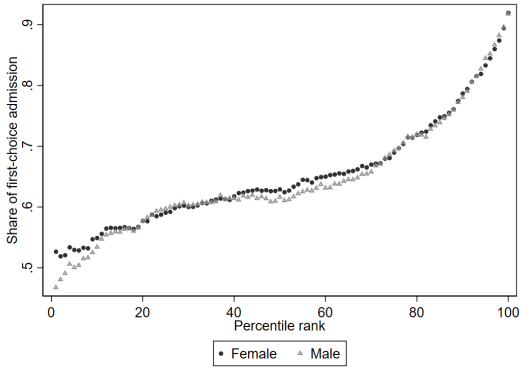
Do you think attending a university outside the home province provides:

(1) Greater benefits for boys
(2) Greater benefits for girls
(3) Similar benefits for both boys and girls

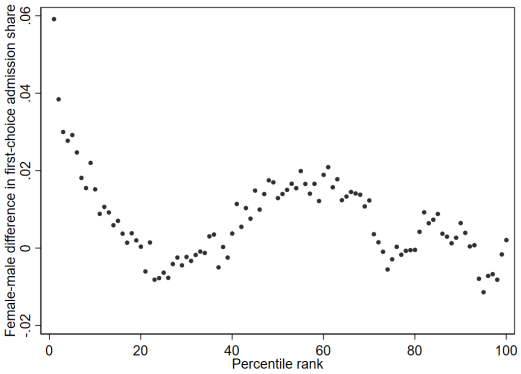
Variable Construction: Perceived benefit is an indicator variable equal to 1 if the parent believes that studying outside the province yields greater benefits for the gender of their own child (i.e., selects option (1) for a son or option (2) for a daughter), and 0 otherwise.

D. Gender Differences in Decision-making under Uncertainty and in Parental SES

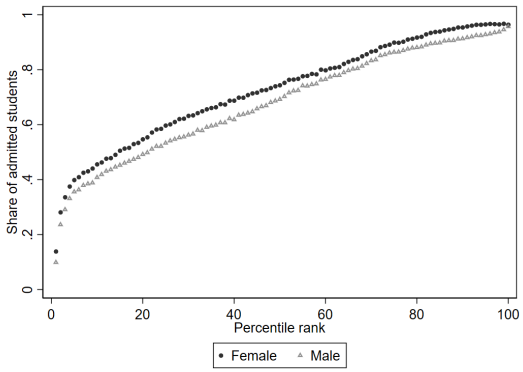
Figure D1: Gender Differences in (First-Choice) College Admission



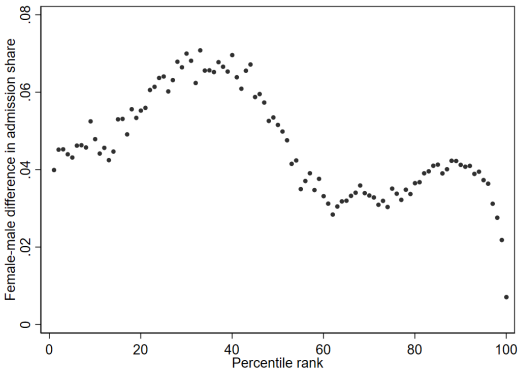
(a) Raw pattern of first-choice admission by gender



(b) Female–male difference in first-choice admission



(c) Raw pattern of overall college admission by gender



(d) Female–male difference in overall college admission

Notes: This figure plots college admission outcomes and their corresponding female–male differences across NCEE score percentile ranks. Panels (a) and (b), based on the admitted students sample, show the raw patterns by gender and the female–male difference (female mean minus male mean) for first-choice college admission. Panels (c) and (d), based on all NCEE exam takers, show the raw patterns and the female–male difference for overall college admission. The x-axis represents the student's percentile rank, divided into 100 integer bins, constructed based on the score distribution of the sample of admitted students. In Panels (a) and (c), grey hollow triangles and black solid circles denote male and female students, respectively.

Table D1: Gender Differences in Family Background (CCSS Data)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variable:	Ln(family income)	Father education	Mother education	Father agricultural hukou	Mother agricultural hukou	Number of siblings	Ln(family income per child)
Female	0.105***	0.480***	0.682***	-0.053***	-0.050***	0.064***	0.069***
	(0.021)	(0.093)	(0.112)	(0.013)	(0.014)	(0.015)	(0.024)
City-track-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ventile rank FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Dependent mean	10.587	10.988	9.961	0.551	0.569	1.757	10.145
Observations	17,311	16,341	16,343	17,311	17,311	17,293	17,293

Notes: This table reports gender differences in family background using survey data from the China College Student Survey (2011–2013). The independent variable of interest is an indicator for being female. The dependent variable is the natural logarithm of total household income in the previous year in column (1); father's and mother's years of schooling in columns (2) and (3), respectively; indicators for the father and mother holding agricultural hukou in columns (4) and (5), respectively; the total number of siblings in column (6); and the natural logarithm of family income per child in column (7). Family income per child is calculated as the total household income divided by the total number of children in the family (the number of siblings plus one). All columns control for hometown city-track-year fixed effects and fixed effects for ventiles of the NCEE score calculated within each province-track-year cohort. We use city-track-year fixed effects instead of county-track-year fixed effects, and ventile rank fixed effects instead of percentile rank fixed effects, due to the limited sample size within each cell. Robust standard errors, clustered at the hometown city level, are reported in parentheses. *** Significant at the 1 percent level; ** significant at the 5 percent level; * significant at the 10 percent level.

Table D2: Gender Gap in College Migration across Admission Information Regimes

	(1)	(2)	(3)
Dependent variable:	Out of province		
Female	-0.043***		
	(0.001)		
Female × Low information		0.004***	0.005***
		(0.001)	(0.001)
Female × High information		0.002	0.001
		(0.001)	(0.002)
County-track-year FE	Yes	Yes	Yes
Percentile rank FE	Yes	Yes	Yes
Demographic characteristics	Yes	Yes	Yes
Female-province FE		Yes	Yes
Female-year FE			Yes
Dependent mean	0.246	0.246	0.246
Observations	5,138,624	5,138,624	5,138,624

Notes: This table reports estimates of the heterogeneity in gender migration gaps across different college admission regimes, with sample restricted to the 13 provinces that experienced changes in admission policies between 1999 and 2003. The dependent variable equals 1 if a student's admitted college is located in a different province from their home province, and 0 otherwise. The admission regimes are defined as follows: "Low Information" (preference submission before the exam), "Medium Information" (submission after the exam but before knowing scores), and "High Information" (submission after knowing scores). Column (1) reports the average gender gap for the sample. Columns (2) and (3) include interactions between the female indicator and the regime dummies (Low Information and High Information); the reference category is the Medium Information regime. Demographic characteristics include agricultural hukou dummy, ethnic minority dummy, and age dummies. Robust standard errors, clustered at the home county level, are reported in parentheses. *** Significant at the 1 percent level; ** significant at the 5 percent level; * significant at the 10 percent level.