NBER WORKING PAPER SERIES

OPTION VALUE NEGLECT:
EVIDENCE FROM CENTRALIZED COLLEGE ADMISSION

Ao Wang
Shaoda Wang
Xiaoyang Ye

Working Paper 33704
http://www.nber.org/papers/w33704

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2025

We thank Nikhil Agarwal, Ned Augenblick, Yan Chen, Stefano DellaVigna, David Gill, Dorothea Kübler, Muriel Niederle, Alex Rees-Jones, Dmitry Taubinsky, Christopher Walters, and various conference and seminar participants for insightful suggestions. Financial support from J-PAL, CEGA, and BFI is gratefully acknowledged. Haoyang Xie and Kaicheng Luo provided excellent research assistance. The research described in this article was approved by UC Berkeley IRB, protocol 2019-02-11806. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

At least one co-author has disclosed additional relationships of potential relevance for this research. Further information is available online at http://www.nber.org/papers/w33704

Option Value Neglect: Evidence from Centralized College Admission
Ao Wang, Shaoda Wang, and Xiaoyang Ye
NBER Working Paper No. 33704
April 2025
JEL No. D47, D81, D91, I21, I28

ABSTRACT

Sequential choices are ubiquitous in daily life, yet making optimal decisions in such settings—where properly accounting for option value is crucial—can be challenging. This paper provides field experimental evidence on the neglect of option value in high-stakes decisions and quantifies the associated welfare consequences. We study a centralized college admissions system where students submit brief preference lists. Although option value should encourage riskier top choices, many students are overly cautious and fail to include safer lower-ranked options. We argue that directed cognition—making myopic decisions while ignoring the value of subsequent options—explains this behavior. An in-field experiment targeting a key framing-based prediction shows that about 50% of applicants exhibit this pattern, especially among disadvantaged students. Counterfactual analysis suggests that de-biasing interventions could significantly reduce outcome gaps and improve overall efficiency.

Ao Wang
Department of Economics
Faculty of Arts and Social Sciences
National University of Singapore
AS2 #06-02, 1 Arts Link
Singapore 117570
Singapore
ao.wang@nus.edu.sg

Shaoda Wang
Harris School of Public Policy
University of Chicago
1307 E. 60th Street office 2055
Chciago, IL 60637
and NBER
shaoda@uchicago.edu

Xiaoyang Ye
Brown University
164 Angell St.,
Providence, RI 02912
xiaoyang.ye26@gmail.com

A randomized controlled trials registry entry is available at
https://www.socialscienceregistry.org/trials/4952

1 Introduction

Sequential choices are ubiquitous in daily life. For example, decisions about how much to consume versus save, when to stop looking for job opportunities, or how long to hold a stock depend on the option value—that is, the tradeoffs associated with future decisions.[1] Models from behavioral economics, such as the model of directed cognition (Gabaix and Laibson, 2005; Gabaix et al., 2006), posit that appreciating the implications of option value may be difficult, and in that case decision-makers could simplify the problem by making isolated decisions without fully considering subsequent options. Despite its potentially far-reaching economic implications, direct field evidence on such neglect of option value remains limited. In this paper, we exploit a stylized setting in which college applicants make a sequence of high-stakes choices and provide among the first field experimental evidence on the prevalence of option value neglect and its associated welfare consequences.

Centralized college admissions provide a prime example of sequential decision-making where option value is critical. In these systems, students are assigned based on their ranked college preferences (i.e., rank order lists or ROLs) and priority scores, yet with a limited number of choices, applicants must carefully decide which colleges to list. Making isolated decisions—option value neglect—can lead to suboptimal educational and career outcomes.[2] To examine this, we study application behavior in a Chinese province that uses a constrained Deferred Acceptance Algorithm allowing students to list up to four colleges from 239 first-tier institutions.[3] The algorithm processes applications in the order of students' scores from the College Entrance Exam (CEE): it first considers a student's top choice, assigning the student if that college's quota remains open; if not, it sequentially moves through the subsequent choices, ultimately relegating the student to a second-tier college if all four selected options are filled. This setting offers a unique window into how option value neglect operates in a high-stakes, real-world environment.

The optimal choice for each spot depends on option value. To avoid relegation, a student needs to list a safe college at the fourth spot. The risk associated with the first three choices, however, is less obvious because the consequence of being rejected from these is not as severe as being relegated. Intuitively, not getting into one's first, second, or third choice is less devastating than not securing one's fourth choice—if these earlier choices are missed, the system can simply consider the next option without triggering relegation. This reasoning

[1] When people decide how much to save versus consume today, considering option value means recognizing that saving provides a "real option" for future flexibility.

[2] The impacts of postsecondary education on labor market income have been documented in many countries, including Chile (Hastings et al., 2013), China (Jia and Li, 2016), and the US (Chetty et al., 2020).

[3] As detailed in Section 2.2, this is also a serial dictatorship mechanism.

suggests that, rather than evaluating each spot in isolation, a student should formulate a contingency plan for the entire portfolio by accounting for the option value at each spot.

We obtained access to administrative data on students' application lists and found that students with identical CEE scores often attend colleges that differ substantially in selectivity. This heterogeneity in admission outcomes is especially pronounced among higher-achieving students. For example, holding exam scores constant, students in the bottom quartile of the SES index attend colleges whose selectivity—measured by the mean cutoff scores from 2014 to 2018—is approximately 0.13 standard deviations lower than that of students in the most advantaged quartile, a difference equivalent to 20% of the baseline exam performance gap between these two groups.

As the admission outcome is jointly determined by CEE scores and students' rank-order lists, we examine students' choices in terms of the selectivity of the colleges they choose, as measured by their historical admission cutoff scores. We find that even for their first choices, high-achieving students tend to be quite cautious. For example, among those whose CEE scores are well above the minimum requirement for first-tier colleges, roughly a quarter pick a college whose unconditional probability of admission exceeds 90% despite the availability of more selective options.[4] Meanwhile, roughly a third of the students exhibit "competitiveness reversals," defined as ranking a less-selective college above their last choice on the rank-order list.[5]

Moreover, these behaviors correlate with students' SES index and contribute to the inequality in admission outcomes. Students with lower SES are substantially more cautious in their first and second choices, but not in their fourth choices. In addition, they are more likely to exhibit competitiveness reversals on their rank-order lists. This evidence suggests that the decision quality of the disadvantaged may be substantially worse, consistent with findings from many other contexts (Lucas and Mbiti, 2012; Ajayi, 2013; Larroucau et al., 2021).

To explain these empirical patterns, we propose that a boundedly rational decision rule—inspired by the Directed Cognition Model (henceforth the DC Model) (Gabaix and Laibson, 2005; Gabaix et al., 2006)—can naturally account for these findings. In our context, the DC Model predicts that students focus entirely on the single spot they are contemplating, thereby neglecting the impact of lower-ranked choices on overall option value. This rule reduces a portfolio choice problem to repeated discrete choice problems in which decisions are made as if the option value of each choice is determined by a cognitive default, such as

[4] Unconditional admission probability is the probability that a student's College Entrance Exam (CEE) score meets the cutoff of a given college, assuming the college is placed first on the student's rank-order list and ignoring other list positions.

[5] Competitiveness is used interchangeably with selectivity throughout the paper.

one's perceived outside options.[6]

Because most centralized mechanisms share the feature that optimal choices across different spots are interdependent, our hypothesis, if substantiated, could have general implications for the design of matching systems. However, testing decision optimality is difficult due to heterogeneous preferences (Agarwal and Somaini, 2018, 2019). This may be why limited progress has been made in understanding whether decision-makers respond to such interdependency in a systematically suboptimal way[7].

We make progress with an incentivized survey experiment conducted among a subset of student applicants immediately after they submitted their ROLs but before admission results were announced. The sample comprises 10% of the applicants in a Chinese province and spans a wide range of socioeconomic statuses. This in-field experiment was designed to distinguish our hypothesis from alternatives and to test a core prediction of the DC Model. Specifically, the DC Model predicts that the DC Type takes more risks when the ROL problem is transformed into a mathematically equivalent lottery formulation—where the option value of a choice is properly incorporated. Based on responses, our analysis suggests that 50.4% (SE = 1.7%) of the students behave according to the predictions of the DC Type. Consistent with our administrative data, a one standard deviation increase in the socioeconomic status index is associated with a 5.3% (SE = 1.8%) decrease in the propensity to be the DC Type. Moreover, we find that this measure systematically predicts students' reported behavior on their college applications, such as the degree of cautiousness in top choices and competitiveness reversals between top and bottom choices.

To assess whether the estimates above explain the observational data, we conduct structural estimation for several specifications using simulated method of moments. We compare a model in which all students play optimally with a mixture model in which some applicants follow the DC Rule while the rest play optimally. The share of the DC Type is set equal to the estimates from the survey experiment. Preferences are identified by exploiting how college choices respond to variation in priority scores, which induces differential rates of change

[6] See Enke and Graeber (2023) for a general treatment of this notion and its application in other domains.

[7] Significant progress has been made in understanding mistakes in the absence of strategic concerns and risk-taking considerations. Artemov et al. (2017) posit that, in a strategy-proof environment, the mistakes on the lists are primarily inconsequential. Shorrer and Sóvágó (2018) find that, in Hungary, dominated choices in college applications are more likely when the expected cost is lower; they argue that multiple imperfections in decision-making may contribute to their findings. Hastings and Weinstein (2008) document a substantial presence of information frictions in school choices using randomized interventions. In settings with strategic considerations, Rees-Jones et al. (2020) find that decision-makers neglect correlation in admission chances, even when such correlation is of first-order concern. Kapor et al. (2020) document mistaken beliefs in centralized system applications using survey data. Finally, Dreyfuss et al. (2019) propose that Koszegi-Rabin expectation-based reference dependence may explain the dominated choices observed in Li (2017)'s experimental data.

in assignment probabilities across colleges. Even though both types are assumed to have the same distribution of preferences, the mixture model fits the data better than the single-type model, with the fit improving by about 40% in terms of predicted moments. Moreover, the mixture model yields substantially better out-of-sample predictions than the model where all students play optimally.

We further estimate a mixture model without imposing the same preference distribution across the two types. The results suggest that college selectivity is the predominant factor in students' choices—much more so than the consideration of geographic distance. While higher-achieving students exhibit more intense preferences for more selective colleges, their SES index has a negligible impact on the marginal utility of these colleges. If anything, more advantaged students seem to derive slightly less marginal utility from selectivity. These estimated preferences indicate that preference heterogeneity alone cannot substitute for the explanatory power of the DC Model. Reassuringly, the estimated preferences are also qualitatively consistent with the satisfaction data we collected in the survey experiment.

Finally, we conduct counterfactual analysis based on the estimated preferences. The main result is that when the DC Type is de-biased such that all students behave optimally, the gap in admissions outcomes between high- and low-SES students (holding exam scores constant) is nearly eliminated. At the same time, the system becomes more efficient for students across all SES index quartiles and CEE performance quintiles. Moreover, due to the higher prevalence of the DC Type and the increased marginal utility for selectivity among high-achieving students, the welfare gain for high-achieving disadvantaged students (those in the lowest SES quartile and highest CEE quintile) can be particularly high—approximately 6% of the utility difference between the best and worst colleges.

This paper relates to three strands of literature. First, it provides among the first field experimental evidence on option value neglect in high-stakes decisions and quantifies the associated welfare consequences. The model of directed cognition, as proposed by Gabaix and Laibson (2005) and Gabaix et al. (2006), is consistent with a broad range of suboptimal behaviors in multi-stage decision contexts—such as failure to learn about selective colleges (Hoxby and Avery, 2012), misperceptions about energy efficiency (Allcott and Taubinsky, 2015), applying for competitive jobs (Abebe et al., 2021), and experimenting with new commuting routes (Larcom et al., 2017). However, isolating this mechanism from other behavioral biases in field settings has been challenging. Our experimental evidence on option value neglect in high-stakes college choices demonstrates the economic significance of the model of directed cognition and echoes emerging laboratory experimental findings that illustrate the fundamental challenge of complexity on revealed preference (Oprea, 2020; Enke and Graeber, 2023; Oprea, 2024).

Second, our paper adds to the literature that studies the systematic presence of behavioral agents in mechanism design (Pathak and Sönmez, 2008; Hassidim et al., 2016; Li, 2017; Rees-Jones and Skowronek, 2018; Dreyfuss et al., 2019). We contribute by showing that option value neglect, a cognitive limitation that is potentially generalizable to other decision-making environments, is empirically relevant and influences the equitability and efficiency of centralized systems. Our analysis demonstrate that such deviation from optimality can be constructively studied using a portable model where the option value is replaced with a cognitive default. The advantage of the DC Model makes it easy to be incorporated into the framework of revealed preference analysis laid out in Agarwal and Somaini (2018, 2019), thereby showcasing the benefits from structural modeling in the analysis of behavioral agents (DellaVigna, 2018).

Third, the paper also adds to the empirical studies on suboptimal decision making in student-school matching (Lucas and Mbiti, 2012; Shorrer and Sóvágó, 2018; Song et al., 2020; Kapor et al., 2020; De Haan et al., 2023). It focuses on the extent to which decision makers take advantage of the theoretical considerations of centralized matching systems, such as the intensity of cardinal preference (Abdulkadiroğlu et al., 2011) and the utilization of insurance options (Chen and Kesten, 2017). Such investigation reveals a previously neglected mechanism that results in decision making inequality, and in that sense our work is related to the literature on the distributional consequences of behavioral biases (Campbell, 2016; Bhargava et al., 2017; Allcott et al., 2019; Rees-Jones and Taubinsky, 2020).

The paper proceeds as follows. Section 2 introduces the empirical setting and data sources. Section 3 lays out the problem mathematically and discusses both the Rational Rule and the DC Decision Rule. Section 4 presents the evidence concerning Cautiousness and Competitiveness Reversals. Section 5 tests Framing Effects using the survey experiment data. Section 6 presents results from the structural estimation. Section 7 concludes and discusses the interpretations of our findings.

2 Empirical Setting

2.1 Summary of Timeline

Figure A1 presents the timeline of the admission procedure. All student applicants are required to take the College Entrance Exam (CEE), a nationwide exam that takes place less than one month before the start of college admissions[8]. Immediately after the exam, the Provincial Education Authorities grade it, a process that takes about two weeks.

[8] See E.2 for the exam structure.

Upon completion of grading, students are notified of their test scores and corresponding provincial rankings; then, the online college application system opens. At that time, students learn whether their scores meet the minimum requirements for applying to schools in the 1st-tier college category, as determined by the Provincial Education Authorities—a criterion that has remained relatively stable over time.[9]

The application process is both time-constrained and cognitively demanding. Students must choose four colleges from a pool of 239 but have few opportunities to learn about the system through trial and error before submitting their final decisions, despite the novelty of the decision environment. Anecdotal evidence suggests that misunderstandings are common. According to several college application advisory platforms on the Chinese Internet, one of the most frequent mistakes is treating the four spots on the ROL as separate and equal, effectively ignoring the order in which the ROL is processed.[10]

2.2 Admission Rule

After the deadline for ROL submission, the centralized admission system assigns students to colleges using a deferred acceptance algorithm based on their priority score and the colleges' pre-announced admission quotas. Since each student's priority score is the same for all colleges, the mechanism effectively functions as a serial dictatorship, with colleges simply specifying a priority score cutoff to determine which applicants are admitted—regardless of a college's position on the student's submitted ROL. Each student knows her priority score and ranking when she applies, and past cutoff scores are publicly available; the only uncertainty arises from the cutoffs for the current year.

Table B1 presents examples of how cutoffs determine admission outcomes. In Example 1, the student is admitted to College A because her score exceeds A's cutoff, and the system ignores all her lower-ranked choices. In Example 2, the student is admitted to College B because her score does not meet A's cutoff but exceeds B's cutoff; as a result, the system assigns her to College B, ignoring C and D. In Example 3, the student is unassigned because she does not meet the cutoff of any college she listed. In Example 4, the student is assigned to College D because she does not meet the cutoffs of A, B, or C but meets the cutoff for D.

After first-tier college assignments are completed, students are notified of their admission decisions within a month. Students not admitted to any first-tier colleges advance to the next stage, where the centralized system assigns them to lower-tier colleges using the same priority score and similar algorithms.[11]

[9] See Appendix E.1 for additional details on college category.

[10] See here for an example of comments about mistakes in college applications (in Chinese).

[11] Students could rank four second-tier colleges and four third-tier colleges.

2.3 Data

The first dataset comprises administrative data generated by the centralized admission system, which records the application behavior of students from 2014 to 2018. This dataset is maintained by the Provincial Education Authorities of Ningxia, where the study took place. The second dataset comes from an online survey experiment conducted in 2020; the timeline for this survey is also displayed in Figure A1. The survey targets CEE takers in Ningxia who are applying to first-tier colleges—the same group of students analyzed in the administrative dataset.

(I) Administrative Data We obtained access to administrative data for first-tier student applicants from 2014 to 2018. As shown in Table B2, the sample includes approximately 6,000 to 8,000 first-tier eligible science track students in each admission cycle. The number is considerably lower for humanities track students, amounting to roughly 1,500 to 2,000 each year. The data contains students' CEE scores, ROLs, admission outcomes, and some demographic information, including the county (available for all years) and the city and street of their residence address (available only in 2015, 2017, and 2018).

(II) Online Survey In August 2020, we conducted an online survey targeting high school students who had applied for first-tier colleges that year. Four local high schools actively encouraged their students to participate, so while any first-tier eligible applicants in Ningxia could respond, our sample mainly consists of students from these four schools. We collected 1,412 complete and effective responses, representing roughly 15% of the total number of first-tier college applicants in 2020. As shown in Figure A1, the survey was administered immediately after students submitted their ROL for first-tier colleges, but before they were notified of the admission outcomes.[12] The survey consisted of three parts. The first part elicited basic information, such as students' final ROL, high school, gender, age, parents' education and occupation, CEE score, source of application advice, and preferences over college characteristics. The second part gathered their beliefs about the unconditional admission probability of the colleges on their ROL (ranging from 0% to 100%) and how satisfied they would feel (on a 0–100 scale) if admitted to a particular college.[13] The third part consisted of

[12] We chose this timing for two reasons. First, to best approximate the informational environment in which the decisions were made—the survey took place after students were notified of their score and had spent time researching colleges, but before the admission results were announced. Second, high school officials believed that administering the survey during the high-stakes, time-sensitive application process could distract students, so it was postponed until after the college application deadline.

[13] The question asked students to guess the probability of meeting the admission cutoff for the college in question. Based on pilot feedback, students were familiar with the concept of admission cutoffs, making it a natural basis for eliciting beliefs.

several incentivized risk-taking questions presented in the form of a ROL and lottery, which are discussed in detail in Section 5.2.

3 Decision Problem

3.1 Setup and Mathematical Notations

To mathematically describe the decision problem in our setting, consider a student, Mei, who needs to list four colleges from a set of colleges on her ROL. The set is denoted by $\mathcal{C}$. After she submits her ROL, along with other students, the centralized admission system will process their ROLs using the Constrained Deferred Acceptance Algorithm (DAA). Then, Mei will either be assigned to one of the four colleges that she listed on her ROL, or will be rejected by all four colleges and end up with her outside option. We assume that the number of students is much larger than the number of colleges $|\mathcal{C}|$.

The information Mei has about other students is incomplete—she does not know the ROLs submitted by other students.[14] As discussed in Section 2.1, because students have been notified of their scores and corresponding provincial rankings when they apply for colleges, the admission outcomes solely depend on the cutoffs of the colleges that they apply for, unknown at the time of list submission. Therefore, Mei only needs to form her beliefs about the chance that her score meets the cutoffs of the colleges that she is interested in. Anecdotally, the notion of cutoff appears easy to grasp, as past cutoffs, which are publicly available, are highly indicative of the competitiveness of a college.

Based on beliefs about the distribution of cutoffs, Mei will assign probabilities of meeting the cutoffs to the colleges that she is contemplating. Denote the probability of vacancy availability (i.e., unconditional assignment probabilities or probability of meeting the cutoff) of the n colleges by p_1, p_2, ..., $p_{|\mathcal{C}|}$ respectively. Denote the utility of admission to colleges by u_1, u_2, ..., $u_{|\mathcal{C}|}$ respectively. The utility of Mei's outside option is $\underline{u}$. For any single college j, the only two characteristics of each choice in this decision making problem are admission utility u_j and unconditional admission probability p_j.

3.2 Estimate Probability p_j from Data

To construct measures that approximate students' perceptions of unconditional probability p_j, we use admission cutoffs in the past, which are publicly available shortly after the end of each previous admission cycle. In Figure 1, we plot the cutoff in its converted form

[14] This assumption is conventional in this literature. Examples include Agarwal and Somaini (2018); Kapor et al. (2020); Calsamiglia et al. (2020).

in a specific year (e.g., 2018) against the cutoff from the previous year (e.g., 2017) for all the colleges that admitted Ningxia students during this time period. We find that the correlation between a cutoff and its past-year counterpart is around 0.95. The uncertainty is larger among less competitive colleges, as the scatter at the bottom left of each graph is more likely to deviate from the 45-degree line.

To calculate probabilities, we assume that for college j, in year t, the cutoff c_{jt} is normally distributed:

$$c_{jt} \sim N(\mu_j, \sigma(\mu_j))$$

where the value of μ_j is the average of admission cutoffs for college j during 2014-2018, which reflects the overall competitiveness of a college across years. The assumptions about the value of μ_j are motivated by the informativeness of past cutoffs. We compute the distribution of $c_{jt} - \mu_j$ in Figure A2a, and find that the distribution function is a bell-shaped function that is centered around zero, with reasonably thin tails, suggesting that it is possible to approximate the true distribution with a hybrid of normal distributions.

Based on the observation from Section 1 that the predictability is heterogeneous across colleges of different competitiveness, we assume that the mean of $ln(\sigma_j)$ is a fourth-order polynomial of μ_j:

$$E[ln(\sigma_j)|\mu_j] = \beta_0 + \sum_{k=1}^{4} \beta_k \mu_j^k$$

We estimate the model using maximum likelihood, for the science and humanity tracks, respectively. We then use the estimated $\hat{\beta}$ to predict σ_j:

$$\hat{\sigma_j} \equiv exp(\hat{\beta}_0 + \sum_{k=1}^{4} \hat{\beta}_k \mu_j^k)$$

With the estimation, the probability of meeting the cutoff of college j is:

$$\hat{p}_{ij} \equiv \Phi(\frac{s_i - \mu_j}{\hat{\sigma_j}})$$

where $\Phi(.)$ is the CDF of standard Gaussian.

We follow Kapor et al. (2020) to validate the estimates of admission probabilities. Specifically, we analyze individual-level data on admission outcomes by running the following regression:

$$1(\text{Admitted to First Choices})_i = \alpha_1 + \beta_1 \hat{p}_{ij}$$

where subscript j represents students' first choices. The null hypothesis that the estimated admission probability is accurate implies that $\alpha_1 = 0$ and $\beta_1 = 1$.

The results suggest that our estimates provide a reasonable approximate to actual probability. As shown in Columns (1) and (2) of Table B3, $\hat{\alpha_1} = -0.004$ (SE=0.002) and -0.005 (SE=0.004) for the science and humanity tracks, respectively, and $\hat{\beta_1} = 1.002$ (SE=0.004) and 1.012 (SE=0.008) in the science and humanity tracks, respectively.

In Figure A2b, we divide colleges into four groups of equal size according to their competitiveness and plot the kernel density estimation of the distribution of $c_{jt} - \mu_j$ for each group, respectively. The figure shows that our model approximates these empirical patterns well.

3.3 The Optimal Rank-Order List

Given the beliefs about the unconditional admission probability of colleges, as well as the utility of admission for each college, the optimal play reduces to calculating the optimal portfolio. Specifically, suppose that Mei's ROL is (a, b, c, d), and the utilities and probabilities are (u_a, p_a), (u_b, p_b), (u_c, p_c), (u_d, p_d), respectively. Given her list, she will be admitted to college a with probability p_a. Under constrained DAA, she will be considered by college b only when college a has rejected her; thus, the probability of being admitted to college b is $(1 - p_a)p_b$[15]. Similarly, the probabilities of being admitted to colleges c and d are $(1 - p_a)(1 - p_b)p_c$ and $(1 - p_a)(1 - p_b)(1 - p_c)p_d$, respectively. Her expected utility from the portfolio $\{a, b, c, d\}$ is:

$$\begin{aligned} EU([a, b, c, d]) \equiv\ & p_a u_a + (1 - p_a)p_b u_b + (1 - p_a)(1 - p_b)p_c u_c + (1 - p_a) \\ & (1 - p_b)(1 - p_c)p_d u_d + (1 - p_a)(1 - p_b)(1 - p_c)(1 - p_d)\underline{u} \end{aligned} \tag{1}$$

Considering joint assignment probabilities for a group of colleges at the same time substantially complicates the decision problem because the chance of admission at one college depends on the chance of admission at the colleges that the student has ranked above it. This interdependence means decisions should not be made by evaluating each program in isolation. Instead, Mei should consider admissions probabilities arising from a complete ROL, and thus optimal decision-making requires picking an optimal portfolio out of a large number[16].

Given the structure of the ROL, an optimal list (a^*, b^*, c^*, d^*) must be ordered to satisfy

[15] Here we are assuming that admission probability is independent because we believe that this is a reasonable approximation in our empirical setting. See Section E.3 for a detailed discussion of why this assumption is justified.

[16] To be precise, the number is $\binom{n}{4}$, which amounts to 4 billion in our context.

the property that $u_a^* \geq u_b^* \geq u_c^* \geq u_d^*$. A more difficult and important question in this decision problem is which four colleges to choose from hundreds of alternatives. Chade and Smith (2006) find that a decision rule, Marginal Improvement Algorithm, can achieve the global optimum. This algorithm decompose this decision into four steps, each of which feature a distinct discrete choice problem. In each step, the optimum depends on the colleges that have already been included in the portfolio in previous steps.

To understand what an optimal list looks like, note that optimality entails the following conditions to hold for the top choice a:

$$p_a u_a + (1-p_a)EU([b,c,d]) \geq p_x u_x + (1-p_x)EU([b,c,d]) \text{ for any } x \in \mathcal{C}/[b,c,d] \tag{2}$$

Similarly, for the second choice b:

$$p_b u_b + (1-p_b)EU([c,d]) \geq p_x u_x + (1-p_x)EU([c,d]) \text{ for any } x \in \mathcal{C}/[a,c,d]$$

For the third choice c:

$$p_c u_c + (1-p_c)EU([d]) \geq p_x u_x + (1-p_x)EU([d]) \text{ for any } x \in \mathcal{C}/[a,b,d]$$

and for the fourth choice d:

$$p_d u_d + (1-p_d)\underline{u} \geq p_x u_x + (1-p_x)\underline{u} \text{ for any } x \in \mathcal{C}/[a,b,c] \tag{3}$$

Therefore, it is the colleges that the student ranks below the current choice, not the colleges ranked above it, that matter most for the optimal choices. These necessary conditions thus reflect the logic of backward induction, as visualized below:

- (blank) $\rightarrow$ (blank) $\rightarrow$ (blank) $\rightarrow$ d^* $\rightarrow$ (outside option)
- (blank) $\rightarrow$ (blank) $\rightarrow$ c^* $\rightarrow$ d^* $\rightarrow$ (outside option)
- (blank) $\rightarrow$ b^* $\rightarrow$ c^* $\rightarrow$ d^* $\rightarrow$ (outside option)
- a^* $\rightarrow$ b^* $\rightarrow$ c^* $\rightarrow$ d^* $\rightarrow$ (outside option)

3.4 Formulation of the Directed Cognition Model

The optimal benchmark in Section 3.3 proposes that the best portfolio can be reached by decomposing the problem into four distinct discrete choice problems. The correct decomposition of this problem requires student applicants to appreciate the interdependence between choices, because it is the colleges they list below a given rank—not those above—that affect the optimal choice for that rank. Literature in experimental economics, however, has established that even in simplified settings, subjects have trouble grasping the concept of backward induction (Camerer et al., 1993; Johnson et al., 2002). Moreover, laboratory evidence suggests that decision-makers lack the ability to cope with uncertainty in simple decisions (Martínez-Marquina et al., 2019), a skill that is necessary to assess the distribution of utility for the lower-ranked choices.

In this subsection, we continue to use the setting in Section 3.3, and introduce an alternative boundedly rational decision rule that is inspired by the Model of Directed Cognition (the DC Model) (Gabaix et al., 2006). We hypothesize that this decision rule explains students' suboptimal strategies when they cannot apply the optimal strategy as prescribed in Chade and Smith (2006). The rule prescribes that, instead of tracking all the information and acting optimally upon it, another student applicant, Hua, due to cognitive limitations, myopically focuses on the spot where he is actively contemplating which college to fill in, and neglects the impact of that choice on the rest of the list. He fills out his ROL in a sequential order, from top to bottom, rather than using backward induction. Mathematically, Hua chooses i to maximize

$$p_i u_i + (1 - p_i)\underline{u} \tag{4}$$

among options available. As a result, in each step Hua is making choices for essentially the *same* decision problem. That is, he maximizes the expected utility of a portfolio that consists of his current choice and a cognitive default $\underline{u}$, such as his perceived outside option. Let (a', b', c', d') denote the result of such decision rule, the choice process is described graphically below:

- **Step 1**: a' → (blank) → (blank) → (blank) → (Outside Option)
- **Step 2**: a' → b' → (blank) → (blank) → (Outside Option)
- **Step 3**: a' → b' → c' → (blank) → (Outside Option)
- **Step 4**: a' → b' → c' → d' → (Outside Option)

This decision rule requires less cognitive capacity for two reasons. First, because Hua is making choices for each spot in isolation, the source of uncertainty is reduced and he only needs to consider at most two states: admission to his choice in question, or getting the utility of the outside option. Second, Hua is proceeding in a natural order by considering the first choices before the rest of the ROL.

Such myopic decision making comes at a cost. Compared to the necessary condition for optimality (Equation 2), the condition for DC (Equation 4) ignores the colleges that are ranked below the choice. When the perceived outside option is equal across two decision rules, the DC rule appears more cautious. Take the first choice on his list, for example; when he uses DC rule, his perceived utility from backups is merely $\underline{u}$, far below the actual expected utility from the colleges that follow it $EU([b, c, d])$. This difference is most pronounced at the first choice of the ROL because that is where the difference between *actual* and *perceived* expected utility from the options ranked below is maximal. In contrast, a DC Type's risk-taking behavior would not differ substantially from its rational counterparts for the fourth choices of ROL because that is where the *actual* (Equation 3) and *perceived* (Equation 4) expected utility coincide.

4 Risk-Taking Behavior from Administrative Data

Sections 3.3 and 3.4 have introduced and compared the Optimal Rule and the DC Rule and show that intuitively they yield different predictions regarding risk taking behavior for different spots on the ROL. Section 4.1 discusses their differences when horizontal preferences are present. Section 4.2 discusses the admission outcome gap between students of different SES. Sections 4.3 and 4.4 discuss how the differences in risk-taking behavior between the advantaged and disadvantaged students appear in line with the prediction in Section 4.1.

4.1 Predictions: Cautiousness and Competitiveness Reversal

Following the notations in previous sections, let (a^*, b^*, c^*, d^*) denote the set of choices by the Optimal Rule, and let (a', b', c', d') denote the set of choices by the DC Rule. For a given student, $\vec{u} \equiv (u_1, u_2, ..., u_{|C|})$ is his preference profile for all colleges, and $\vec{p} \equiv (p_1, p_2, ..., p_{|C|})$ is the vector of his admission probabilities.

Prediction 1 (Top-Choice Cautiousness) *For any $\vec{u}$ and $\vec{p}$, assuming outside option is zero and there do not exist $x \neq y$ such that $p_x u_x = p_y u_y$, we have:*

1. $p_{a^*} \leq p_{a'}$;
2. *When $p_{a^*} = p_{a'}$, we have $a^* = a'$ and $p_{b^*} \leq p_{b'}$;*
3. *When $p_{a^*} = p_{a'}$ and $p_{b^*} = p_{b'}$, we have $b^* = b'$ and $p_{c^*} \leq p_{c'}$;*
4. *When $p_{a^*} = p_{a'}$, $p_{b^*} = p_{b'}$, and $p_{c^*} = p_{c'}$, we have $c^* = c'$ and $p_{d^*} \leq p_{d'}$;*

The proof is in Appendix D. This prediction shows that the DC Rule always predicts cautiousness for the top choices compared to the Optimal Rule, no matter what value $\vec{u}$

takes. It clarifies that the two rules are not distinguishable in terms of risk-taking only if the option value of non-top spots does not change the consideration of the choices for higher-ranked spots, in which case option value neglect does not lead to sub-optimal choices.

Prediction 2 (Competitiveness Reversals) *For any $\vec{u}$ and $\vec{p}$, assuming outside option is zero and there does not exist $x \neq y$ such that $p_x u_x = p_y u_y$. If $max\{p_{a^*}, p_{b^*}, p_{c^*}\} > p_{d^*}$ (i.e. reversal between non-bottom and bottom choices under the Optimal Rule), then there must be competitiveness reversal in list (a', b', c', d').*

The proof is in Appendix D. This prediction shows that for any preference profiles, the DC Rule will only generate more instance of the competitiveness reversals compared to the Optimal Rule.

The presence of cautiousness and competitiveness reversals jointly characterizes the set of preferences that renders the two decision rules indistinguishable.

Observation 1 *Assuming outside option is zero, consider the ratio of probability among pairs of choices $L_{ij} = \frac{p_i}{p_j}$, $i, j \in \{a, b, c, d\}$ where i, j constitutes a competitiveness reversal, then the two decision rules yield identical risk taking behavior only if for any $x \notin \{a, b, c, d\}$ and $i, j \in \{a, b, c, d\}$, $L_{ij} p_x u_x \leq p_i u_i$.*

The proof is in Appendix D. The more cautious choices are, the higher L_{ij} is. The more severe the reversal is, the higher L_{ij} is. Higher L_{ij} implies that the set of $\vec{u}$ and $\vec{p}$ that render the two decision rules indistinguishable will be smaller. In the extreme, the two rules yield identical outcomes if there is satiation regarding four colleges among the hundreds of alternatives. Observation 1 characterizes the middle ground between this extreme form of horizontal preferences and pure vertical preferences, in which case any competitiveness reversal would be sufficient to indicate the presence of the DC Rule.

4.2 Outcome Gap in terms of Competitiveness

A large literature documents that socioeconomically disadvantaged students are worse at strategizing in centralized systems (Lucas and Mbiti, 2012; Ajayi, 2013; De Haan et al., 2015; Shorrer and Sóvágó, 2018; Kapor et al., 2020). To examine whether students of low SES end up in worse colleges, as suggested by existing research, we approximate socioeconomic status using average educational attainment at the township level[17]. We obtained statistics on the average years of education among adults between 40 and 65 at the township level from the China Census in 2010, as well as access to students' home addresses in 2015, 2017, and

[17]Roughly speaking, township is equivalent to zip code in the US.

2018[18]. We match the township-level educational attainment data to individual students in the administrative dataset and plot the distribution of this measure in Figure A3. We split students into four groups according to their SES and focus on the most advantaged quartile and the most disadvantaged quartile in this subsection.

To analyze the gap in admission outcomes, we run the following regression:

$$\text{Outcome}_i = \beta_1 \mathbb{1}_i(\text{1st Quartile}) + \beta_4 \mathbb{1}_i(\text{4th Quartile}) + f(\text{CEE Score}_i) + \mathbf{X}_i + \epsilon_i \qquad (5)$$

Admission outcome is constructed by the average standardized cutoffs during 2014-2018 for the colleges that admit the student in question[19]. $f(\cdot)$ is a fourth-order polynomial function. The coefficients of interest are β_1 and β_4, which indicate, compared to the second and third SES quartiles, whether being in the lowest or highest SES quartile leads to admission to more competitive colleges, conditional on having the *same* CEE score. Since admission outcomes are jointly determined by students' CEE scores and their application lists, the estimates of β_1 and β_4 can only be attributed to systematic differences in ROLs.

Table 1 shows the outcome gap obtained by running regressions on different samples with different sets of controls. The gap between the first quartile (β_1) and the fourth quartile (β_4), $\beta_1 - \beta_4$, can be as large as 0.105 standard deviations (Panel A, Column 1). The gap amounts to 0.129 standard deviations when the sample is restricted to higher-achieving students (Panel B, Column 1).

When more controls are added, the estimated gaps shrink by up to 50%. We note, however, that these control variables are themselves reflective of students' socioeconomic backgrounds. These variables apparently capture omitted aspects of socioeconomic background compared to a more straightforward measure of adults' average educational attainment at the township level. The main takeaway of this exercise is that the gap appears quite robust and consistent, and is not easily explained by any single observable characteristic.

Overall, when compared to the middle (second and third) SES quartiles, the estimates of β_4 for the fourth SES quartile clearly show that the most advantaged students' outcomes are substantially higher. Similarly, the estimates of β_1 for the first SES quartile show that the most disadvantaged students' outcomes are consistently worse. The signs of these estimates appear robust, and the effect is slightly larger for high-achieving students, as shown in Panel B. When restricted to the STEM track (Columns 4-6), the estimates are similar to those obtained when both tracks are included (Columns 1-3).

[18] For 2014 and 2016, we cannot obtain data at such a granular level and have only been able to obtain county-level statistics.

[19] Figure A4 shows the distribution of the un-standardized cutoffs of the colleges that admit a student by their CEE score quintile. The figure helps visualize the competitiveness of the placement among students of different CEE performance.

4.3 Top-Choice Cautiousness: Advantaged vs. Disadvantaged

As previous literature also suggests that disadvantaged students may be more likely to make strategic mistakes. We seek to understand in this subsection whether the observed outcome gap could be associated with option value neglect, as prescribed by the DC Decision Rule. We focus on the comparison between the fourth (most advantaged) quartile and other quartiles and consider the following regression:

$$\text{Prob}_i = \beta_4 \mathbb{1}_i(\text{4th Quartile}) + f(\text{CEE Score}_i) + \mathbb{1}_i(\text{4th Quartile}) * f(\text{CEE Score}_i) + \mathbf{X}_i + \epsilon_i \quad (6)$$

where the dependent variable captures students' risk-taking behavior for a specific choice on their lists, the degree of risk-taking is measured by the unconditional probability (i.e., the probability that a student's score meets the cutoff of a college in question, disregarding other choices on the list). We allow the gap between the fourth quartile and other quartiles to vary with the percentile of CEE score by including the interaction term between the dummy "Fourth Quartile" and the fourth-order polynomial of CEE score percentile.

As discussed in Sections 3.4 and 4.1, a key difference between the DC Rule and the Optimal Rule is that the DC Rule, in particular, predicts substantially more cautiousness for top choices but not for bottom choices. In Table 2, we show in Panel B Column 1 that the first-choice probability for the most advantaged quartile is 4.90% lower (i.e., more risk taking) than that of other quartiles, conditional on having the same CEE score. The explanation that risk aversion—or lack of interest in competitive colleges—cannot easily account for this gap is reinforced by Panel B Column 2, which shows that the most advantaged quartile is actually *more* cautious than their counterparts. Finally, in Panel B Column 3, we use the difference between the fourth and the first choice as a measure of diversity in risk-taking, and find that the most advantaged quartile has a substantially more diverse portfolio than other quartiles, with the probability gap between the first and fourth choices being roughly 6.98% larger than that of other quartiles.

The above findings appear to be quite uniform across students' CEE scores. Figure 2a shows a bin-scatter plot of the mean unconditional probability for all four choices on the list, by CEE score percentile. While students with different CEE scores exhibit varying levels of risk, the average risk-taking among the most advantaged quartile appears substantially more diverse than in the other quartiles. Specifically, for high-achieving students (e.g., CEE percentile $\geq 40\%$), they are substantially more risk-taking for their first and second choices, but slightly more cautious for their fourth choices. For lower-achieving students (e.g., CEE percentile $\leq 40\%$), the diversity in risk-taking seems to stem mainly from greater caution in bottom choices. Table 2 Panel C quantifies this visual evidence and shows precisely estimated gaps.

The gap also holds when comparing the cumulative distribution of the unconditional probability for the highest quartile with those of the other quartiles. As shown in Figure 2b, for the first and second choices, the CDF of the other quartiles stochastically dominates that of the highest quartile. The gap is quite substantial—for example, among the most advantaged quartile, approximately 26.91% of students have a first-choice unconditional probability exceeding 75%, whereas in the other quartiles, this share is about 34.72%, roughly one-third higher than that for the most advantaged quartile. In contrast, for the fourth choices, the CDF of the other quartiles appears to be weakly stochastically dominated by that of the highest quartile.

These patterns jointly show that the most advantaged students' portfolios are uniformly more diverse than those of their disadvantaged counterparts. They also suggest that disadvantaged students may not be applying for colleges optimally, effectively neglecting option value. While we cannot determine the optimal level of diversity without making assumptions about preferences, Panel A in Table 2 shows that many high-achieving students appear to be extremely cautious even with their first choice (e.g., one quarter of these students apply for a college whose unconditional probability exceeds 90.96%). Such lack of risk-taking seems incompatible with their high motivation and substantial investment in CEE preparation.

4.4 Competitiveness Reversals: Advantaged vs. Disadvantaged

Another prediction from the discussion in Section 4.1 is that for a given set of preferences, the DC rule will always generate more competitiveness reversals, meaning a riskier college is placed at a lower-ranked spot. This pattern contradicts the risk-taking strategy intuitively prescribed in Section 3.3, where students should take more risk for higher-ranked spots.

These predictions intuitively suggest that in order for students to construct a portfolio with diverse levels of risk, it is often necessary that the unconditional probability of their fourth choices is substantially larger than that of their first choices. Otherwise, it either indicates a direct competitiveness reversal or suggests a lack of diverse risk-taking, which potentially leads to competitiveness reversal elsewhere on the ROL. In Figure 3a, we show the cumulative distribution of the difference between the fourth-choice probability and the first-choice probability for the most advantaged quartile and for other quartiles. Echoing previous findings on cautiousness, the figure displays a significant gap in this difference, with the distribution for the most advantaged quartile stochastically dominating that of the other quartiles. Moreover, this gap is especially substantial in the area of interest, where roughly 41.42% of the students in the most advantaged quartile have a fourth-first probability difference smaller than 25%, whereas the corresponding share climbs to roughly 51.14% for other quartiles. If we instead consider the more severe case where there is a competitiveness reversal between the first and fourth choices, roughly 8.61% of the students in the most advantaged quartile exhibit a reversal, whereas the share nearly doubles to 15.54

In Figure 3b, we also include the second and third choices in the analysis by examining the CDF of the following statistic: $R = max\{p_a, p_b, p_c\} - p_d$. Here, R will be negative when at least one of the top three choices constitutes a risk-taking reversal with the fourth choice. The more negative R is, the more severe the reversal. The graph demonstrates that the CDF of the highest quartile stochastically dominates that of the other quartiles for $R < 0$, suggesting that regardless of the degree of reversals examined, the gap between advantaged and other quartiles persists. Specifically, when we include all such reversals ($R \leq -1\%$), 35.70% of the students in the other quartiles commit reversals, whereas the share is only 25.49% for students in the most advantaged quartile. When we look at the more severe reversals ($R \leq -25\%$), 9.53% of the students in the most advantaged quartile commit such reversals, while the share nearly doubles to 17.01% for students in the other quartiles.

5 Survey Experiment: Testing Framing Effect

Section 4 shows that disadvantaged students tend to have less diverse portfolios and exhibit more reversals. This evidence is consistent with the possibility that more disadvantaged students neglect option value, as predicted by the DC Model. However, students' decisions may also be influenced by horizontal preferences, information frictions, or subjective beliefs.

To address these concerns, we design a survey experiment in which students make incentivized, hypothetical choices involving colleges and lotteries. The monetary incentives and risks associated with these hypothetical colleges are structured to test predictions of the DC Rule in a way that is not confounded by the factors mentioned above. One of the questions presented in this section will be randomly selected to "count," and students will be paid based on their response to that question.

5.1 Prediction: Framing Effect

The correct utility of choosing college a as the rth choice is

$$U_a \equiv p_a u_a + (1 - p_a) U_r,$$

where U_r is the expected utility of the portfolio consisting of all options ranked below the rth position. The challenge for a DC Type arises when the decision is presented in the form of a ROL (Rank-Ordered List): they fail to compute U_r and instead treat it as a default fallback utility u_0.

However, when the same decision is presented as a mathematically equivalent lottery$(p_a, u_a; 1-p_a, U_r)$. The payoff in the event of rejection is explicitly specified. This clarity prevents the DC Type from misperceiving the utility of backup options. In effect, presenting choices in the form of a lottery "de-biases" the decision by drawing the decision-maker's attention to the utility of backup options, reducing the cognitive focus on a single outcome.

More generally, let U_r' denote the perceived expected utility of the portfolio ranked below the rth choice. A higher U_r' increases the willingness to take risk, as the downside appears more favorable. For an Optimal Type, $U_r' = U_r$, regardless of whether the decision is framed as a lottery or a ROL. For a DC Type, $U_r' = U_r$ in the lottery frame, but $U_r' = u_0 < U_r$ in the ROL frame:

Prediction 3 (Framing Effect) *An Optimal Type will behave consistently across the ROL frame and the lottery frame. A DC Type will be more risk-taking in the lottery frame compared to the ROL frame.*

5.2 Design of the Survey Experiment

The core of the survey experiment consists of three groups of incentivized questions, presented as three Multiple Price Lists (MPL).[20] The first and third groups ask students to indicate their preferred level of risk by choosing either College X or College Y as the first choice in a hypothetical Rank-Ordered List (ROL). Each list includes seven questions, as shown in Panels A2 and C2 of Table B4. The payoff associated with College X remains constant—an admission probability of 50% and a reward of 25 CNY if admitted—while the payoff associated with College Y increases across questions, holding the admission probability fixed at 25% but raising the reward from 30 to 60 CNY in 5 CNY increments.

The second, third, and fourth positions on the ROL are fixed in advance, as displayed in Panels A1 and C1 of Table B4, with one of these colleges guaranteed to admit the student if she is not admitted to her first choice. The "rules of admission" for both question groups mirror those of the actual college admission process.

The difference between the first and third MPL lies in the value of the backup payoffs: in Question Group 1, admission to lower-ranked colleges yields 20 Chinese Yuan (CNY), while in Question Group 3, the amount is only 5 CNY. These scenarios capture the core trade-off in our setting: students who wish to take on more risk by selecting a more desirable college

[20] Appendix F provides sample screenshots of these comprehension questions.

as their top choice must also face a greater chance of being admitted to backup options with substantially lower payoffs.

To ensure that the questions closely resemble the high-stakes college application decisions students make for themselves, the specific names of the colleges are customized at the individual level based on the exam score students report at the beginning of the survey. The benefit of this approach is that it eliminates the need to introduce an entirely new decision environment—something that focus group results suggest may be less effective in a general survey context. The downside, however, is that students who misinterpret the instructions may project their own preferences onto the named colleges, potentially overriding the monetary incentives intended to discipline their stated preferences. To minimize this possibility, we implement two measures. First, we clarify that participants are assisting another student in making college application decisions, and this hypothetical student's preferences are explicitly shown via satisfaction scores assigned to College X, College Y, and other options. To avoid confusion over currency conversion, these satisfaction scores also represent the monetary incentive (in Chinese Yuan) awarded upon "admission." Second, we include comprehension questions to assess students' understanding of how each college corresponds to potential monetary payoffs. Participants are provided feedback and corrected if they respond incorrectly.

The second group of questions is mathematically equivalent to the first but is presented in a lottery frame. This group also consists of seven questions. Lottery X has a fixed payoff structure of (25 CNY, 50%; 20 CNY, 50%), which replicates the payoff distribution of College X in Question Group 1 and remains constant across questions. The payoff structure of Lottery Y varies across questions—(30 CNY, 25%; 20 CNY, 75%), (35 CNY, 25%; 20 CNY, 75%), ..., up to (60 CNY, 25%; 20 CNY, 75%)—and is mathematically equivalent to the payoff profile of College Y in Question Group 1. However, the framing in Question Group 2 differs from that of Group 1: the probability and payoff associated with rejection from the first choice (i.e., the downside) are explicitly included in each option, as shown in Panel B of Table B4. This distinction matters because DC Types in our setting evaluate their top college choice in isolation and neglect the utility of lower-ranked options, thereby failing to correctly translate the ROL into its equivalent lottery. By explicitly incorporating the downside payoff, the lottery frame corrects this cognitive error and neutralizes the bias introduced by the DC decision rule.

Students were instructed to carefully read the explanations provided before answering the questions. Since the payoff from Option X is fixed while the payoff from Option Y increases across questions, a coherent pattern of responses should involve at most one switch from choosing X to choosing Y. This principle is clearly communicated in the instructions, and students are allowed to switch from X to Y at most once in their responses.

5.3 Analysis of Survey Data

Panel A of Table 3 examines whether students from disadvantaged socioeconomic backgrounds, as measured by their parents' average years of education, are more likely to exhibit the framing effects predicted by the DC Decision Rule. Linear probability models with varying sets of controls show that a one standard deviation increase in the normalized SES index is associated with a 4.3 to 4.9 percentage point increase in the likelihood of belonging to the red block (i.e., students who exhibit substantial framing effects).

To estimate the share of DC Types in the survey sample, we model students' risk-taking behavior using a power utility function over monetary outcomes:

$$u = (c + B)^{\rho},$$

where B represents background consumption, set to 10 CNY[21]. The curvature parameter ρ is assumed to lie in the range $0.05 \leq \rho \leq 1$, and varies with SES (standardized parental education) and CEE (standardized priority score):

$$\rho = \min\left\{\max\left\{0.05,\ \alpha_0 + \alpha_1 \text{SES} + \alpha_2 \text{CEE} + \epsilon_\rho\right\}, 1\right\},$$

where $\epsilon_\rho \sim N(0, \sigma_\rho^2)$.

We incorporate decision noise by assuming that students perceive the payoff of Choice Y with error. In Question Group 2, let the lotteries be defined as $\tilde{X}(20) \equiv (25,\ 50\%;\ 20,\ 50\%)$ and $\tilde{Y}(m, 20) \equiv (m,\ 25\%;\ 20,\ 75\%)$. For Optimal Types, the riskier option $\tilde{Y}(m, 20)$ is perceived as $\tilde{Y}(m + \epsilon,\ 20)$, where $\epsilon \sim N(0, \sigma_\epsilon^2)$ is i.i.d. across individuals.

For Question Groups 1 and 3, we define the lottery equivalents of College X and Y as $X(b) \equiv (25,\ 50\%;\ b,\ 50\%)$ and $Y(m, b) \equiv (m,\ 25\%;\ b,\ 75\%)$, where m is the first-choice payoff and b is the backup payoff. Optimal Types are assumed to correctly process the ROL questions as equivalent to these lotteries, and we apply the same noise structure as in Question Group 2.

DC Types, in contrast, misperceive the backup payoff in the ROL frame as 0. Thus, they interpret the questions as involving lotteries $X(0)$ and $Y(m, 0)$, leading to more cautious behavior. As with Optimal Types, noise is added to m.

We also include a third behavioral type established in the literature: the Sincere Type (Pathak and Sönmez, 2008; Calsamiglia et al., 2020). In the ROL frame, Sincere Types

[21] This amount roughly corresponds to the cost of a meal in Ningxia.

always choose the college with the highest nominal payoff, ignoring admission probabilities. Noise is added to m, which corresponds to the higher of the two payoffs.

The decision noise variance σ_ϵ^2 is allowed to vary across question groups but is assumed to be constant across behavioral types.

We further assume in this mixture model that the share of DC Types varies with socioeconomic status:

$$\Pr(\text{DC Type} \mid \text{SES}, \text{Score}) = \frac{\exp(\beta_0 + \beta_1 \text{SES} + \beta_2 \text{Score})}{\exp(\beta_0 + \beta_1 \text{SES} + \beta_2 \text{Score}) + 1}.$$

We estimate this mixture model and present the main results in Panel B of Table 3. Including the DC Type in the estimation (Column 2) substantially improves model fit (log-likelihood = −8781.778, 9 parameters) compared to the baseline model in Column 1, which only allows for the Optimal Type (log-likelihood = −9100.574, 6 parameters). The improvement is substantial: the Bayesian Information Criterion (BIC) also favors the mixture model (BIC = 17628.83) over the single-type rational model (BIC = 18244.66).

Adding the Sincere Type (Column 3) provides a further improvement in fit, but the effect is limited relative to the DC Type (log-likelihood = −8780.909, 10 parameters). Consistent with this, the estimated share of the DC Type in our preferred specification is 50.4%. The marginal effect of socioeconomic status is statistically significant: a one standard deviation increase in the SES index is associated with a 5.5 percentage point decrease (SE = 1.8%) in the likelihood of being classified as a DC Type.

In contrast, the CEE score is not significantly associated with DC Type classification. This may seem surprising, but it is important to note that our sample includes only students eligible for first-tier colleges—that is, those already in the top 25–30 percent of the cohort. Moreover, as all students have been intensively trained for the CEE over several years, it is unclear whether CEE scores are informative about their ability to make decisions in a novel, real-life context like this one.

5.4 Framing Effect Predicts Risk-Taking Behavior in College Choices

While we do not have access to administrative data from 2019 and are therefore unable to link students' responses directly to the centralized application system, we did ask them a series of questions about their real college applications, as well as their decisions in other counterfactual but incentivized scenarios. These questions allow us to connect their responses—analyzed in the previous subsections—to a broader set of outcomes.

Specifically, regarding their actual college applications, we asked students the following:

1. The Rank-Ordered List (ROL) they submitted.

2. For each choice on the list, their perceived probability of meeting its cutoff score.

If responses to the incentivized risk-taking questions genuinely reflect the presence of DC-Type behavior in real college application decisions—rather than merely a pattern of responses specific to the hypothetical survey setting—then we would expect their reported ROLs to exhibit behaviors consistent with Predictions 1 and 2 as well.

In addition, we ask participants to engage in a separate college selection task in which they construct a four-college Rank-Ordered List (ROL) from a set of colleges recommended to them by the system. These recommendations are generated based on their individual information and stated considerations regarding college applications.

Similar to the questions introduced in Section 6.2, the recommended colleges are assigned satisfaction scores, which serve as the basis for monetary rewards in the event that a participant's response is selected for payment. In this task, the event of "admission" is defined as the participant's actual CEE score exceeding the publicly announced cutoff score for a given college in the 2020 admission cycle.

Following the construction of the full ROL, we ask participants the following counterfactual questions:

1. If only two colleges could be included in the list, which two would you choose?

2. If only one college could be included in the list, which college would you choose?

We compare students' answers on the shortened list tasks to their original four-college lists. The rationale is that the DC Decision Rule differs from the Optimal Decision Rule by prescribing that decision-makers make choices in a forward-looking manner, rather than by backward induction—particularly when more competitive colleges are also more desirable.[22] This implies that if students are DC Types, they should be more likely to select their top-choice colleges—rather than safer, lower-ranked options—when constrained to include only one or two colleges on an artificially shortened list.

Panel C of Table 3 presents summary statistics of students' college choices, disaggregated by whether they exhibit a substantial framing effect (i.e., whether they belong to the red block in Figure 4). Applicants who are very cautious in the ROL frame but not in the equivalent lottery frame are significantly more likely to state that they would have chosen their top-ranked colleges had they been allowed to include only one or two options. Furthermore, these students are significantly more likely to report that they selected safer colleges for their top

[22] As discussed in Section 4.1, in the presence of substantial horizontal preferences, there are exceptions where the Optimal Rule leads to the same outcome as the DC Rule, such that backward induction is not required.

choice (i.e., exhibit top-choice cautiousness), but not for their bottom choice, and are also more likely to report committing competitiveness reversals.

6 Structural Estimation using Administrative Data

In this section, our primary objective is to estimate students' preferences by fitting their application choices under different assumptions about the decision rule they employ. Section 6.1 describes the available variation in the administrative data. Sections 6.2 and 6.3 outline the model setup and estimation procedures. Finally, Sections 6.4 and 6.5 present the main results and discuss their welfare implications.

6.1 Identification

As discussed in Section 4.1, when students are satiated with a few non-competitive colleges, applying either the Optimal Rule or the DC Rule yields identical outcomes. In other words, it is always theoretically possible to explain observed decisions using some configuration of the preference vector $\vec{u}$. This implies that, without imposing any distributional assumptions on $\vec{u}$, we must rely on additional variables or sources of variation for identification, as emphasized in Agarwal and Somaini (2018).

However, Section 4.1 also provides positive results: for *any given* $\vec{u}$, the DC Rule generates a *higher* degree of top-choice cautiousness and a *greater* incidence of competitiveness reversals compared to the Optimal Rule. This implies that, given a parameterized distribution of college preferences, one can separately identify preferences and decision rules. Specifically, the reduced-form patterns documented in Section 4 can be exploited to achieve identification of the behavioral decision rule.

As the estimated share of DC Types from structural decomposition may be sensitive to the assumed distribution of preferences, we adopt a more direct approach by imposing the share of DC Types to match the survey-based estimate in the structural estimation. This additional data source enables the core exercise of this section, in which we estimate and compare three specifications:

1. All students use the Optimal Rule.

2. Students use either the Optimal Rule or the DC Rule. Conditional on students' SES and CEE scores, the share of DC Types is fixed to match the estimates from Section 5. The distribution of preferences is assumed to be identical across both types.

3. Students use either the Optimal Rule or the DC Rule. Conditional on students' SES and CEE scores, the share of DC Types is fixed to match the estimates from Section 5. However, the distribution of preferences is allowed to differ across the two types.

The comparison between the first and second specifications provides a test of the explanatory power of the DC Model for real college application behavior, under the assumption that the distribution of decision types among 2020 survey respondents is similar to that of applicants from 2014 to 2018. The resulting improvement in fit may be viewed as a lower bound on the explanatory potential of incorporating the DC Model, since violations of this assumption would typically attenuate the model's predictive power. The third specification offers the most flexible framework, allowing for heterogeneity in preferences across types and enabling an evaluation of the welfare implications of debiasing.

While the survey-based estimates offer key advantages, one may wonder what variation the administrative data can provide. In our setting, it is difficult to leverage some of the common sources of variation found in other datasets, such as applicants' geographic location. Most first-tier colleges are located in major coastal cities, while only two such colleges are situated in Ningxia. Furthermore, Ningxia is geographically small compared to other provinces and is distant from the coastal regions. As a result, the location of applicants' homes within Ningxia exhibits very limited variation in terms of access to colleges, since all major institutions under consideration are roughly equally distant from any two points within the province.

The main source of variation we exploit is in students' priority scores. We assume that the distribution of preferences changes smoothly with respect to the priority score—that is, as the priority score increases, preferences evolve more gradually than admission probabilities. This assumption is justified on two grounds.

First, students cannot perfectly control their total scores, and anecdotal evidence suggests that mock exam performance is a poor predictor of actual CEE outcomes. Second, exam difficulty varies across years and across subjects, generating exogenous variation in priority scores. Since the priority score is the sum of raw scores across all subjects—and subject-specific scores are not standardized—this variation is not purely reflective of overall ability. For instance, students with high math ability may achieve a higher ranking when the math exam is particularly discriminating.[23]

Such variation affects the assignment probabilities of all colleges—often at *different rates*. To illustrate this point, consider two colleges, A and B, with admission cutoffs distributed as $c_A \sim N(\mu_A, \sigma_A^2)$ and $c_B \sim N(\mu_B, \sigma_B^2)$, respectively, where $\mu_A > \mu_B$. For a student with

[23] Figure A6a plots the standard deviation of exam scores, normalized by total subject score, from 2014 to 2018. The figure shows clear year-to-year variation in the exams' ability to differentiate among students.

priority score s, the unconditional admission probabilities are given by $p_A(s) = \Phi\left(\frac{s-\mu_A}{\sigma_A}\right)$ and $p_B(s) = \Phi\left(\frac{s-\mu_B}{\sigma_B}\right)$, where $\Phi(\cdot)$ denotes the standard normal cumulative distribution function.

We note the following:

Observation 2 *Consider the probability ratio $\mathcal{L}(s) = \frac{p_A(s)}{p_B(s)}$. With sufficient variation in s, we have:*

- *If $\sigma_A = \sigma_B$, then the range of $\mathcal{L}(s)$ includes $(0, 1)$.*
- *If $\sigma_A < \sigma_B$, then the range of $\mathcal{L}(s)$ includes $(0, 1)$.*
- *If $\sigma_A > \sigma_B$, then the range of $\mathcal{L}(s)$ includes $(1, +\infty)$.*

The derivation is provided in Appendix D. The range of $\mathcal{L}(s)$ determines how variation in s can be used to trace out preference intensity. Suppose a student, who perceives the outside option as yielding zero utility, is choosing between colleges A and B for a single available spot. Then, the student chooses A if and only if:

$$p_A(s)u_A \geq p_B(s)u_B \quad \iff \quad \mathcal{L}(s) \geq \frac{u_B}{u_A}.$$

In the language of Agarwal and Somaini (2018), the wider the range of $\mathcal{L}(s)$, the larger the "cone" one can construct to trace out the distribution of preference intensity for college A relative to B. Observation 2 provides a rough lower bound on how large the range of $\mathcal{L}(s)$ can be.

Given the fitted unconditional probabilities described in Section 3, for two colleges of comparable quality, it is typically the case that $\sigma_A \approx \sigma_B$. This implies that $\mathcal{L}(s)$ generally spans the interval $(0, 1)$, making this variation especially well-suited for identifying the intensity of vertical preferences.

6.2 Setup

Structure of College Preferences We parameterize college preferences in order to implement the structural estimation. For individual i, the utility of being admitted to college j is given by:

$$\begin{aligned} u_{ij} = {} & \theta^X X_j + \theta^{SX} SES_i X_j + \theta^{CX} CEE_i X_j + (\eta_i^X + \kappa\nu_i) X_j \\ & + f_C(C_j, SES_i, CEE_i) + O_i(SES_i, CEE_i) + \epsilon_{ij}, \end{aligned} \tag{7}$$

where X_j is a 10-dimensional vector of characteristics of college j. This includes the standardized competitiveness of the college, denoted C_j. The variable SES_i is the standardized average educational attainment among adults in student i's township of residence, and CEE_i is the standardized CEE score.

The function $f_C(\cdot)$ is a polynomial that captures differential preferences for competitive colleges based on socioeconomic background and academic achievement. The term $O_i(SES_i, CEE_i)$ is a polynomial that allows students with different backgrounds and test performance to have heterogeneous outside options.

The parameter θ^X, θ^{SX}, θ^{CX} are 10-dimensional vectors. The term η_i^X follows a 10-dimensional normal distribution, with independently distributed elements and variances given by the vector σ_X^2, and ν_i is a scalar component capturing correlated preferences across characteristics, scaled by a 10-dimensional vector κ. The term ϵ_{ij} represents an individual-college-specific shock to utility.

Further details on variables and model specification are provided in Appendix C. Given this specification, we simulate utility and generate predictions for a given set of parameters and decision rules.

Specification 1: All Are Optimal Agents. In this specification, we assume all students follow the Optimal Decision Rule. Utility is simulated using the specification described above, and students' predicted ROLs are generated by solving for the utility-maximizing list. The model includes a total of 60 parameters.

Specification 2: Mixture of Optimal and DC Agents with Identical Preferences. This specification allows for heterogeneity in decision rules: each student may follow either the Optimal Rule or the DC Rule. The probability that a student is a DC Type is modeled as a function of her SES and CEE score, using the estimates obtained in Section 5. Predicted ROLs are generated conditional on the student's type and utility function of the same form as Specification 1, except for the outside option mean-shifter, which—representing the cognitive default for DC Types—is allowed to differ from that of Optimal Types. This specification includes a total of 61 parameters.

Specification 3: Mixture of Optimal and DC Agents with Different Preferences. As in Specification 2, students may follow either the Optimal or DC Rule, and the probability of being a DC Type is determined by SES and CEE scores using estimates from Section 5. Unlike Specification 2, however, we allow preferences to differ across types. Specifically, the parameters θ^X, θ^{SX}, θ^{CX}, the polynomial function $f_C(\cdot)$, and the outside option function

$O_i(\cdot)$ are all type-specific. This specification includes 99 parameters in total.

6.3 Estimation Strategy

The estimation strategy is the Simulated Method of Moments (SMM).[24] The estimation uses 160 moments, which fall into two groups.

The first group contains 40 moments designed to identify preference parameters related to X_j. These include:

- The first 10 moments are the means of X_j across all choices on students' ROLs, used to identify the parameters θ_j.
- The next 10 moments are the means of $X_j \cdot SES_i$, used to identify θ_j^{SX}.
- The third set of 10 moments are the means of $X_j \cdot CEE_i$, used to identify θ_j^{CX}.
- The final 10 moments are the means of X_{ij}^2, where X_{ij} represents the average characteristics of all colleges on student i's ROL. These moments help identify the variance parameters σ_X^2 and the correlation scaling factor κ.

The second group consists of 120 moments, constructed to differentiate DC Types from Optimal Types for any given preference configuration. These moments are motivated by the reduced-form predictions introduced earlier.

To detect top-choice cautiousness (Prediction 1), we include moments such as the mean assignment probabilities of the first, second, third, and fourth choices on students' ROLs. To capture competitiveness reversals (Prediction 2), we compute (i) the share of ROLs that exhibit any reversal in competitiveness ranking, and (ii) the share of ROLs in which the probability of admission for a higher-ranked college exceeds that of a lower-ranked college by at least 25%.

There are thus six types of moments in total. To account for heterogeneity, we divide the sample into 20 subgroups based on a cross-classification of SES quartiles (4 groups) and CEE score quintiles (5 groups). We compute the mean of each of the six moment types within each subgroup, yielding $4 \times 5 \times 6 = 120$ moments.

We use the Simulated Method of Moments (SMM) to estimate the model. For each student, we simulate choice behavior 50 times and compute the simulated moments by

[24] An alternative is to use Maximum Likelihood Estimation (MLE) with the computational techniques proposed in Larroucau and Rios (2018). The main advantage of SMM is its intuitive alignment with the empirical patterns the model aims to match, though this comes at the cost of potential information loss relative to MLE.

averaging across these simulation draws. The estimation procedure minimizes the weighted distance between the simulated moments $m(\theta)$ and the empirical moments m_0:

$$\min_{\theta} \, (m(\theta) - m_0)' W (m(\theta) - m_0),$$

where W is the weighting matrix. The estimator is asymptotically normal, with an estimated variance given by:

$$(\hat{G}' W \hat{G})^{-1} \left[\hat{G}' W \left(1 + \frac{1}{50}\right) \left(\frac{\hat{\Omega}}{N}\right) W \hat{G} \right] (\hat{G}' W \hat{G})^{-1},$$

where 50 corresponds to the number of simulated draws per observation (Laibson et al., 2007; DellaVigna et al., 2016). Here, $\hat{G} \equiv \frac{\partial m(\theta)'}{\partial \theta}$ is the Jacobian of the moment vector, and $\hat{\Omega} = \mathrm{Var}(m(\hat{\theta}))$ is the estimated variance of the empirical moments.

To improve efficiency, we set the weighting matrix W to be the inverse of a diagonal matrix whose entries correspond to the estimated variance components in $\hat{\Omega}$.

6.4 Estimation Results

Since detailed township-level SES data are only available for students in 2015, 2017, and 2018, we restrict the estimation sample to these three cohorts. For students in 2014 and 2016, only county-level SES data are available, and these cohorts are used for out-of-sample validation.

The odd-numbered columns in Panel A of Table 4 report the in-sample model fit across the three specifications. The introduction of DC Types in Specification 2 leads to a substantial improvement in fit—over 40% better compared to Specification 1, which assumes all students follow the Optimal Rule. Allowing preferences to differ across types in Specification 3 yields further gains in fit, but the incremental improvement is modest relative to the impact of introducing the DC Type alone.

The even-numbered columns show the out-of-sample prediction results. Here too, Specification 2 significantly outperforms Specification 1, improving fit by approximately 20%. This highlights the robustness of the DC Type's explanatory power even outside the estimation sample.

Figure 5 illustrates the extent of improvement in predicting the second group of moments, which are designed to capture the presence of DC Types. Panel 5a shows that prediction accuracy for average unconditional assignment probabilities improves most notably for first-choice colleges, while the improvements for third and fourth choices are minimal. This pattern is consistent with the prediction of top-choice cautiousness. Panel 5b presents the improvement in predicting the share of ROLs exhibiting competitiveness reversals. The model's fit improves across all reversal measures, with larger gains observed for more severe reversals—again consistent with the theoretical prediction of competitiveness reversal behavior under the DC Rule.

Panel B of Table 4 summarizes the simulated utilities generated using the fitted parameters from Specification 3. In Columns 1–4, we compute the mean and interquartile range (75th minus 25th percentile) of simulated utility across all students and colleges, repeated over 200 simulation draws. We then regress these statistics on two key characteristics of colleges: competitiveness and distance, using ordinary least squares (OLS).

For the median utility reported in Columns 1 and 3, the preference for more competitive colleges emerges as the dominant explanatory factor—unsurprisingly—far outweighing the role of distance. In fact, the effect of competitiveness is more than an order of magnitude larger than that of distance, which is typically considered important in other contexts. This may be due to competitiveness functioning as a market price that already incorporates geographic desirability, as well as the widespread preference in China for attending college in coastal regions where employment opportunities are more abundant. Notably, simulated DC Types exhibit slightly stronger preferences for competitive colleges than their Optimal Type counterparts.

The college-level analysis is less precise due to the aggregation of simulated utilities at the college level, resulting in only 239 observations. However, summary analyses based on student-level utility yield consistent patterns. Quantitatively, the estimated preference for competitiveness is approximately 24 times stronger than the preference for proximity, with more advantaged students showing relatively weaker preferences for competitive colleges.

Columns 2, 4, and 6 show that preference heterogeneity is greater for both more competitive and more distant colleges, suggesting greater dispersion in perceived value across students for these attributes.

We additionally analyze the satisfaction data from the survey experiment introduced in Section 5, as a complementary approach to quantify the relative preference intensity for college selectivity and distance. Compared to Column 5 in Panel B of Table 4, the only difference is that the dependent variable is stated satisfaction rather than simulated utility.

The estimates exhibit consistent signs: applicants' stated satisfaction increases more strongly with competitiveness than with proximity. This reinforces the finding that college selectivity plays a dominant role in shaping perceived value. Interaction terms with SES and CEE appear to play a minor role. If anything, students from more advantaged backgrounds tend to exhibit slightly weaker preferences for highly competitive colleges.

6.5 Welfare

The estimated preferences allow us to evaluate the potential welfare implications of a counterfactual in which all DC-Type students are de-biased and behave optimally. Ex ante, the effect of such a change is theoretically ambiguous. On the one hand, the reallocation may worsen outcomes for some students, as reshuffling can lead to displacement in competitive assignments. On the other hand, because DC-Type students are not optimally expressing their preferences, they may fail to fully benefit from the strategic advantages offered by such mechanisms (Abdulkadiroğlu et al., 2011; Chen and Kesten, 2017). Thus, whether de-biasing improves welfare overall—or specifically benefits disadvantaged students—is ultimately an empirical question.

Table 5 reports the average change in welfare across all SES $\times$ CEE cells. Since utility in our framework is invariant to affine transformations, we normalize utility such that the best college available to each student in each simulation is assigned a value of 1, while the outside option is normalized to 0. The results show that de-biasing leads to welfare improvements across all SES $\times$ CEE groups. The gains are particularly pronounced among students with high academic performance but low socioeconomic status. For example, students in the top 20% of the CEE distribution who fall in the lowest SES quartile experience an average welfare gain of 6.68%. In contrast, students in the bottom 20% of the CEE distribution but in the highest SES quartile see a more modest average gain of 1.01%. These results are visualized in Figure 6a.

The rationale behind these results is intuitive. Students with lower CEE scores face inherently limited choice sets, as they are ineligible for many selective colleges. As a result, improvements in decision quality—such as correcting for DC-Type behavior—have relatively little impact on their final outcomes. In contrast, students from lower SES backgrounds are more likely to be classified as DC Types, according to our survey-based estimates. De-biasing thus directly improves the decision quality for this group, leading to greater welfare gains. As shown in Table B5 and visualized in Figure 6b, de-biasing almost entirely eliminates the welfare gap between DC Types and Optimal Types, despite the underlying preference differences across these groups.

7 Conclusion and Further Discussions

This paper examines high-stakes sequential decision-making in centralized college admissions, revealing that many students neglect the option value when deciding on their rank-order lists for college preferences. While sequential optimization that accounts for option value should encourage riskier top choices and a more diversified portfolio, our field experimental evidence shows that approximately 50% of applicants—especially those from disadvantaged backgrounds—consistently make overly cautious decisions while simultaneously failing to include safer backup options. Structural estimation and counterfactual analyses further demonstrate that such option value neglect, driven by directed cognition, significantly undermines both equity and efficiency in the admissions process, suggesting that de-biasing interventions could substantially reduce outcome gaps and enhance overall welfare.

In the remainder of this section, we discuss several issues pertinent to the interpretation of our findings.

7.1 Alternative Modeling Considerations

We discuss the assumptions made in our analysis and the extent to which our findings regarding the DC type and welfare implications are affected by them.

Partial Neglect We believe the current approach is a first step in systematically tackling such problems. The best formulation of option value neglect may often be partial and differ across contexts and mechanisms. One parametric form of partial neglect could involve a convex combination between the correct option value and a perceived outside option. Another parametric form could involve level-k style reasoning, where decision-makers can only foresee the impact of a limited number of lower-ranked choices. Our current formula, though extreme, makes it relatively easier to characterize the behavior of applicants who neglect option value and to understand the allocation effects between optimal and non-optimal students. It is also supported by the survey experiment, which finds that many students appear extremely cautious in the college choice questions.

Beliefs about Admission Probability In our analysis of administrative data, we use the estimated probability of admission as a proxy for students' beliefs because, in our context, past cutoffs are publicly available, appear quite predictable across years, and are the most familiar characteristic of the schools. Our survey experiment provides two ways to examine whether the observed risk-taking in rank-order lists can be explained by mistaken beliefs. First, the questions that test the framing effect are not prone to biases in perceived admission probability. Second, in the survey we also elicit students' beliefs about the admission probabilities of their choices. Table B6 shows that, in terms of subjective beliefs, students with higher SES exhibit less top-choice cautiousness (Columns 1 and 2) and fewer risk-taking reversals (Columns 5 and 6). At the same time, their bottom choices do not exhibit significant differences (Columns 3 and 4).

Preferences for Majors The centralized system assigns students at the college level without processing students' expressed major preferences. Students' choices over colleges can be regarded as choices over a set of majors within that college that they could possibly pursue in the future. For those who regard major as the most important consideration, there is an incentive to behave cautiously for their top choices, hoping that, conditional on being admitted to their top choices, their score will out-compete those of other admitted students in major assignment. Using data from the survey experiment, in Table B7 we explore this consideration. We find that only 13.17% of the students regard major as their top consideration, and this share is not significantly different across students of different SES. Moreover, among those who regard major as the most important factor, the estimated admission probability of their top choices is marginally higher than for those who do not. While the perceived admission probability is significantly higher, the resulting impact on overall top-choice cautiousness is merely around 1%. We conclude that the consideration of major is quantitatively irrelevant for the cautiousness gap between advantaged and disadvantaged students.

7.2 Cognitive Limitation vs. Exotic Preferences

To construct the best rank-order lists, student applicants need to process and make choices regarding uncertainty that is resolved in multiple stages. The nature of such decision problems makes several types of models applicable, as discussed below.

Anticipatory Utility / Expectation-Based Reference Dependence As the resolution of uncertainty occurs several weeks after the submission of rank-order lists, students who wish to avoid disappointment might behave cautiously for their top choices. However, subjects' responses in the survey experiment show that this cautiousness for top choices largely disappears when they choose lottery equivalents, where preference-based models predict identical behavior.

Ambiguity Aversion Rank-order list choice can be viewed as a compound lottery problem. The established association between ambiguity aversion and the failure to reduce compound lotteries suggests that top-choice cautiousness may be a manifestation of ambiguity preferences. However, our survey experiment shows that such top-choice cautiousness persists even when uncertainty in later stages (i.e., non-top choices) is eliminated.

Salience The Model of Directed Cognition could be viewed as an example of distortion in decision weights in a particular environment, where the outcomes of non-focused choices receive less attention due to decreased prominence in the language of Bordalo et al. (2022). The evidence, taken as a whole, seems consistent with this broader interpretation.

7.3 Implications for School Application Systems

A large literature documents that disadvantaged students select worse schools in the presence of school choice (Hastings and Weinstein, 2008; Hoxby and Avery, 2012; Walters, 2018). The main contribution of this paper is to establish in our context that much of this inequality in choices can be explained by inequality in decision quality. The finding is potentially generalizable because the neglect of option value, as the psychological underpinning, is applicable to other contexts and mechanisms. We demonstrate that the Model of Directed Cognition captures the psychology well and has great portability in preference estimation and counterfactual analysis. In our case, behavioral biases not only result in increased inequality but also decrease allocation efficiency. We hope that future studies can apply such models to other contexts to evaluate their relevance under different mechanisms.

References

Abdulkadiroğlu, Atila, Yeon-Koo Che, and Yosuke Yasuda (2011) "Resolving conflicting preferences in school choice: The" boston mechanism" reconsidered," *American Economic Review*, 101 (1), 399–410.

Abebe, Girum, A Stefano Caria, and Esteban Ortiz-Ospina (2021) "The selection of talent: Experimental and structural evidence from ethiopia," *American Economic Review*, 111 (6), 1757–1806.

Agarwal, Nikhil and Paulo Somaini (2018) "Demand Analysis using Strategic Reports: An application to a school choice mechanism," *Econometrica*, 86 (2), 391–444.

——— (2019) "Revealed Preference Analysis of School Choice Models," *Annual Review of Economics*, 12.

Ajayi, Kehinde F (2013) "School choice and educational mobility: Lessons from secondary school applications in Ghana," *Institute for Economic Development Working Paper*, 259.

Allcott, Hunt, Benjamin B Lockwood, and Dmitry Taubinsky (2019) "Regressive sin taxes, with an application to the optimal soda tax," *The Quarterly Journal of Economics*, 134 (3), 1557–1626.

Allcott, Hunt and Dmitry Taubinsky (2015) "Evaluating behaviorally motivated policy: Experimental evidence from the lightbulb market," *American Economic Review*, 105 (8), 2501–2538.

Arteaga, Felipe, Adam J Kapor, Christopher A Neilson, and Seth D Zimmerman (2021) "Smart matching platforms and heterogeneous beliefs in centralized school choice,"Technical report, National Bureau of Economic Research.

Artemov, Georgy, Yeon-Koo Che, and Yinghua He (2017) "Strategic 'mistakes': Implications for market design research," *NBER Working Paper*.

Bhargava, Saurabh, George Loewenstein, and Justin Sydnor (2017) "Choose to lose: Health plan choices from a menu with dominated option," *The Quarterly Journal of Economics*, 132 (3), 1319–1372.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer (2022) "Salience," *Annual Review of Economics*, 14 (1), 521–544.

Calsamiglia, Caterina, Chao Fu, and Maia Güell (2020) "Structural estimation of a model of school choices: The Boston mechanism versus its alternatives," *Journal of Political Economy*, 128 (2), 642–680.

Camerer, Colin F, Eric Johnson, Talia Rymon, and Sankar Sen (1993) "Cognition and framing in sequential bargaining for gains and losses," *Frontiers of game theory*, 104, 27–47.

Campbell, John Y (2016) "Restoring rational choice: The challenge of consumer financial regulation," *American Economic Review*, 106 (5), 1–30.

Chade, Hector and Lones Smith (2006) "Simultaneous search," *Econometrica*, 74 (5), 1293–1307.

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, 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.

De Haan, Monique, Pieter A Gautier, Hessel Oosterbeek, and Bas Van der Klaauw (2015) "The performance of school assignment mechanisms in practice."

——— (2023) "The performance of school assignment mechanisms in practice," *Journal of Political Economy*, 131 (2), 388–455.

DellaVigna, Stefano (2018) "Structural behavioral economics," in *Handbook of Behavioral Economics: Applications and Foundations 1*, 1, 613–723: Elsevier.

DellaVigna, Stefano, John A List, Ulrike Malmendier, and Gautam Rao (2016) "Voting to tell others," *The Review of Economic Studies*, 84 (1), 143–181.

Dreyfuss, Bnaya, Ori Heffetz, and Matthew Rabin (2019) "Expectations-based loss aversion may help explain seemingly dominated choices in strategy-proof mechanisms,"Technical report, National Bureau of Economic Research.

Enke, Benjamin and Thomas Graeber (2023) "Cognitive uncertainty," *The Quarterly Journal of Economics*, 138 (4), 2021–2067.

Gabaix, Xavier and David Laibson (2005) "Bounded rationality and directed cognition," *Harvard University*, 20.

Gabaix, Xavier, David Laibson, Guillermo Moloche, and Stephen Weinberg (2006) "Costly information acquisition: Experimental analysis of a boundedly rational model," *American Economic Review*, 96 (4), 1043–1068.

Hassidim, Avinatan, Assaf Romm, and Ran I Shorrer (2016) """ Strategic" Behavior in a Strategy-Proof Environment," in *Proceedings of the 2016 ACM Conference on Economics and Computation*, 763–764.

Hastings, Justine S, Christopher A Neilson, and Seth D Zimmerman (2013) "Are some degrees worth more than others? Evidence from college admission cutoffs in Chile,"Technical report, National Bureau of Economic Research.

Hastings, Justine S and Jeffrey M Weinstein (2008) "Information, school choice, and academic achievement: Evidence from two experiments," *The Quarterly journal of economics*, 123 (4), 1373–1414.

Hoxby, Caroline M and Christopher Avery (2012) "The missing" one-offs": The hidden supply of high-achieving, low income students,"Technical report, National Bureau of Economic Research.

Jia, Ruixue and Hongbin Li (2016) "Access to elite education, wage premium, and social mobility: the truth and illusion of China's college entrance exam," *Stanford Center for International Development, Working Paper* (577).

Johnson, Eric J, Colin Camerer, Sankar Sen, and Talia Rymon (2002) "Detecting failures of backward induction: Monitoring information search in sequential bargaining," *Journal of Economic Theory*, 104 (1), 16–47.

Kapor, Adam J, Christopher A Neilson, and Seth D Zimmerman (2020) "Heterogeneous beliefs and school choice mechanisms," *American Economic Review*, 110 (5), 1274–1315.

Laibson, David, Andrea Repetto, and Jeremy Tobacman (2007) "Estimating discount functions with consumption choices over the lifecycle,"Technical report, National Bureau of Economic Research.

Larcom, Shaun, Ferdinand Rauch, and Tim Willems (2017) "The benefits of forced experimentation: Striking evidence from the London underground network," *The Quarterly Journal of Economics*, 132 (4), 2019–2055.

Larroucau, Tomás, Manuel Martínez, Christopher Neilson, and Ignacio Rios (2021) "Application Mistakes and Information frictions in College Admissions."

Larroucau, Tomás and Ignacio Rios (2018) "Do "short-list" students report truthfully? Strategic behavior in the Chilean college admissions problem," *Preprint, submitted September*, 1 (10.13140).

——— (2020) "Dynamic college admissions and the determinants of students' college retention,"Technical report, Technical Report 2020. and,"Do "Short-List" Students Report Truthfully.

Li, Hongbin, Lingsheng Meng, Kai Mu, and Shaoda Wang (2021) "English Language Requirement and Educational Inequality: Evidence of 16 Million College Applicants in China."

Li, Shengwu (2017) "Obviously strategy-proof mechanisms," *American Economic Review*, 107 (11), 3257–87.

Lucas, Adrienne M and Isaac M Mbiti (2012) "The determinants and consequences of school choice errors in Kenya," *American Economic Review*, 102 (3), 283–88.

Martínez-Marquina, Alejandro, Muriel Niederle, and Emanuel Vespa (2019) "Failures in Contingent Reasoning: The Role of Uncertainty," *American Economic Review*, 109 (10), 3437–74.

Oprea, Ryan (2020) "What Makes a Rule Complex?" *American Economic Review*, 110 (12), 3913–51.

——— (2024) "Decisions under risk are decisions under complexity," *American Economic Review*, 114 (12), 3789–3811.

Pathak, Parag A and Tayfun Sönmez (2008) "Leveling the playing field: Sincere and sophisticated players in the Boston mechanism," *American Economic Review*, 98 (4), 1636–52.

Rees-Jones, Alex, Ran Shorrer, and Chloe J Tergiman (2020) "Correlation Neglect in Student-to-School Matching,"Technical report, National Bureau of Economic Research.

Rees-Jones, Alex and Samuel Skowronek (2018) "An experimental investigation of preference misrepresentation in the residency match," *Proceedings of the National Academy of Sciences*, 115 (45), 11471–11476.

Rees-Jones, Alex and Dmitry Taubinsky (2020) "Measuring "schmeduling"," *The Review of Economic Studies*, 87 (5), 2399–2438.

Shorrer, Ran I (2019) "Simultaneous search: Beyond independent successes," in *Proceedings of the 2019 ACM Conference on Economics and Computation*, 347–348.

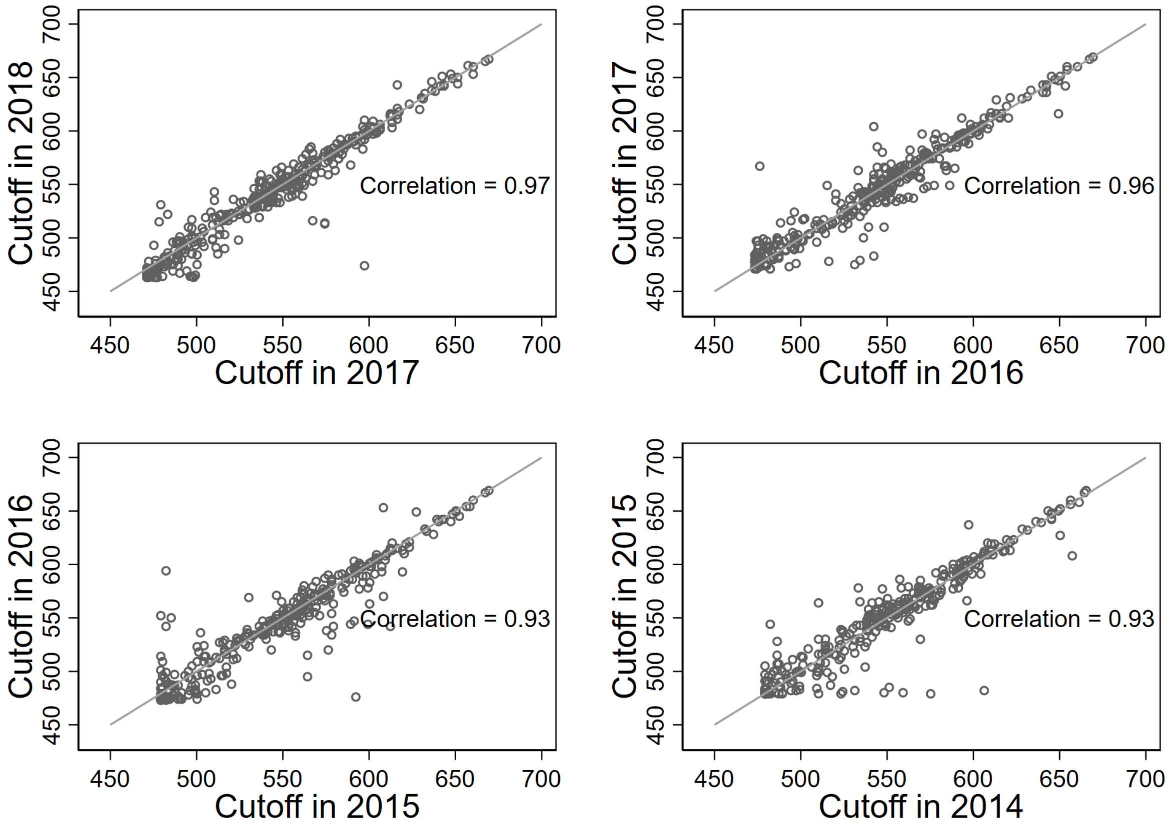
Shorrer, Ran I and Sándor Sóvágó (2018) "Obvious mistakes in a strategically simple college admissions environment: Causes and consequences," *Available at SSRN 2993538.*

Song, Yan, Kentaro Tomoeda, and Xiaoyu Xia (2020) "Sophistication and Cautiousness in College Applications," *Available at SSRN 3711118.*

Walters, Christopher R (2018) "The demand for effective charter schools," *Journal of Political Economy*, 126 (6), 2179–2223.

Wang, Ao, Shaoda Wang, and Xiaoyang Ye (2023) "When information conflicts with obligations: the role of motivated cognition," *The Economic Journal*, 133 (654), 2533–2552.

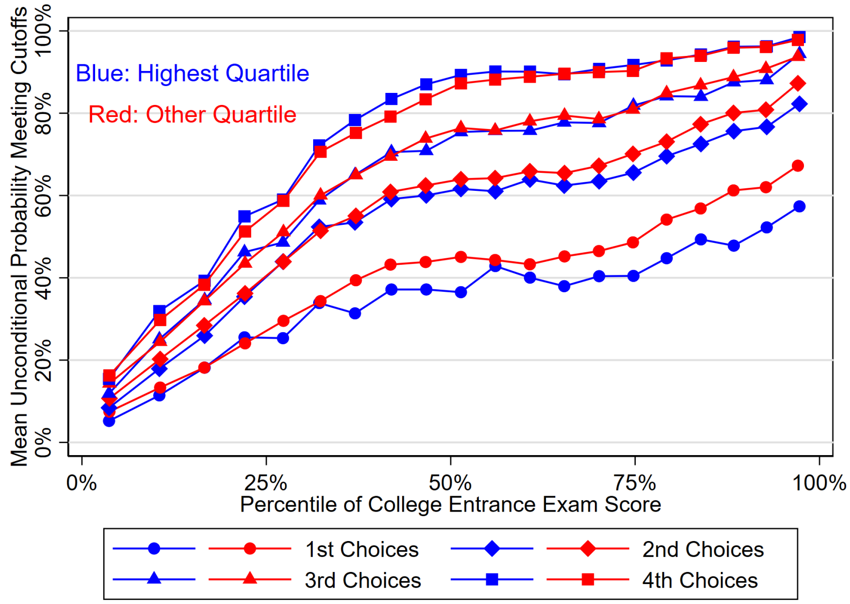
Figure 1: Admission Cutoffs Across Years



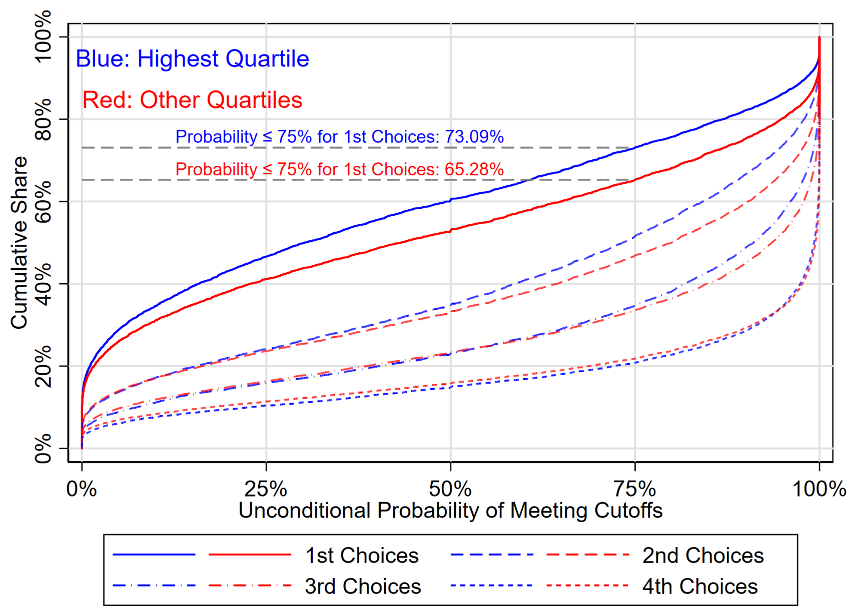
Note: This figure plots the admission cutoffs for all colleges that admit Ningxia students during 2014-2018, for the discussion in Section 3.2. The admission cutoffs are converted to its 2018 rank-preserving equivalents. Each dot represents the admission cutoffs of a specific college in two consecutive years. Top left graph plots the cutoffs in 2018 against 2017. Similarly, top right graph plots the cutoffs in 2017 against 2016; bottom left graph plots the cutoffs in 2016 against 2015; bottom right graph plots the cutoffs in 2015 against 2014.

Figure 2: Mean of Vacancy Probability and Share of Reversals by SES Groups

(a) Mean of Vacancy Probability by CEE Score Percentile



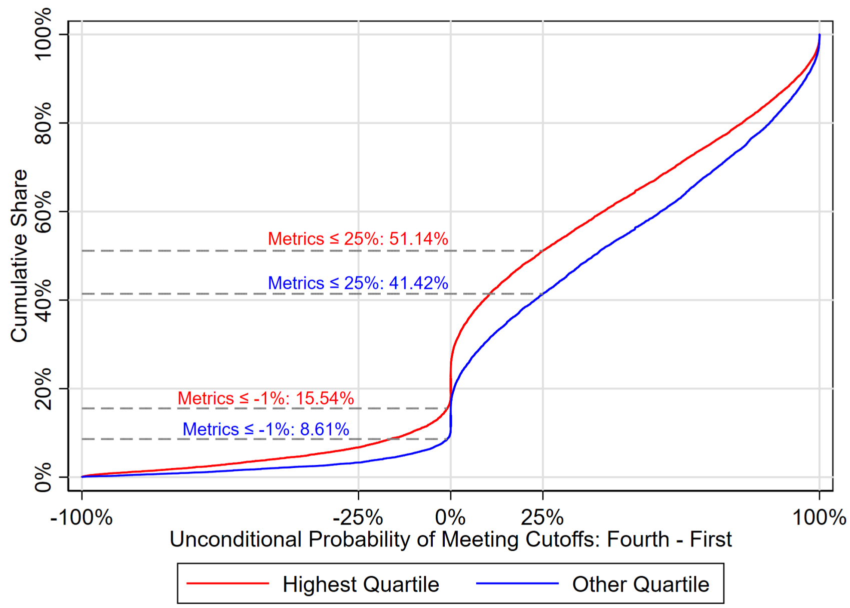
(b) Cumulative Distribution of Unconditional Probability



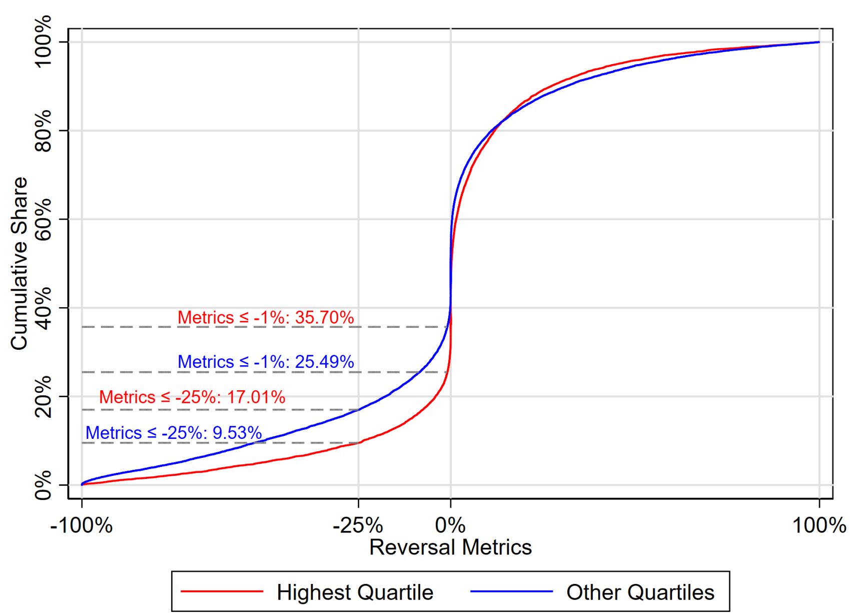
Note: This figure presents statistics for the most advantaged quartile (in blue) and the other quartiles (in red), respectively. Subfigure (a) plots the mean of vacancy probability (i.e. unconditional probability of meeting the cutoff) by choice (first, second, third, fourth) and SES (most advantaged quartile vs. other quartiles). Subfigure (b) plots the cumulative distribution of the vacancy probability i.e. unconditional probability of meeting the cutoff) for by choice (first, second, third, fourth) and SES (most advantaged quartile vs. other quartiles) where the distribution for other quartiles are computed by re-weighting according to CEE Score, to control for the difference in CEE Score between the most advantaged quartile and other quartiles. This figure is first referenced in Section 4.3.

Figure 3: Mean of Vacancy Probability and Share of Reversals by SES Groups

(a) CDF of Fourth-First Choice Gap



(b) CDF of Reversal Metric



Note: This figure presents statistics for the most advantaged quartile (in blue) and the most disadvantaged quartile (in red), respectively. Subfigure (a) plots the cumulative distribution of the vacancy probability (i.e. unconditional admission probability) of the fourth choice minus that of the first choice. Subfigure (b) plots the share of students whose strategy exhibits at least one pair of competitiveness reversals comparing the top three choices to the fourth choice. To quantify such pattern, the statistic we use to classify reversal is $R \equiv min\{p_4 - p_i | \text{for } 1 \leq i \leq 3\}$. The graph plots the cumulative distribution of R. This figure is first referenced in Section 4.4.

Figure 4: Distribution of Response to Incentivized Survey Questions

			Equivalent Lottery: Indifferent Between (25 CNY, 50%; 20 CNY, 50%) (L CNY, 25%; 20 CNY, 75%)		
			Very Cautious L≥50	Not Very Cautious L<50	Total
ROL 1st Choice:Payoff of non-first colleges is 20 CNY Indifferent Between (25 CNY, 50%) (R CNY, 25%)	Very Cautious	R≥50	72 5.1%	435 30.8%	507 35.9%
	Not Very Cautious	R<50	81 5.7%	824 58.4%	905 64.1%
	Total		153 10.8%	1259 89.2%	1412 100%

Note: This figure displays a coarsened joint distribution of responses to Question Group 1 (ROL 1st Choice) and Question Group 2 (Equivalent Lottery), for the discussion in Section 5.

Blue cells indicate that students' responses are both classified as "not very cautious" in the "ROL first-choice" problem and the equivalent lottery problem, as predicted by Rational Decision Rule.

Red cells indicate that students' responses are very cautious in "ROL first-choice" problem but not very cautious in the equivalent lottery problem as predicted by Directed Cognition (DC) Decision Rule.

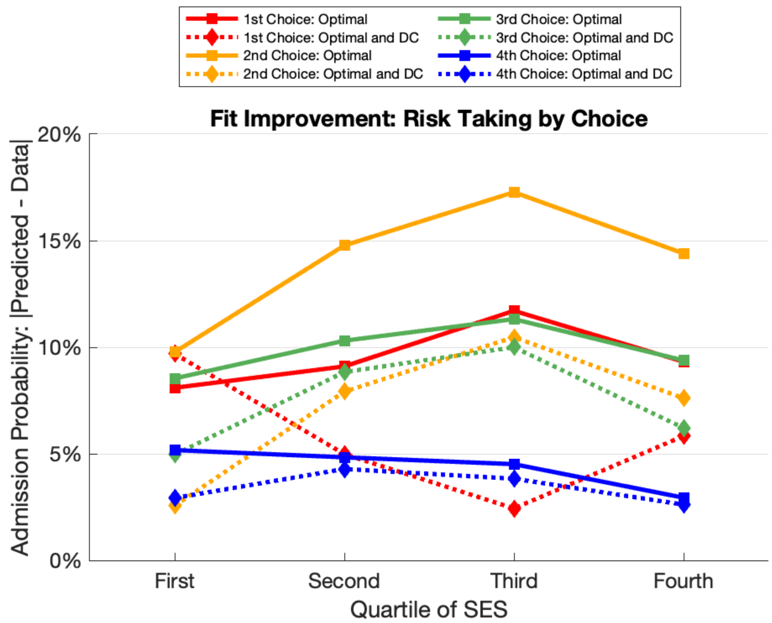
Gray cells indicate that students' responses are not very cautious in "ROL first-choice" problem but very cautious in the equivalent lottery problem, which cannot be predicted by neither decision rule.

"ROL first-choice" question (vertical) corresponds to Question Group 1. In this question, students need to choose from college X, whose admission probability is 50% and payoff of admission is 25 CNY, or college Y, whose admission probability is 25% and payoff of admission is R CNY, where R = 30, 35, 40, 45, 50, 55, 60, respectively in the MPL, and put the college of their choice on the top of their list. If students are not admitted to their first choices, they will be admitted to one of the bottom choices in this scenario, which corresponds to the payoff of 20 CNY.

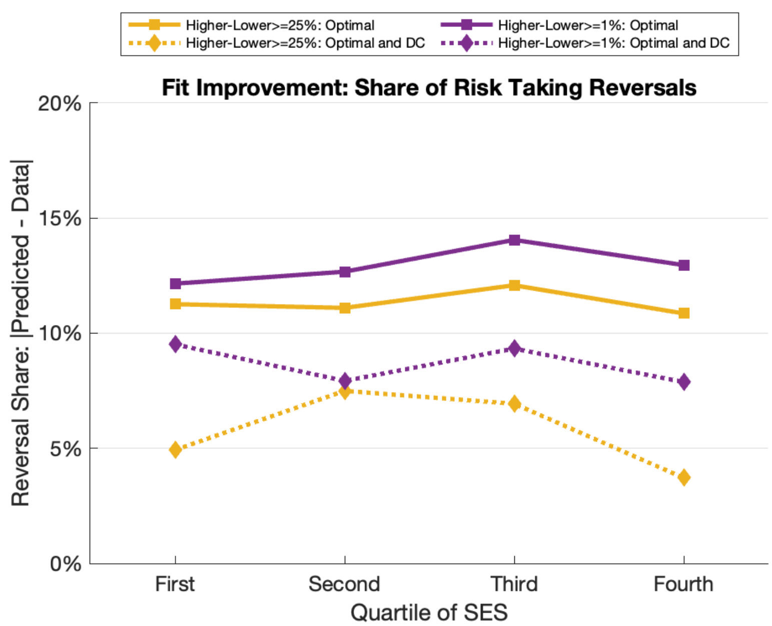
"Equivalent Lottery" question (horizontal) corresponds to Question Group 2. In this question, students need to choose from lottery X, whose payoff is 25 CNY with 50% of chance, 20 CNY with 50% of chance, or lottery Y, whose payoff is L CNY with 25% of chance, 20 CNY with 75% of chance, where L = 30, 35, 40, 45, 50, 55, 60, respectively. Note that "Equivalent Lottery Question" is mathematically equivalent to the aforementioned "ROL first-choice Question".

Figure 5: Fit of Moments by SES Groups

(a) Mean of Probability Probability



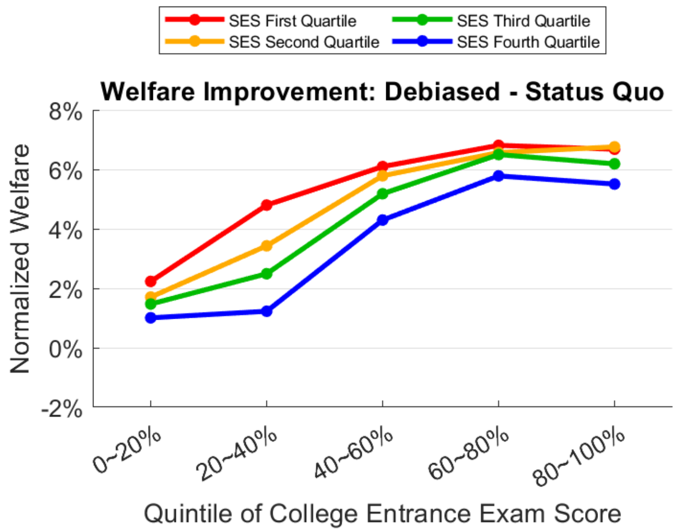
(b) Share of Competitive Reversals



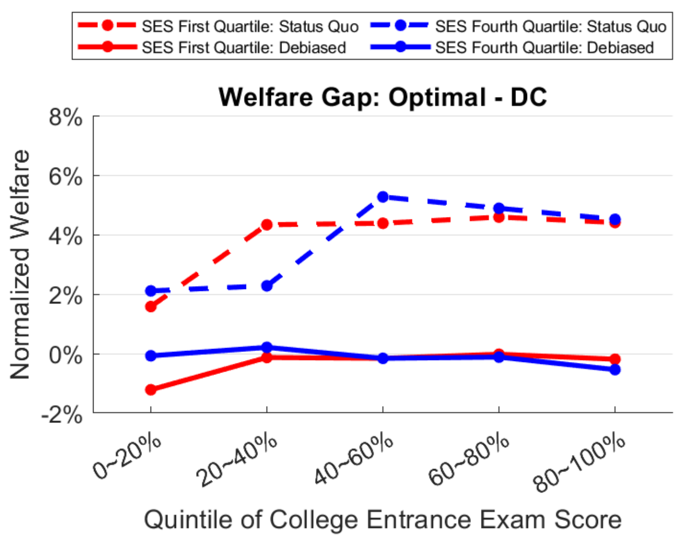
Note: This figure presents the distance (absolute value) between prediction and data for key moments by SES Quartile. The graphs focus on the comparison of the optimal-only model (in solid lines) and the model that allows for the presence of the optimal and the DC type. Subfigure (a) plots the mean of admission probability (unconditional) for the first choice in red, the second choice in orange, the third choice in green, the fourth choice in blue respectively. Subfigure (b) plots the share of students whose strategy exhibits at least one pair of competitiveness reversals anywhere on her ROL, as a function of the threshold above which the reversal is counted in the statistic. The statistic we use to classify reversal is $R \equiv max\{p_i - p_j | \text{for any} 1 \leq j < i \leq 4\}$, only those whose R is above threshold $X\%$ will be counted. Purple lines represent the share where $X = 1$. Yellow lines represent the share where $X = 25$. This figure is referenced in Section 6.

Figure 6: Welfare Analysis by SES and CEE Score

(a) Welfare Gain from Debiasing



(b) Welfare Difference between Optimal and DC:
Status Quo vs. Debiased



Note: this figure presents the key information from welfare analysis detailed in Table 5 and B5. The welfare is normalized under the assumption that for each student, the best first-tier college is 100% and the worst first-tier college is 0%. We calculate the average welfare by students' socioeconomic quartile, as measured by the average adult education attainment at the applicant's township of residence, as well as students' CEE score quintile. Subfigure 6a shows the average welfare gain in debiasing by students' CEE and SES cell. Subfigure 6b showcases the change in welfare gap between the Optimal Type and the DC Type, before (dashed line) and after (solid) debiasing. This figure is referenced in Section 6.

Table 1: Normalized Admission Outcome by SES Quartiles

Panel A: Full Sample						
	STEM & Humanity			STEM Only		
SES:1st Quartile	-0.0305	-0.0111	-0.0081	-0.0298	-0.0109	-0.0040
	(0.0041)	(0.0044)	(0.0052)	(0.0044)	(0.0048)	(0.0057)
SES:4th Quartile	0.0742	0.0588	0.0519	0.0695	0.0549	0.0490
	(0.0042)	(0.0045)	(0.0059)	(0.0048)	(0.0050)	(0.0066)
Point Estimate: 1st - 4th Quartile	-.105	-.07	-.06	-.099	-.066	-.053
Standard Error: 1st - 4th Quartile	(.005)	(.006)	(.008)	(.005)	(.006)	(.008)
CEE Score	Yes	Yes	Yes	Yes	Yes	Yes
Demographic Variables	No	Yes	Yes	No	Yes	Yes
County Fixed Effects	No	No	Yes	No	No	Yes
Number of Observations	44201	44192	44192	35999	35990	35990
R squared	0.875	0.876	0.876	0.874	0.875	0.875
Panel B: CEE Percentile $\geq$ 40%						
	STEM & Humanity			STEM Only		
SES:1st Quartile	-0.0441	-0.0167	-0.0175	-0.0437	-0.0163	-0.0109
	(0.0064)	(0.0069)	(0.0081)	(0.0071)	(0.0076)	(0.0089)
SES:4th Quartile	0.0845	0.0650	0.0550	0.0823	0.0630	0.0538
	(0.0059)	(0.0064)	(0.0086)	(0.0068)	(0.0072)	(0.0096)
Point Estimate: 1st - 4th Quartile	-.129	-.082	-.072	-.126	-.079	-.065
Standard Error: 1st - 4th Quartile	(.007)	(.009)	(.011)	(.008)	(.01)	(.013)
CEE Score	Yes	Yes	Yes	Yes	Yes	Yes
Demographic Variables	No	Yes	Yes	No	Yes	Yes
County Fixed Effects	No	No	Yes	No	No	Yes
Number of Observations	27011	26999	26999	22012	22000	22000
R squared	0.827	0.829	0.830	0.823	0.826	0.827

Note: This table reports the outcome gap between students of different SES Quartiles, where the benchmark is the average outcome of the second and third Quartiles. The outcome is measured by the average normalized competitiveness of the colleges that admit the students in question during 2014 and 2018. For students who were not admitted to any first-tier colleges, their outcome is coded as the lowest value among all first-tier colleges. Panel A analyze the outcome gap without any restriction on students' CEE score. Panel B analyze the outcome gap by restricting the sample to those whose CEE Percentile is 40% or higher. Column 1-3 analyze the gap for both tracks, with progressively longer list of controls. Column 4-6 analyze the gap for the STEM track only, with progressively longer list of controls. CEE Score refers to the fourth-order polynomial of normalized CEE Score. Demographics refers to a list of variables including gender, ethnicity, age, whether students live in rural area, whether students have taken the CEE in previous years. This table is first referenced in Section 4.2.

Table 2: Risk Taking in First and Fourth Choices:
Most Advantaged Quartile vs. Other Quartiles

Unconditional Admission Probability	(1) First	(2) Fourth	(3) Fourth-First
Panel A: Summary Statistics of Admission Probability for CEE Percentile $\geq$ ***40%***			
Mean	49.95%	90.61%	40.29%
25th Percentile	8.56%	95.44%	3.49%
50th Percentile	50.05%	99.81%	37.71%
75th Percentile	90.96%	100.00%	80.65%
Panel B: Highest vs. Other Quartiles - Mean of Gap			
Highest (Most Advantaged) Quartile	-4.90%	1.54%	6.98%
	(0.41%)	(0.34%)	(0.50%)
4th-Order Polynomial of Priority Score	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes
Track (Science or Humanity) Fixed Effects	Yes	Yes	Yes
Panel C: Highest vs. Other Quartiles - Gap by Quantile of Priority Score			
E[Highest-Other|Priority Score 20%]	-2.68%	2.82%	6.04%
	(0.73%)	(0.61%)	(0.89%)
E[Highest-Other|Priority Score 40%]	-4.27%	2.38%	7.01%
	(0.62%)	(0.50%)	(0.75%)
E[Highest-Other|Priority Score 60%]	-5.04%	1.11%	6.85%
	(0.44%)	(0.37%)	(0.54%)
E[Highest-Other|Priority Score 80%]	-6.63%	-0.22%	7.17%
	(0.88%)	(0.72%)	(1.08%)

Note: This table reports the results from analyzing the unconditional admission probability (i.e. the probability of students' scores meeting the cutoff of a college) of students' first choices and fourth choices using administrative data. Column 1 reports statistics related to the unconditional probability of the first choices. Column 2 reports statistics related to the unconditional probability of the fourth choices. Column 3 reports statistics related to the fourth-first choice gap, in terms of unconditional probability. Panel A reports the summary statistics of aforementioned variables. Statistics in Panel B and C are generated by a fixed effect regression that regresses outcome variables on the interaction of an indicator whether students belong to the most advantaged quartile, and a fourth-order polynomial of priority scores. Panel B reports the estimate of the mean behavior gap between the most advantaged quartile and other quartiles in terms of the aforementioned variables. Panel C examines the heterogeneity of this gap by reporting the predicted mean of gap conditional on CEE score percentile. Section 4.3 references this table.

Table 3: Testing Framing Effect - Incentivized Survey Responses

	(1)	(2)	(3)
Panel A: Framing Effect and Socio-Economic Status			
		Prob(Red Block) Very Cautious in ROL But Not Very Cautious in Lottery	
SES Index Normalized	-4.8%	-4.9%	-4.3%
	(1.3%)	(1.6%)	(1.4%)
Control: Priority Score	Yes	Yes	Yes
Control: Demographic Variables	No	Yes	Yes
Control: College Preferences	No	No	Yes
Panel B: Estimated Share of Type from Survey Responses			
Share of Directed Cognition Type: Mean	0.0%	50.4%	50.4%
	-	(1.7%)	(1.7%)
Marginal Effect of Normalized SES	0.0%	-5.3%	-5.5%
	-	(1.8%)	(1.8%)
Marginal Effect of Normalized SES	0.0%	0.6%	0.6%
	-	(1.9%)	(1.8%)
Share of Sincere Type	0.0%	0.0%	0.8%
	-	-	(0.6%)
CRRA Coefficient Mean	0.490	0.240	0.242
	(0.014)	(0.012)	(0.012)
Number of Parameters	6	9	10
Log Likelihood	-9100.574	-8781.778	-8780.909
Number of Observations	1412	1412	1412
Bayesian Information Criterion	18244.66	17628.83	17634.35
Panel C: Framing Effect and Elicited College Application Behavior			
	Not More Cautious in ROL than Lottery (Red Block)	More Cautious in ROL than Lottery (Blue / Gray Block)	Difference
List the 1st Choice If Only One Spot	27.6%	50.8%	23.2%
	(1.4%)	(2.4%)	(2.7%)
List Top Two Choices If Only Two Spots	27.3%	34.5%	7.2%
	(1.4%)	(2.3%)	(2.6%)
Subject Probability of Meeting the Cutoff of the 1st Choice	48.0%	56.6%	8.5%
	(0.9%)	(1.5%)	(1.7%)
Subject Probability of Meeting the Cutoff of the 4th Choice	72.5%	70.6%	-1.8%
	(1.0%)	(1.6%)	(1.9%)
Share of Reversal: Subjective Probability Higher Ranked - Lower Ranked > 25%	18.4%	23.2%	4.8%
	(1.2%)	(2.0%)	(2.3%)

Note: This table reports empirical analysis from incentivized survey response, as described in Section 5. Panel A analyzes to what extent the framing effect, that is, being very cautious in college choice problem than its lottery equivalent, is correlated with students' socioeconomic status and their priority scores, controlling for other variables elicited in the survey. Panel B jointly estimates students' propensity of being a Directed-Cognition Type and their CRRA risk preferences using different specifications. Column (1) excludes any behavioral type. Column (2) estimates a mixture model of the DC and the optimal type. Column (3) estimates a model that additionally allows for sincere type. Panel C summarizes whether the average behavior in reported college choices among students who exhibit framing effect differ from those who do not.

Table 4: Preference Estimation from Administrative Data and Survey

Panel A: Comparison of Different Models						
	(1)	(2)	(3)	(4)	(5)	(6)
	Optimal Only		Optimal and DC Same Preferences		Optimal and DC Different Preferences	
	Fit	Out of Sample	Fit	Out of Sample	Fit	Out of Sample
Distance	8596.39	10313.22	5086.02	8227.27	4440.28	7130.52
# Moments	160		160		160	
# Parameters	50		51		99	

Panel B: Summary of Estimated College Preferences from Administrative Data - Regressing Simulated Utility on Key College and Student Characteristics						
	(1)	(2)	(3)	(4)	(5)	(6)
	College-Level Analysis				Student-Level Analysis	
	Optimal Type		DC Type			
	Median	75th-25th%	Median	75th-25th%	Median	75th-25th%
Quality	100.40	113.83	104.69	112.41	121.60	35.94
	(4.44)	(4.85)	(5.00)	(4.77)	(0.04)	(0.01)
SES * Quality					-15.56	0.06
					(0.04)	(0.01)
CEE * Quality					104.88	0.02
					(0.04)	(0.01)
Distance	-5.90	10.83	-6.88	10.21	-5.36	8.39
	(4.94)	(5.40)	(5.07)	(5.31)	(0.04)	(0.01)
SES * Distance					3.66	0.00
					(0.04)	(0.01)
CEE * Distance					2.05	0.00
					(0.04)	(0.01)

Panel C: Summary of Estimated College Preferences from Survey						
	Quality	SES * Quality	CEE * Quality	Distance	SES * Distance	CEE * Distance
Dependent Stated: Stated Satisfaction (0 100)	9.02	-0.18	0.91	-1.40	0.85	0.28
	(0.25)	(0.17)	(0.17)	(0.16)	(0.17)	(0.18)

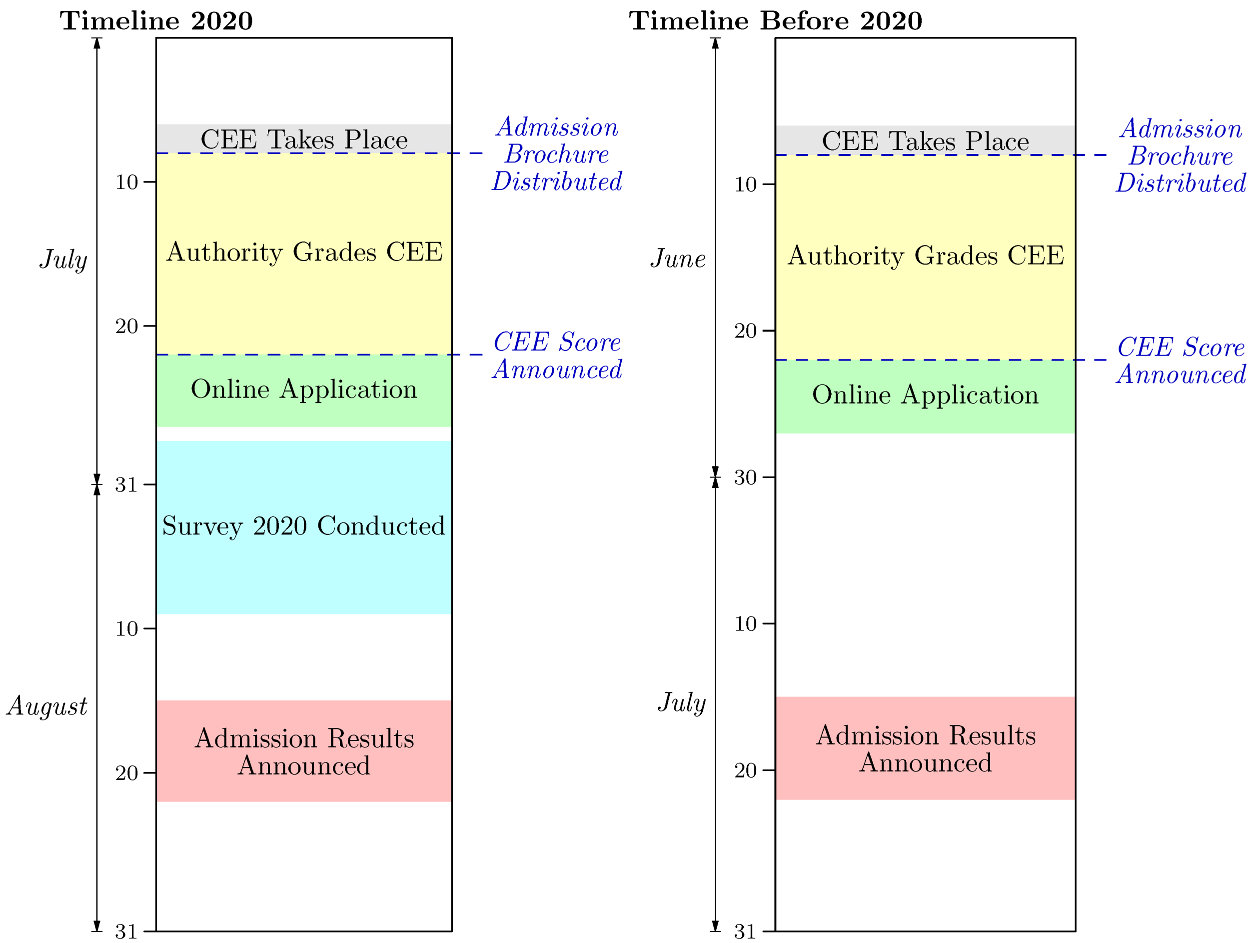
Note: This table reports the results of preference estimation in Sections 6. Panel A Column 1, 3, 5 reports the structural model fit using administrative data in 2015, 2017, 2018. Panel A Column 2, 4, 6 reports the distance between out-of-sample predictions about the moments in year 2014 and 2016 using the aforementioned estimated model. There are three versions of structural model, a model that only allow Optimal Type is present (Column 1, 2), a model that allows for both Optimal and DC Type with same preferences (Column 3, 4), a model that allows for both Optimal and DC Type with different preferences (Column 5, 6). Panel B summarizes how estimated preferences depends on important college characteristics (Competitiveness and Distance) and student characteristics (socio-economic status and college entrance exam score) using the specification that allows Optimal and DC to have different preferences. Column (1), (2), (3), (4) analyzes the median and the gap between 25% and 75% percentile on college level. Column (5) and (6) analyzes the median and the gap between 25% and 75% percentile on student level. Panel C use the same student-level regression to analyze expressed satisfaction (0-100) in the survey experiment. Standard errors are in parentheses.

Table 5: Welfare Calculation

SES Quartile CEE Quintile		0-25%	25-50%	50-75%	75-100%	Aggregate
0-20%	Status Quo	35.04%	41.63%	48.92%	64.90%	44.29%
	Debiased	37.27%	43.33%	50.39%	65.91%	46.04%
	Change	2.23%	1.70%	1.47%	1.01%	1.74%
20-40%	Status Quo	30.74%	30.62%	30.98%	34.15%	31.32%
	Debiased	35.54%	34.05%	33.47%	35.38%	34.64%
	Change	4.80%	3.43%	2.49%	1.23%	3.32%
40-60%	Status Quo	34.92%	33.66%	33.05%	33.36%	33.89%
	Debiased	41.02%	39.45%	38.23%	37.65%	39.33%
	Change	6.09%	5.79%	5.18%	4.30%	5.44%
60-80%	Status Quo	37.87%	37.26%	36.22%	36.71%	37.05%
	Debiased	44.68%	43.83%	42.72%	42.49%	43.47%
	Change	6.81%	6.56%	6.50%	5.78%	6.42%
80-100%	Status Quo	43.23%	43.59%	44.25%	47.21%	45.24%
	Debiased	49.91%	50.34%	50.43%	52.71%	51.31%
	Change	6.68%	6.76%	6.19%	5.51%	6.08%
Aggregate	Status Quo	35.57%	37.08%	38.81%	43.20%	38.48%
	Debiased	40.58%	41.60%	42.98%	47.40%	42.99%
	Change	5.01%	4.52%	4.17%	4.20%	4.51%

Note: this table presents the average welfare by student subgroups. The welfare is normalized by the assumption that the best first-tier college for each student is 100%, and the worse first-tier college is 0%. Based on students' exam score in the CEE, we divide them into five groups: 0-20%, 20-40%, 40-60%, 60-80%, 80-100% respectively. Higher quintile means the exam performance was better. The results of each CEE group are presented in rows. Based on the student applicants' township of residence, we divide them into four groups according to average adult education attainment in the corresponding township: 0-25%, 25-50%, 50-75%, 75-100%. Higher quartile means the socioeconomic status was better. The results of each SES group are presented in columns. The table therefore calculate the average welfare for 5*4 cells of student groups. The total welfare by aggregating along the dimension of CEE is shown at the bottom of the table. The total welfare by aggregating along the dimension of SES is shown to the rightmost column of the table. The aggregated welfare along both dimensions is presented at the bottom right corner of the table. For each subgroup in question, we calculate three numbers. "Status Quo" refers to the average welfare if no intervention was conducted. "Debiased" refers the average welfare if all DC students played optimally in the application system. "Change" refers to the gain in average welfare if we debiased all the DC students.

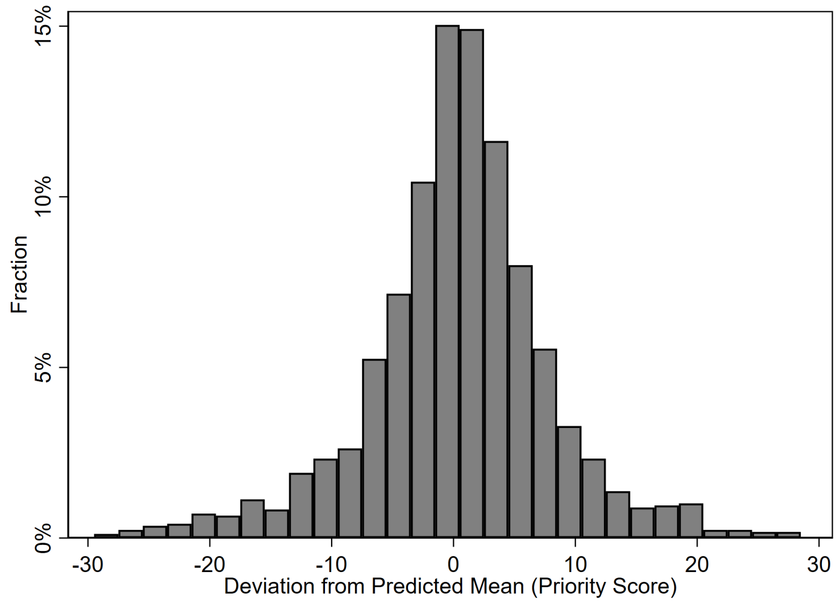
A Supplementary Figures



Note: This figure presents the timeline of college admission process, as detailed in Section 2. Note that the timeline remains unchanged until 2020, when the college entrance exam (CEE) was postponed by exactly one month due to COVID-19. As a result, all the admission procedures thereafter were postponed by exactly one month as well. As shown in the 2020 timeline, the survey was conducted right after application deadline, but before students were informed of their admission outcomes.

Figure A1: Timeline of College Admission Process

Figure A2: Distribution of Realized Cutoff-Predicted Mean Differences

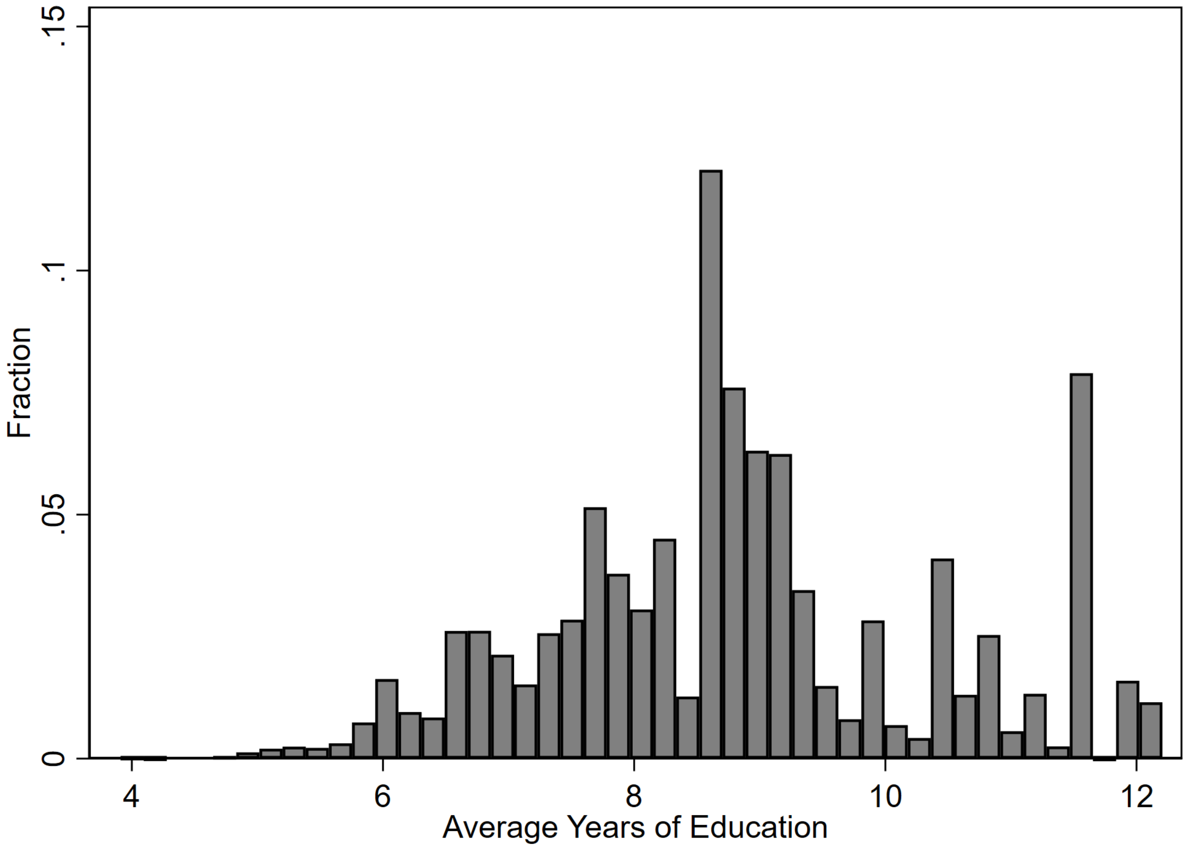


(a) Distribution of Differences: All Colleges

(b) Distribution of Differences by College Competitiveness

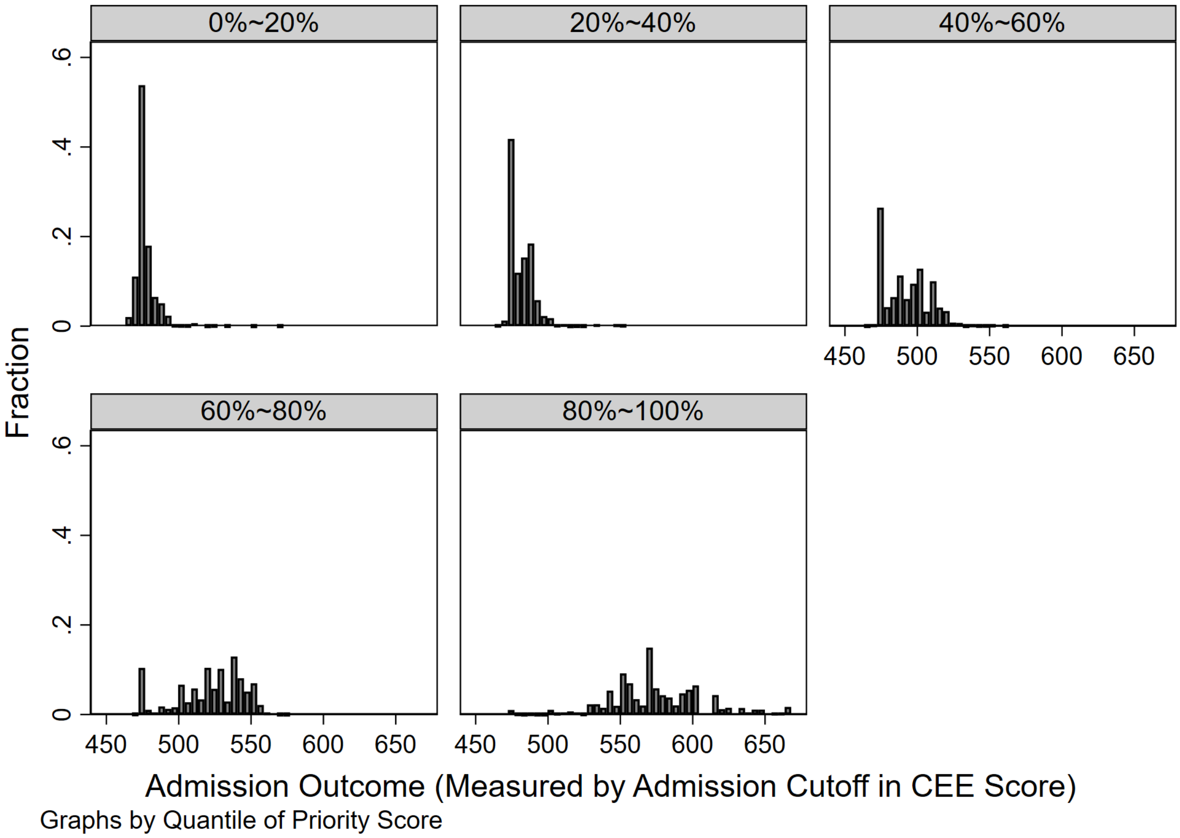
Note: This figure plots the distribution of distance between realized cutoffs and predicted mean as defined in Section 3.2. Subfigure (a) plots the distribution of differences between admission cutoffs and predicted mean among all colleges. Subfigure (b) plots the distribution of differences by college competitiveness, where the dashed lines in the top left, top right, bottom left, and bottom right graph are the empirical distribution for the least competitive quarter of colleges, second from the least competitive quarter, second from the most competitive quarter, and the most competitive quarter of colleges, respectively. The solid lines in subfigure (b) are fitted distributions using the median estimate of the standard deviation of cutoff-mean difference for corresponding quarter of colleges. Both subfigures omit outliers which is more than 30 points away from the predicted mean.

Figure A3: Average Years of Education in Administrative Data (Township Level)



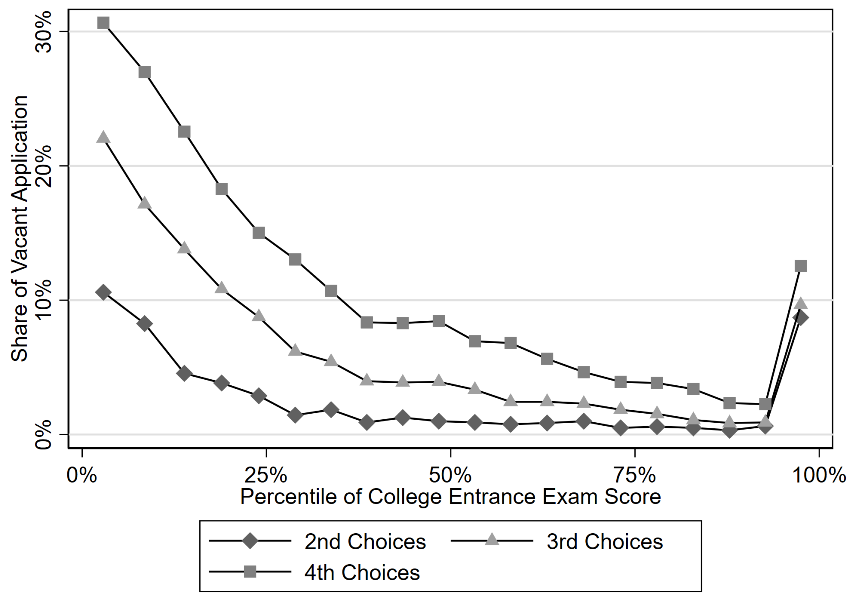
Note: This figure plots the cumulative distribution of probability of meeting cutoffs for students' first (in red), second (in yellow), third (in green) and fourth (in blue) choices, respectively, during 2014-2018. The figure demonstrates sizable heterogeneity in terms of educational attainment across different townships in Ningxia (25th percentile: 7.80 years; median: 8.73 years; 75th percentile: 9.88 years). This figure is first referenced in Section 4.2.

Figure A4: Competitiveness of Admitting College: by CEE Score Quintile

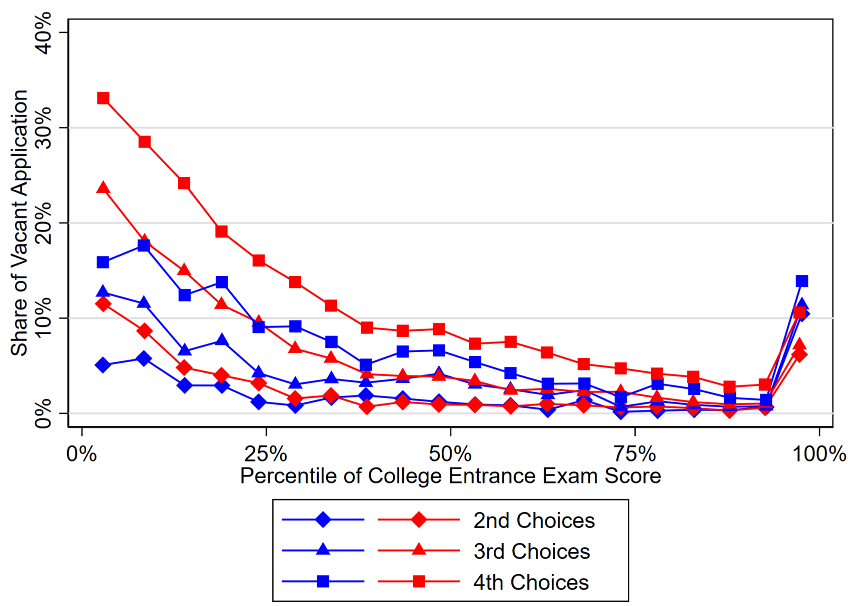


Note: This figure plots the distribution of colleges that admit students during 2014-2018. We split the entire sample of STEM applicants into five groups according to their CEE Score, with the first quintile being the group with lowest CEE Score and the fifth being the highest. This graph is first referenced in Section 4.2.

Figure A5: Share of Vacant Application for Non-Top Choices



(a) All Students

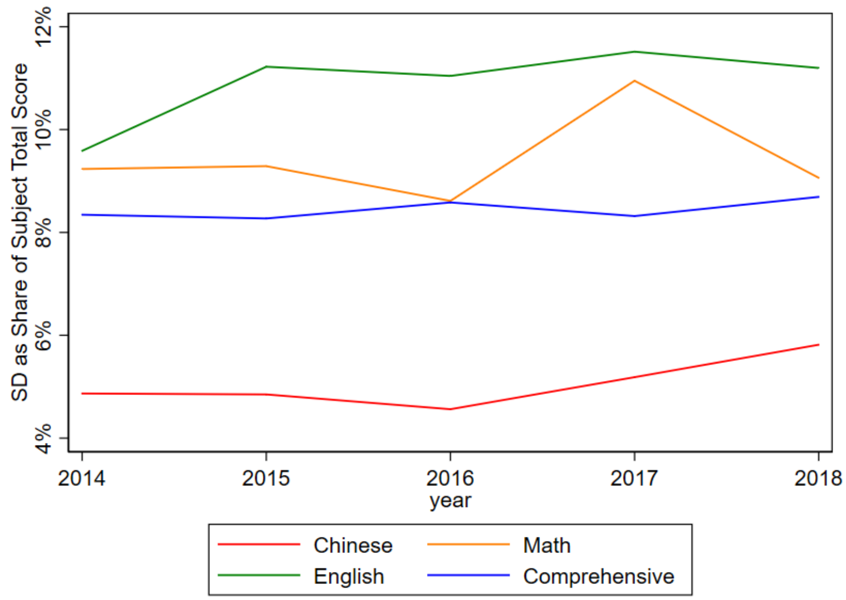


(b) Highest SES Quartile vs. Other SES Quartiles

Note: This figure plots the share of vacant applications among first-tier college eligible applicants for their non-top choices. Subfigure (a) presents the bin-scatter plot of the share for all students conditional on their percentile in college entrance exam score. Subfigure (b) plot the same statistics in Subfigure (a) for the highest SES quartile (in blue) versus other quartiles (in red). The diamonds are the share for vacant second choices. The triangles are the share for vacant third choices. The squares are the share for vacant fourth choices.

Figure A6: Impact of Subject Difficulty Variation on Admission Probability

(a) Difficulty of Subjects Varies Across Years



Note: This figure explains the source of exogenous variation in admission probability. Subfigure (a) presents the standard deviation of raw exam scores as share of subject total scores for each subject during 2014-2018. This figure is referenced in Section 6.1.

B Supplementary Tables

Table B1: Examples of Admission Rules

(1)	(2)	(3)	(4)	(5)	(6)	(7)
		Admission Cutoffs				
Ex. Number	Priority Score	A	B	C	D	Admission Outcome
1	600	595	590	587	580	A
2	580	581	572	583	550	B
3	577	595	590	587	580	None
4	580	595	590	587	580	D

Note: This table presents four examples in which students with different priority scores applying for different sets of colleges in their ROLs. The serial number of examples is in Column (1). Each example is associated with a hypothetical student. The priority scores of the students are recorded in Column (2). Columns (3)-(6) present the admission cutoffs of the colleges that the students put in their ROLs. Column (7) presents the admission outcome of each hypothetical student as a result of their scores and ROLs. This table is helpful for the discussion in Section 2.2.

Table B2: Number of Students

Track	2014	2015	2016	2017	2018
Humanity	1387	1587	1586	1606	2064
Science	6427	6644	7134	7545	8354

Note: This table presents the number of students by track and year for the administrative data sample that we analyze.

Table B3: Validity of Estimates of Unconditional Probability

	(1)	(2)	(3)	(4)
	Admitted to First		Admitted to Second	
Estimated Uncond. Prob. to First	1.002	1.012		
	(0.004)	(0.008)		
Estimated Uncond. Prob. to Second			0.990	1.026
			(0.006)	(0.011)
Constant	-0.004	-0.005	0.003	-0.009
	(0.002)	(0.004)	(0.004)	(0.006)
Subsample (Science/Humanity)	Science	Humanity	Science	Humanity
Empirical Mean Probablity of Vacancy	.41	.314	.482	.42
Predicted Mean Probability of Vacancy	.413	.316	.484	.418

Note: This table reports the empirical exercise that tests the validity of probability estimates that we construct in Section 3.2. Columns 1 and 2 present the results of the regression where we regress the outcome of being admitted to the first choices on the estimated probability of meeting the cutoff of first choices, using the full sample of science-track and humanity-track students, respectively. Columns 3 and 4 present the results of the regression where we regress the outcome of being admitted to the second choices on the estimated probability of meeting the cutoff of second choices using the subsample of students who are not admitted to their first choices, for science-track and humanity track, respectively. This table is helpful for the discussion in 3.2.

Table B4: Design of Incentivized Questions in Survey

	Panel A1: ROL for Question Group 1			Panel C1: ROL for Question Group 3		
	ROL #	Payoffs	Prob(Meeting Cutoffs)	ROL #	Payoffs	Prob(Meeting Cutoffs)
	1st	?	?	1st	?	?
	2nd	20 CNY	100%	2nd	5 CNY	100%
	3rd	20 CNY	100%	3rd	5 CNY	100%
	4th	20 CNY	100%	4th	5 CNY	100%

	Panel A2: MPL for Question Group 1		Panel C2: MPL for Question Group 3	
Question #	College X	College Y	College X	College Y
(1)	25 CNY, 50%	30 CNY, 25%	25 CNY, 50%	30 CNY, 25%
(2)	25 CNY, 50%	35 CNY, 25%	25 CNY, 50%	35 CNY, 25%
(3)	25 CNY, 50%	40 CNY, 25%	25 CNY, 50%	40 CNY, 25%
(4)	25 CNY, 50%	45 CNY, 25%	25 CNY, 50%	45 CNY, 25%
(5)	25 CNY, 50%	50 CNY, 25%	25 CNY, 50%	50 CNY, 25%
(6)	25 CNY, 50%	55 CNY, 25%	25 CNY, 50%	55 CNY, 25%
(7)	25 CNY, 50%	60 CNY, 25%	25 CNY, 50%	60 CNY, 25%

	Panel B: MPL for Question Series 2	
Question #	Choice X (Payoff, Prob)	Choice Y (Payoff, Prob)
(1)	(25 CNY, 50%; 20 CNY, 50%)	(30 CNY, 25%; 20 CNY, 75%)
(2)	(25 CNY, 50%; 20 CNY, 50%)	(35 CNY, 25%; 20 CNY, 75%)
(3)	(25 CNY, 50%; 20 CNY, 50%)	(40 CNY, 25%; 20 CNY, 75%)
(4)	(25 CNY, 50%; 20 CNY, 50%)	(45 CNY, 25%; 20 CNY, 75%)
(5)	(25 CNY, 50%; 20 CNY, 50%)	(50 CNY, 25%; 20 CNY, 75%)
(6)	(25 CNY, 50%; 20 CNY, 50%)	(55 CNY, 25%; 20 CNY, 75%)
(7)	(25 CNY, 50%; 20 CNY, 50%)	(60 CNY, 25%; 20 CNY, 75%)

Note: This table presents the content of the three groups of incentivized MPL questions. Panels A1 and C1 present the ROL for Question Groups 1 and 3, respectively. Names of colleges for the second, third, and fourth spot are replaced with safe colleges that students are familiar with. Panels A2 and C2 present the MPL for Question Groups 1 and 3, respectively. In both question groups, students need to select either College X or College Y from each row, and the selected college will be put at the first spot in the corresponding ROL. Panel B presents the MPL for Question Group 2. Similarly, students need to choose one from Lottery X and Y in each row. This table is helpful for the discussion in Section 5.2.

Table B5: Welfare Gap between the Optimal and the DC Type

SES Quartile / CEE Quintile		0-25%	25-50%	50-75%	75-100%
0-20%	Status Quo	1.58%	1.13%	1.22%	2.11%
	Debiased	-1.22%	-1.41%	-1.26%	-0.07%
	Change	-2.80%	-2.54%	-2.48%	-2.18%
20-40%	Status Quo	4.33%	3.82%	3.23%	2.28%
	Debiased	-0.13%	-0.02%	-0.03%	0.21%
	Change	-4.46%	-3.84%	-3.26%	-2.07%
40-60%	Status Quo	4.38%	4.98%	4.49%	5.26%
	Debiased	-0.16%	-0.06%	-0.05%	-0.16%
	Change	-4.54%	-5.04%	-4.54%	-5.43%
60-80%	Status Quo	4.59%	4.77%	4.68%	4.88%
	Debiased	-0.02%	-0.08%	-0.06%	-0.12%
	Change	-4.61%	-4.84%	-4.74%	-5.00%
80-100%	Status Quo	4.41%	5.20%	4.90%	4.51%
	Debiased	-0.19%	-0.15%	-0.39%	-0.54%
	Change	-4.60%	-5.35%	-5.29%	-5.05%

Note: this table presents the difference in average welfare between the Optimal Type and the DC Type. Positive difference means that the Optimal Type is better off compared to the DC Type. The welfare is normalized by the assumption that the best first-tier college for each student is 100%, and the worse first-tier college is 0%. Based on students' exam score in the CEE, we divide them into five groups: 0-20%, 20-40%, 40-60%, 60-80%, 80-100% respectively. Higher quintile means the exam performance was better. The results of each CEE group are presented in rows. Based on the student applicants' township of residence, we divide them into four groups according to average adult education attainment in the corresponding township: 0-25%, 25-50%, 50-75%, 75-100%. Higher quartile means the socioeconomic status was better. The results of each SES group are presented in columns. The table therefore calculate the average welfare for 5*4 cells of student groups. For each subgroup in question, we calculate three numbers. "Status Quo" refers to the welfare difference if no intervention was conducted. "Debiased" refers the welfare difference if all DC students played optimally in the application system. "Change" refers to the change in welfare gap if we debiased all the DC students. This table is referenced in Section 6.

Table B6: Subjective Beliefs

	(1)	(2)	(3)	(4)	(5)	(6)
	First Choice		Fourth Choice		Reversal Share	
SES Index (Normalized)	-1.95%	-1.75%	-0.80%	-0.88%	-2.81%	-3.02%
	(0.82%)	(0.83%)	(0.83%)	(0.83%)	(1.24%)	(1.25%)
CEE Score	Yes	Yes	Yes	Yes	Yes	Yes
College Preferences	No	Yes	No	Yes	No	Yes

Note: This table presents analysis of student applicants' subjective beliefs about the unconditional admission probability for colleges of choices. SES index is parents' average years of education attainment. Columns 1 and 2 analyzes the beliefs about meeting the cutoff of the first choices. Columns 3 and 4 analyzes the beliefs about meeting the cutoff of the fourth choices. Column 5 and 6 analyzes the share of severe reversal, where a list is said to exhibit reversal if there the probability of any non-bottom choice exceeds that of the bottom choice by 25%. Besides SES Index, the only additional covariate in Columns 1, 3, and 5, is the fourth-order polynomial of CEE score. In Columns 2, 4, and 6, we additionally control for stated college preferences. This table is referenced in Section 7.1.

Table B7: Major Preference and its Impact on Risk Taking

	(1)	(2)	(3)	(4)	(5)	(6)
	Major Top Concern		Estimated Prob		Subjective Beliefs	
SES Index (Normalized)	-1.53%	-1.24%				
	(0.92%)	(0.96%)				
Major Top Concern			2.15%	1.53%	9.63%	8.04%
			(2.93%)	(3.08%)	(2.36%)	(2.43%)
Share of Major Top Concern	13.17%					
Implied Impact on 1st Choices			0.28%	0.29%	1.27%	1.06%
Priority Score	Yes	Yes	Yes	Yes	Yes	Yes
College Preferences	No	Yes	No	Yes	No	Yes

Note: Columns 1 and 2 present regressions that examine whether share of survey takers who think major is their top concern is correlated with SES Index. Columns 3 and 4 present the regression that examines whether those who declare major to be of top concern take exhibit different degree of cautiousness their first choices, in terms of the estimated unconditional admission probability. Columns 5 and 6 examine whether those who declare major to be of top concern take different amount of risks on their first choices, in terms of subjective beliefs. Columns 1, 3, 5 only control for priority score, whereas Columns 2, 4, 6 additionally control for demographics as well as stated college preferences. The share of people who think major is most important have been reported below the estimate for Columns 1 and 2. The implied impact of major consideration on the risk-taking of first choices is reported below the estimated coefficient in Columns 3, 4, 5, 6. This is calculated as the product of the share of people who regard major as the most important consideration, and the average of additional cautiousness for the first choices among this group of people. This table is referenced in Section 7.1.

Table B8: Parameter Estimation for Specification 1 - Optimal Type Only

O_i	-75.893 (4.235)	O^{Cmptv2}	-63.306 (1.955)	O^{Cmptv3}	8.325 (0.912)	$O^{SCmptv2}$	2.903 (1.054)	$O^{CCmptv2}$	-1.665 (0.986)	θ^{Cmptv}	53.165 (1.994)
θ^{Mega}	-12.949 (1.074)	θ^{d}	-2.768 (0.274)	θ^{C985}	-10.380 (2.160)	θ^{Gra}	-17.249 (2.722)	θ^{Sci}	8.934 (1.603)	θ^{Cpmrh}	1.121 (2.221)
θ^{Fin}	-16.464 (3.054)	θ^{Med}	-21.738 (2.473)	θ^{Tch}	9.559 (3.425)	θ^{SCmptv}	1.757 (1.503)	θ^{SMega}	-2.031 (0.906)	θ^{Sd}	4.815 (0.276)
θ^{SC985}	-1.119 (1.171)	θ^{SGra}	-0.224 (0.766)	θ^{SSci}	4.569 (2.175)	θ^{SCpmrh}	1.556 (2.172)	θ^{SFin}	0.773 (1.752)	θ^{SMed}	-4.365 (1.047)
θ^{STch}	3.948 (4.077)	θ^{CCmptv}	103.522 (3.222)	θ^{CMega}	-1.877 (1.348)	θ^{Cd}	1.021 (0.266)	θ^{CC985}	-3.223 (2.670)	θ^{CGra}	1.841 (4.356)
θ^{CSci}	11.213 (1.520)	θ^{CCpmrh}	12.737 (1.564)	θ^{CFin}	4.346 (3.105)	θ^{CMed}	6.850 (1.732)	θ^{CTch}	7.428 (2.186)	σ^{Cmptv}	3.172 (0.132)
σ^{Mega}	-0.765 (66.461)	σ^{d}	-0.429 (9.208)	σ^{C985}	0.310 (18.329)	σ^{Gra}	-1.066 (27.216)	σ^{Sci}	2.328 (0.373)	σ^{Cpmrh}	3.210 (0.105)
σ^{Fin}	3.170 (0.075)	σ^{Med}	3.362 (0.065)	σ^{Tch}	0.855 (1.627)	κ^{Cmptv}	-69.830 (2.609)	κ^{Mega}	34.691 (2.243)	κ^{d}	-17.933 (0.573)
κ^{C985}	15.222 (3.245)	κ^{Gra}	40.954 (3.988)	κ^{Sci}	-37.017 (3.044)	κ^{Cpmrh}	-16.755 (3.721)	κ^{Fin}	13.438 (7.567)	κ^{Med}	21.930 (3.950)
κ^{Tch}	-28.706 (14.308)	σ^{O}	-0.151 (1032.300)	θ^{OS}	-5.515 (2.207)	θ^{OC}	-22.437 (4.717)	θ^{OS2}	3.504 (1.855)	θ^{OC2}	38.187 (3.179)

Note: This table presents the point estimate and standard error of the second specification for the structural estimation exercise in Section 6. Standard errors are in the parenthesis. If the standard errors are omitted for a parameter, the moments' derivatives with regard to that parameter

Table B9: Parameter Estimation for Specification 2 -
Optimal and DC Type with Identical Preference Distribution

O_i	-69.222 (7.036)	O^{Cmptv2}	-18.933 (4.756)	O^{Cmptv3}	28.915 (3.180)	$O^{SCmptv2}$	-6.498 (2.492)	$O^{CCmptv2}$	36.759 (5.556)	θ^{Cmptv}	56.108 (1.911)
θ^{Mega}	-13.188 (2.838)	θ^{d}	-7.548 (0.660)	θ^{C985}	-11.331 (6.063)	θ^{Gra}	-11.084 (3.699)	θ^{Sci}	9.664 (1.777)	θ^{Cpmrh}	12.396 (2.077)
θ^{Fin}	-10.953 (9.010)	θ^{Med}	-5.455 (3.972)	θ^{Tch}	5.174 (5.894)	θ^{SCmptv}	-12.255 (2.564)	θ^{SMega}	-0.401 (3.413)	θ^{Sd}	5.436 (0.382)
θ^{SC985}	-0.593 (1.293)	θ^{SGra}	0.691 (0.960)	θ^{SSci}	2.671 (0.597)	θ^{SCpmrh}	-0.266 (0.518)	θ^{SFin}	-0.289 (.)	θ^{SMed}	-5.226 (0.739)
θ^{STch}	0.396 (0.713)	θ^{CCmptv}	85.551 (3.936)	θ^{CMega}	-9.217 (1.541)	θ^{Cd}	1.363 (0.461)	θ^{CC985}	-0.879 (3.499)	θ^{CGra}	-7.499 (2.352)
θ^{CSci}	1.241 (1.220)	θ^{CCpmrh}	-1.109 (1.597)	θ^{CFin}	-5.098 (3.650)	θ^{CMed}	7.291 (1.993)	θ^{CTch}	-3.009 (1.871)	σ^{Cmptv}	-0.873 (40.446)
σ^{Mega}	0.357 (12.580)	σ^{d}	1.151 (0.465)	σ^{C985}	-0.120 (49.035)	σ^{Gra}	3.044 (0.125)	σ^{Sci}	3.886 (0.062)	σ^{Cpmrh}	0.821 (3.472)
σ^{Fin}	3.272 (0.083)	σ^{Med}	3.496 (0.083)	σ^{Tch}	2.559 (0.321)	κ^{Cmptv}	-32.528 (4.078)	κ^{Mega}	27.333 (1.529)	κ^{d}	-15.705 (0.723)
κ^{C985}	-47.914 (6.536)	κ^{Gra}	9.918 (11.973)	κ^{Sci}	-25.151 (4.042)	κ^{Cpmrh}	-51.025 (6.502)	κ^{Fin}	-67.106 (15.787)	κ^{Med}	-64.784 (14.415)
κ^{Tch}	-58.029 (4.364)	ΔO^{DC}	49.939 (.)	σ^{O}	-0.754 (.)	θ^{OS}	-21.496 (7.391)	θ^{OC}	10.639 (3.955)	θ^{OS2}	-6.391 (6.423)
θ^{OC2}	33.103 (5.581)										

Note: This table presents the point estimate and standard error of the second specification for the structural estimation exercise in Section 6. Standard errors are in the parenthesis. If the standard errors are omitted for a parameter, the moments' derivatives with regard to that parameter

Table B10: Parameter Estimation for Specification 3 - Optimal and DC Type with Different Preference Distribution

O_i	-59.871 (11.279)	O^{Cmptv2}	-18.076 (2.504)	O^{Cmptv3}	27.884 (3.634)	$O^{SCmptv2}$	-5.989 (2.718)	$O^{CCmptv2}$	30.292 (3.315)
θ^{Cmptv}	40.175 (3.011)	θ^{Mega}	-12.174 (7.570)	θ^{d}	-5.882 (1.071)	θ^{C985}	-8.704 (12.701)	θ^{Gra}	-8.570 (10.140)
θ^{Sci}	10.272 (1.984)	θ^{Cpmrh}	10.609 (4.172)	θ^{Fin}	-6.382 (9.480)	θ^{Med}	-5.420 (9.605)	θ^{Tch}	4.767 (4.282)
θ^{SCmptv}	-11.747 (2.222)	θ^{SMega}	-0.513 (10.249)	θ^{Sd}	4.109 (0.359)	θ^{SC985}	-0.461 (8.063)	θ^{SGra}	0.576 (4.421)
θ^{SSci}	2.287 (2.408)	θ^{SCpmrh}	-0.239 (3.877)	θ^{SFin}	-0.267 (15.442)	θ^{SMed}	-4.107 (8.723)	θ^{STch}	0.498 (4.910)
θ^{CCmptv}	75.977 (4.169)	θ^{CMega}	-8.214 (10.576)	θ^{Cd}	1.439 (1.490)	θ^{CC985}	-0.747 (22.202)	θ^{CGra}	-6.253 (10.037)
θ^{CSci}	1.403 (3.144)	θ^{CCpmrh}	-1.103 (4.845)	θ^{CFin}	-4.831 (16.187)	θ^{CMed}	6.863 (6.169)	θ^{CTch}	-1.834 (2.463)
σ^{Cmptv}	-0.781 (17.598)	σ^{Mega}	0.384 (14.233)	σ^{d}	0.421 (2.472)	σ^{C985}	-0.132 (.)	σ^{Gra}	2.627 (0.476)
σ^{Sci}	2.870 (0.182)	σ^{Cpmrh}	0.696 (6.627)	σ^{Fin}	1.785 (1.398)	σ^{Med}	3.049 (0.346)	σ^{Tch}	2.071 (0.234)
κ^{Cmptv}	-28.154 (4.415)	κ^{Mega}	22.988 (3.851)	κ^{d}	-11.999 (1.152)	κ^{C985}	-43.986 (9.320)	κ^{Gra}	12.471 (8.223)
κ^{Sci}	-22.761 (3.738)	κ^{Cpmrh}	-39.412 (4.991)	κ^{Fin}	-63.087 (5.866)	κ^{Med}	-59.709 (13.185)	κ^{Tch}	-46.690 (11.608)
ΔO^{DC}	39.859 (.)	σ^{O}	-0.700 (.)	θ^{OS}	-5.159 (5.304)	θ^{OC}	2.553 (2.424)	θ^{OS2}	-1.534 (3.207)
θ^{OC2}	26.771 (6.163)	$\theta^{Cmptv2}(DC)$	-0.883 (2.658)	$\theta^{Cmptv3}(DC)$	3.433 (3.647)	$\theta^{SCmptv2}(DC)$	1.472 (3.012)	$\theta^{CCmptv2}(DC)$	3.533 (2.861)
$\theta^{Cmptv}(DC)$	-13.711 (2.315)	$\theta^{Mega}(DC)$	-1.402 (11.339)	$\theta^{d}(DC)$	-0.826 (1.988)	$\theta^{C985}(DC)$	-3.308 (24.823)	$\theta^{Gra}(DC)$	3.535 (13.936)
$\theta^{Sci}(DC)$	1.819 (3.899)	$\theta^{Cpmrh}(DC)$	3.063 (6.324)	$\theta^{Fin}(DC)$	-1.464 (18.645)	$\theta^{Med}(DC)$	-2.200 (15.374)	$\theta^{Tch}(DC)$	0.206 (.)
$\theta^{SCmptv}(DC)$	0.957 (2.397)	$\theta^{SMega}(DC)$	1.761 (21.064)	$\theta^{Sd}(DC)$	0.023 (.)	$\theta^{SC985}(DC)$	0.099 (.)	$\theta^{SGra}(DC)$	-0.220 (.)
$\theta^{SSci}(DC)$	1.110 (4.056)	$\theta^{SCpmrh}(DC)$	1.063 (6.086)	$\theta^{SFin}(DC)$	3.108 (32.931)	$\theta^{SMed}(DC)$	0.649 (15.172)	$\theta^{STch}(DC)$	1.521 (7.772)
$\theta^{CCmptv}(DC)$	0.422 (2.355)	$\theta^{CMega}(DC)$	2.876 (21.346)	$\theta^{Cd}(DC)$	-2.321 (1.717)	$\theta^{CC985}(DC)$	-2.742 (37.440)	$\theta^{CGra}(DC)$	-13.359 (18.394)
$\theta^{CSci}(DC)$	0.346 (3.653)	$\theta^{CCpmrh}(DC)$	6.476 (8.290)	$\theta^{CFin}(DC)$	-10.228 (46.275)	$\theta^{CMed}(DC)$	0.169 (.)	$\theta^{CTch}(DC)$	0.228 (.)
θ^{OS}	5.888 (.)	θ^{OC}	0.155 (.)	θ^{OS2}	0.670 (.)	θ^{OC2}	1.149 (.)		

Note: This table presents the point estimate and standard error of the third specification for the structural estimation exercise in Section 6. Standard errors are in the parenthesis. If the standard errors are omitted for a parameter, the moments' derivatives with regard to that parameter.

Table B11: Moments and Fitted Value for Specification 1 – Optimal Type Only

Cmptv	-0.147 (0.005) [-0.028]	Mega	0.148 (0.000) [0.149]	Dist	-0.477 (0.005) [-0.463]	C985	0.146 (0.001) [0.141]	Gra	0.228 (0.001) [0.218]	Sci	0.427 (0.001) [0.396]	Cmprh	0.287 (0.000) [0.250]	Fin	0.062 (0.000) [0.058]
Med	0.098 (0.000) [0.097]	Tch	0.038 (0.000) [0.039]	SES*Cmptv	0.202 (0.006) [0.210]	SES*Mega	0.042 (0.001) [0.039]	SES*Dist	0.287 (0.006) [0.281]	SES*C985	0.060 (0.001) [0.056]	SES*Gra	0.070 (0.001) [0.074]	SES*Sci	0.017 (0.002) [0.029]
SES*Cmprh	0.009 (0.001) [0.022]	SES*Fin	0.006 (0.000) [0.008]	SES*Med	-0.025 (0.000) [-0.024]	SES*Tch	-0.006 (0.000) [-0.004]	CEE*Cmptv	0.772 (0.011) [0.854]	CEE*Mega	0.110 (0.001) [0.093]	CEE*Dist	0.199 (0.005) [0.183]	CEE*C985	0.203 (0.002) [0.208]
CEE*Gra	0.249 (0.002) [0.249]	CEE*Sci	-0.013 (0.002) [-0.002]	CEE*Cmprh	0.048 (0.001) [0.056]	CEE*Fin	0.023 (0.000) [0.025]	CEE*Med	-0.012 (0.000) [-0.009]	CEE*Tch	-0.018 (0.000) [-0.018]	$Cmptv^2$	0.732 (0.008) [0.788]	$Mega^2$	0.072 (0.000) [0.079]
$Dist^2$	0.915 (0.017) [1.207]	$C985^2$	0.097 (0.000) [0.095]	Gra^2	0.159 (0.001) [0.168]	Sci^2	0.291 (0.001) [0.313]	$Cmprh^2$	0.143 (0.000) [0.173]	Fin^2	0.029 (0.000) [0.031]	Med^2	0.052 (0.000) [0.054]	Tch^2	0.017 (0.000) [0.016]
C1S1First	0.007 (0.000) [0.012]	C1S1Second	0.011 (0.000) [0.010]	C1S1Third	0.013 (0.000) [0.015]	C1S1Fourth	0.014 (0.000) [0.023]	C1S1R0	0.048 (0.000) [0.029]	C1S1R1	0.032 (0.000) [0.016]	C1S2First	0.009 (0.000) [0.012]	C1S2Second	0.014 (0.000) [0.011]
C1S2Third	0.017 (0.000) [0.018]	C1S2Fourth	0.017 (0.000) [0.026]	C1S2R0	0.047 (0.000) [0.030]	C1S2R1	0.031 (0.000) [0.015]	C1S3First	0.009 (0.000) [0.011]	C1S3Second	0.014 (0.000) [0.011]	C1S3Third	0.017 (0.000) [0.018]	C1S3Fourth	0.018 (0.000) [0.025]
C1S3R0	0.043 (0.000) [0.027]	C1S3R1	0.027 (0.000) [0.012]	C1S4First	0.004 (0.000) [0.006]	C1S4Second	0.008 (0.000) [0.007]	C1S4Third	0.010 (0.000) [0.011]	C1S4Fourth	0.011 (0.000) [0.016]	C1S4R0	0.023 (0.000) [0.017]	C1S4R1	0.013 (0.000) [0.006]
C2S1First	0.020 (0.000) [0.013]	C2S1Second	0.028 (0.000) [0.020]	C2S1Third	0.033 (0.000) [0.025]	C2S1Fourth	0.036 (0.000) [0.036]	C2S1R0	0.045 (0.000) [0.034]	C2S1R1	0.030 (0.000) [0.011]	C2S2First	0.021 (0.000) [0.013]	C2S2Second	0.030 (0.000) [0.020]
C2S2Third	0.034 (0.000) [0.026]	C2S2Fourth	0.037 (0.000) [0.037]	C2S2R0	0.041 (0.000) [0.030]	C2S2R1	0.026 (0.000) [0.009]	C2S3First	0.023 (0.000) [0.013]	C2S3Second	0.032 (0.000) [0.020]	C2S3Third	0.035 (0.000) [0.027]	C2S3Fourth	0.038 (0.000) [0.038]
C2S3R0	0.039 (0.000) [0.029]	C2S3R1	0.022 (0.000) [0.008]	C2S4First	0.013 (0.000) [0.009]	C2S4Second	0.021 (0.000) [0.014]	C2S4Third	0.024 (0.000) [0.019]	C2S4Fourth	0.026 (0.000) [0.027]	C2S4R0	0.026 (0.000) [0.020]	C2S4R1	0.014 (0.000) [0.005]
C3S1First	0.026 (0.000) [0.022]	C3S1Second	0.034 (0.000) [0.027]	C3S1Third	0.039 (0.000) [0.034]	C3S1Fourth	0.043 (0.000) [0.044]	C3S1R0	0.037 (0.000) [0.027]	C3S1R1	0.022 (0.000) [0.013]	C3S2First	0.027 (0.000) [0.022]	C3S2Second	0.038 (0.000) [0.027]
C3S2Third	0.043 (0.000) [0.034]	C3S2Fourth	0.045 (0.000) [0.045]	C3S2R0	0.033 (0.000) [0.025]	C3S2R1	0.017 (0.000) [0.012]	C3S3First	0.028 (0.000) [0.020]	C3S3Second	0.038 (0.000) [0.025]	C3S3Third	0.041 (0.000) [0.033]	C3S3Fourth	0.044 (0.000) [0.043]
C3S3R0	0.031 (0.000) [0.022]	C3S3R1	0.016 (0.000) [0.010]	C3S4First	0.019 (0.000) [0.016]	C3S4Second	0.027 (0.000) [0.020]	C3S4Third	0.031 (0.000) [0.026]	C3S4Fourth	0.035 (0.000) [0.035]	C3S4R0	0.024 (0.000) [0.017]	C3S4R1	0.013 (0.000) [0.008]
C4S1First	0.025 (0.000) [0.024]	C4S1Second	0.032 (0.000) [0.029]	C4S1Third	0.037 (0.000) [0.036]	C4S1Fourth	0.040 (0.000) [0.041]	C4S1R0	0.030 (0.000) [0.030]	C4S1R1	0.017 (0.000) [0.013]	C4S2First	0.025 (0.000) [0.024]	C4S2Second	0.035 (0.000) [0.029]
C4S2Third	0.040 (0.000) [0.036]	C4S2Fourth	0.043 (0.000) [0.042]	C4S2R0	0.027 (0.000) [0.030]	C4S2R1	0.012 (0.000) [0.012]	C4S3First	0.026 (0.000) [0.024]	C4S3Second	0.035 (0.000) [0.029]	C4S3Third	0.041 (0.000) [0.036]	C4S3Fourth	0.044 (0.000) [0.043]
C4S3R0	0.025 (0.000) [0.031]	C4S3R1	0.011 (0.000) [0.012]	C4S4First	0.025 (0.000) [0.027]	C4S4Second	0.038 (0.000) [0.032]	C4S4Third	0.044 (0.000) [0.040]	C4S4Fourth	0.049 (0.000) [0.049]	C4S4R0	0.029 (0.000) [0.037]	C4S4R1	0.011 (0.000) [0.014]

Note: This table presents the moments used - mean of variables used among the four choices on ROL, the standard error estimate of the mean (in parenthesis), and the fitted value by the estimated model [in squared brackets]. Cmptv refers to the standardized competitiveness of the colleges. Mega refers to whether the college is in a mega city. Dist refers to the standardized distance of the university from Ningxia. C985 refers to whether the university belongs to "985-school". Gra refers to whether the university has a graduate school, Sci refers to whether science-related majors are predominant in the university, Cmprh refers to whether all majors are adquautely represented in the university, Fin refers to whether finance & economics majors are predominant in the university. Med refers to whether the university is concentrated on medical-related majors. Tch refers to whether the university is concentrated on teaching-related majors. A variable with superscript 2 means that it is the mean of the square of that variable. CiSjFirst, CiSjSecond, CiSjThird, CiSjFourth refers to the average of unconditional admission probability of the first, second, third, fourth choice times a dummy that indicates whether the observation belongs to the ith quintile in CEE score, jth quartile of SES index. CiSjR0 or CiSjR25 refers to the average of whether an ROL exhibit a comptitiveness reversal of size larger than 0% or 25% times a dummy that indicates whether the observation belongs to the ith quintile in CEE score, jth quartile of SES index.

Table B12: Moments and Fitted Value for Specification 2 - Optimal and DC Type with Identical Preference Distribution

Cmptv	-0.147	Mega	0.148	Dist	-0.477	C985	0.146	Gra	0.228	Sci	0.427	Cmprh	0.287	Fin	0.062
	(0.005)		(0.000)		(0.005)		(0.001)		(0.001)		(0.001)		(0.000)		(0.000)
	[-0.037]		[0.149]		[-0.466]		[0.145]		[0.228]		[0.396]		[0.259]		[0.065]
Med	0.098	Tch	0.038	SES*Cmptv	0.202	SES*Mega	0.042	SES*Dist	0.287	SES*C985	0.060	SES*Gra	0.070	SES*Sci	0.017
	(0.000)		(0.000)		(0.006)		(0.001)		(0.006)		(0.001)		(0.001)		(0.002)
	[0.102]		[0.041]		[0.224]		[0.037]		[0.270]		[0.059]		[0.073]		[0.010]
SES*Cmprh	0.009	SES*Fin	0.006	SES*Med	-0.025	SES*Tch	-0.006	CEE*Cmptv	0.772	CEE*Mega	0.110	CEE*Dist	0.199	CEE*C985	0.203
	(0.001)		(0.000)		(0.000)		(0.000)		(0.011)		(0.001)		(0.005)		(0.002)
	[0.012]		[0.007]		[-0.022]		[-0.006]		[0.905]		[0.092]		[0.200]		[0.206]
CEE*Gra	0.249	CEE*Sci	-0.013	CEE*Cmprh	0.048	CEE*Fin	0.023	CEE*Med	-0.012	CEE*Tch	-0.018	$Cmptv^2$	0.732	$Mega^2$	0.072
	(0.002)		(0.002)		(0.001)		(0.000)		(0.000)		(0.000)		(0.008)		(0.000)
	[0.239]		[-0.019]		[0.012]		[0.026]		[-0.013]		[-0.019]		[0.834]		[0.077]
$Dist^2$	0.915	$C985^2$	0.097	Gra^2	0.159	Sci^2	0.291	$Cmprh^2$	0.143	Fin^2	0.029	Med^2	0.052	Tch^2	0.017
	(0.017)		(0.000)		(0.001)		(0.001)		(0.000)		(0.000)		(0.000)		(0.000)
	[1.050]		[0.105]		[0.163]		[0.329]		[0.166]		[0.031]		[0.057]		[0.017]
C1S1First	0.007	C1S1Second	0.011	C1S1Third	0.013	C1S1Fourth	0.014	C1S1R0	0.048	C1S1R1	0.032	C1S2First	0.009	C1S2Second	0.014
	(0.000)		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)
	[0.013]		[0.010]		[0.011]		[0.013]		[0.039]		[0.015]		[0.013]		[0.012]
C1S2Third	0.017	C1S2Fourth	0.017	C1S2R0	0.047	C1S2R1	0.031	C1S3First	0.009	C1S3Second	0.014	C1S3Third	0.017	C1S3Fourth	0.018
	(0.000)		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)
	[0.013]		[0.017]		[0.037]		[0.014]		[0.012]		[0.011]		[0.014]		[0.020]
C1S3R0	0.043	C1S3R1	0.027	C1S4First	0.004	C1S4Second	0.008	C1S4Third	0.010	C1S4Fourth	0.011	C1S4R0	0.023	C1S4R1	0.013
	(0.000)		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)
	[0.033]		[0.012]		[0.007]		[0.009]		[0.012]		[0.017]		[0.018]		[0.006]
C2S1First	0.020	C2S1Second	0.028	C2S1Third	0.033	C2S1Fourth	0.036	C2S1R0	0.045	C2S1R1	0.030	C2S2First	0.021	C2S2Second	0.030
	(0.000)		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)
	[0.027]		[0.028]		[0.030]		[0.033]		[0.045]		[0.024]		[0.024]		[0.026]
C2S2Third	0.034	C2S2Fourth	0.037	C2S2R0	0.041	C2S2R1	0.026	C2S3First	0.023	C2S3Second	0.032	C2S3Third	0.035	C2S3Fourth	0.038
	(0.000)		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)
	[0.028]		[0.033]		[0.039]		[0.020]		[0.023]		[0.025]		[0.029]		[0.034]
C2S3R0	0.039	C2S3R1	0.022	C2S4First	0.013	C2S4Second	0.021	C2S4Third	0.024	C2S4Fourth	0.026	C2S4R0	0.026	C2S4R1	0.014
	(0.000)		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)
	[0.035]		[0.018]		[0.016]		[0.017]		[0.021]		[0.026]		[0.021]		[0.011]
C3S1First	0.026	C3S1Second	0.034	C3S1Third	0.039	C3S1Fourth	0.043	C3S1R0	0.037	C3S1R1	0.022	C3S2First	0.027	C3S2Second	0.038
	(0.000)		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)
	[0.031]		[0.032]		[0.035]		[0.039]		[0.040]		[0.023]		[0.029]		[0.031]
C3S2Third	0.043	C3S2Fourth	0.045	C3S2R0	0.033	C3S2R1	0.017	C3S3First	0.028	C3S3Second	0.038	C3S3Third	0.041	C3S3Fourth	0.044
	(0.000)		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)
	[0.036]		[0.041]		[0.037]		[0.019]		[0.026]		[0.030]		[0.035]		[0.040]
C3S3R0	0.031	C3S3R1	0.016	C3S4First	0.019	C3S4Second	0.027	C3S4Third	0.031	C3S4Fourth	0.035	C3S4R0	0.024	C3S4R1	0.013
	(0.000)		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)
	[0.033]		[0.017]		[0.020]		[0.024]		[0.029]		[0.034]		[0.024]		[0.011]
C4S1First	0.025	C4S1Second	0.032	C4S1Third	0.037	C4S1Fourth	0.040	C4S1R0	0.030	C4S1R1	0.017	C4S2First	0.025	C4S2Second	0.035
	(0.000)		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)
	[0.029]		[0.031]		[0.035]		[0.039]		[0.035]		[0.016]		[0.027]		[0.030]
C4S2Third	0.040	C4S2Fourth	0.043	C4S2R0	0.027	C4S2R1	0.012	C4S3First	0.026	C4S3Second	0.035	C4S3Third	0.041	C4S3Fourth	0.044
	(0.000)		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)
	[0.035]		[0.040]		[0.033]		[0.014]		[0.026]		[0.030]		[0.036]		[0.042]
C4S3R0	0.025	C4S3R1	0.011	C4S4First	0.025	C4S4Second	0.038	C4S4Third	0.044	C4S4Fourth	0.049	C4S4R0	0.029	C4S4R1	0.011
	(0.000)		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)
	[0.032]		[0.013]		[0.028]		[0.033]		[0.041]		[0.048]		[0.032]		[0.013]

Note: This table presents the moments used - mean of variables used among the four choices on ROL, the standard error estimate of the mean (in parenthesis), and the fitted value by the estimated model [in squared brackets]. Cmptv refers to the standardized competitiveness of the colleges. Mega refers to whether the college is in a mega city. Dist refers to the standardized distance of the university from Ningxia. C985 refers to whether the university belongs to "985-school". Gra refers to whether the university has a graduate school, Sci refers to whether science-related majors are predominant in the university, Cmprh refers to whether all majors are adquautely represented in the university, Fin refers to whether finance & economics majors are predominant in the university. Med refers to whether the university is concentrated on medical-related majors. Tch refers to whether the university is concentrated on teaching-related majors. A variable with superscript 2 means that it is the mean of the square of that variable. CiSjFirst, CiSjSecond, CiSjThird, CiSjFourth refers to the average of unconditional admission probability of the first, second, third, fourth choice times a dummy that indicates whether the observation belongs to the ith quintile in CEE score, jth quartile of SES index. CiSjR0 or CiSjR25 refers to the average of whether an ROL exhibit a comptitiveness reversal of size larger than 0% or 25% times a dummy that indicates whether the observation belongs to the ith quintile in CEE score, jth quartile of SES index.

Table B13: Moments and Fitted Value for Specification 3 - Optimal and DC Type with Different Preference Distribution

Cmptv	-0.147 (0.005) [-0.049]	Mega	0.148 (0.000) [0.144]	Dist	-0.477 (0.005) [-0.480]	C985	0.146 (0.001) [0.142]	Gra	0.228 (0.001) [0.232]	Sci	0.427 (0.001) [0.403]	Cmprh	0.287 (0.000) [0.262]	Fin	0.062 (0.000) [0.065]
Med	0.098 (0.000) [0.097]	Tch	0.038 (0.000) [0.039]	SES*Cmptv	0.202 (0.006) [0.224]	SES*Mega	0.042 (0.001) [0.039]	SES*Dist	0.287 (0.006) [0.264]	SES*C985	0.060 (0.001) [0.060]	SES*Gra	0.070 (0.001) [0.074]	SES*Sci	0.017 (0.002) [0.015]
SES*Cmprh	0.009 (0.001) [0.014]	SES*Fin	0.006 (0.000) [0.007]	SES*Med	-0.025 (0.000) [-0.024]	SES*Tch	-0.006 (0.000) [-0.005]	CEE*Cmptv	0.772 (0.011) [0.912]	CEE*Mega	0.110 (0.001) [0.095]	CEE*Dist	0.199 (0.005) [0.182]	CEE*C985	0.203 (0.002) [0.203]
CEE*Gra	0.249 (0.002) [0.234]	CEE*Sci	-0.013 (0.002) [-0.027]	CEE*Cmprh	0.048 (0.001) [0.028]	CEE*Fin	0.023 (0.000) [0.022]	CEE*Med	-0.012 (0.000) [-0.013]	CEE*Tch	-0.018 (0.000) [-0.022]	$Cmptv^2$	0.732 (0.008) [0.844]	$Mega^2$	0.072 (0.000) [0.080]
$Dist^2$	0.915 (0.017) [1.059]	$C985^2$	0.097 (0.000) [0.100]	Gra^2	0.159 (0.001) [0.161]	Sci^2	0.291 (0.001) [0.307]	$Cmprh^2$	0.143 (0.000) [0.156]	Fin^2	0.029 (0.000) [0.025]	Med^2	0.052 (0.000) [0.052]	Tch^2	0.017 (0.000) [0.014]
C1S1First	0.007 (0.000) [0.012]	C1S1Second	0.011 (0.000) [0.009]	C1S1Third	0.013 (0.000) [0.010]	C1S1Fourth	0.014 (0.000) [0.013]	C1S1R0	0.048 (0.000) [0.040]	C1S1R1	0.032 (0.000) [0.014]	C1S2First	0.009 (0.000) [0.012]	C1S2Second	0.014 (0.000) [0.011]
C1S2Third	0.017 (0.000) [0.012]	C1S2Fourth	0.017 (0.000) [0.017]	C1S2R0	0.047 (0.000) [0.038]	C1S2R1	0.031 (0.000) [0.013]	C1S3First	0.009 (0.000) [0.011]	C1S3Second	0.014 (0.000) [0.011]	C1S3Third	0.017 (0.000) [0.013]	C1S3Fourth	0.018 (0.000) [0.018]
C1S3R0	0.043 (0.000) [0.034]	C1S3R1	0.027 (0.000) [0.012]	C1S4First	0.004 (0.000) [0.007]	C1S4Second	0.008 (0.000) [0.008]	C1S4Third	0.010 (0.000) [0.010]	C1S4Fourth	0.011 (0.000) [0.015]	C1S4R0	0.023 (0.000) [0.019]	C1S4R1	0.013 (0.000) [0.007]
C2S1First	0.020 (0.000) [0.028]	C2S1Second	0.028 (0.000) [0.029]	C2S1Third	0.033 (0.000) [0.031]	C2S1Fourth	0.036 (0.000) [0.034]	C2S1R0	0.045 (0.000) [0.045]	C2S1R1	0.030 (0.000) [0.024]	C2S2First	0.021 (0.000) [0.025]	C2S2Second	0.030 (0.000) [0.027]
C2S2Third	0.034 (0.000) [0.030]	C2S2Fourth	0.037 (0.000) [0.034]	C2S2R0	0.041 (0.000) [0.039]	C2S2R1	0.026 (0.000) [0.020]	C2S3First	0.023 (0.000) [0.025]	C2S3Second	0.032 (0.000) [0.027]	C2S3Third	0.035 (0.000) [0.031]	C2S3Fourth	0.038 (0.000) [0.036]
C2S3R0	0.039 (0.000) [0.035]	C2S3R1	0.022 (0.000) [0.017]	C2S4First	0.013 (0.000) [0.016]	C2S4Second	0.021 (0.000) [0.018]	C2S4Third	0.024 (0.000) [0.022]	C2S4Fourth	0.026 (0.000) [0.027]	C2S4R0	0.026 (0.000) [0.022]	C2S4R1	0.014 (0.000) [0.010]
C3S1First	0.026 (0.000) [0.032]	C3S1Second	0.034 (0.000) [0.034]	C3S1Third	0.039 (0.000) [0.037]	C3S1Fourth	0.043 (0.000) [0.042]	C3S1R0	0.037 (0.000) [0.040]	C3S1R1	0.022 (0.000) [0.023]	C3S2First	0.027 (0.000) [0.030]	C3S2Second	0.038 (0.000) [0.034]
C3S2Third	0.043 (0.000) [0.038]	C3S2Fourth	0.045 (0.000) [0.044]	C3S2R0	0.033 (0.000) [0.037]	C3S2R1	0.017 (0.000) [0.019]	C3S3First	0.028 (0.000) [0.028]	C3S3Second	0.038 (0.000) [0.032]	C3S3Third	0.041 (0.000) [0.037]	C3S3Fourth	0.044 (0.000) [0.043]
C3S3R0	0.031 (0.000) [0.033]	C3S3R1	0.016 (0.000) [0.016]	C3S4First	0.019 (0.000) [0.022]	C3S4Second	0.027 (0.000) [0.026]	C3S4Third	0.031 (0.000) [0.031]	C3S4Fourth	0.035 (0.000) [0.036]	C3S4R0	0.024 (0.000) [0.025]	C3S4R1	0.013 (0.000) [0.011]
C4S1First	0.025 (0.000) [0.029]	C4S1Second	0.032 (0.000) [0.032]	C4S1Third	0.037 (0.000) [0.035]	C4S1Fourth	0.040 (0.000) [0.040]	C4S1R0	0.030 (0.000) [0.035]	C4S1R1	0.017 (0.000) [0.015]	C4S2First	0.025 (0.000) [0.028]	C4S2Second	0.035 (0.000) [0.031]
C4S2Third	0.040 (0.000) [0.036]	C4S2Fourth	0.043 (0.000) [0.041]	C4S2R0	0.027 (0.000) [0.033]	C4S2R1	0.012 (0.000) [0.014]	C4S3First	0.026 (0.000) [0.027]	C4S3Second	0.035 (0.000) [0.031]	C4S3Third	0.041 (0.000) [0.036]	C4S3Fourth	0.044 (0.000) [0.043]
C4S3R0	0.025 (0.000) [0.032]	C4S3R1	0.011 (0.000) [0.012]	C4S4First	0.025 (0.000) [0.029]	C4S4Second	0.038 (0.000) [0.034]	C4S4Third	0.044 (0.000) [0.041]	C4S4Fourth	0.049 (0.000) [0.049]	C4S4R0	0.029 (0.000) [0.032]	C4S4R1	0.011 (0.000) [0.011]

Note: This table presents the moments used - mean of variables used among the four choices on ROL, the standard error estimate of the mean (in parenthesis), and the fitted value by the estimated model [in squared brackets]. Cmptv refers to the standardized competitiveness of the colleges. Mega refers to whether the college is in a mega city. Dist refers to the standardized distance of the university from Ningxia. C985 refers to whether the university belongs to "985-school". Gra refers to whether the university has a graduate school, Sci refers to whether science-related majors are predominant in the university, Cmprh refers to whether all majors are adquautely represented in the university, Fin refers to whether finance & economics majors are predominant in the university. Med refers to whether the university is concentrated on medical-related majors. Tch refers to whether the university is concentrated on teaching-related majors. A variable with superscript 2 means that it is the mean of the square of that variable. CiSjFirst, CiSjSecond, CiSjThird, CiSjFourth refers to the average of unconditional admission probability of the first, second, third, fourth choice times a dummy that indicates whether the observation belongs to the ith quintile in CEE score, jth quartile of SES index. CiSjR0 or CiSjR25 refers to the average of whether an ROL exhibit a comptitiveness reversal of size larger than 0% or 25% times a dummy that indicates whether the observation belongs to the ith quintile in CEE score, jth quartile of SES index.

C Details about Structural Estimation

C.1 Specification

This section lays out the details on the specification and the moments we use for the structural estimation in Section 6.

Let $X_j = (Cmptv_j, Mega_j, d_j, C985_j, Gra_j, Sci_j, Cmprh_j, Fin_j, Med_j, Tch_j)$, where

- $Cmptv_j$ is the standardized competitiveness of university j.
- $Mega_j$ an indicator variable that equals 1 if university j is located in a mega city, and 0 otherwise.
- d_j is the standardized distance of university j from Ningxia.
- $C985_j$ is an indicator variable that equals 1 if the university is classified as "985-school" (more reputable than other first-tier schools), and 0 otherwise.
- Gra_j is an indicator variable that equals 1 if the university has a graduate school;
- Sci_j is an indicator variable that equals 1 if the university is concentrated on natural science-related majors;
- $Cmprh_j$ is an indicator variable that equals 1 if the university does not specifically concentrate in any categories of majors;
- Fin_j is an indicator variable that equals 1 if the university is concentrated on economics and finance majors;
- Med_j is an indicator variable that equals 1 if the university is concentrated on medical-related majors;
- Tch_j is an indicator variable that equals 1 if the university is concentrated on teaching-related majors.

Recall in Section 6.2 that the specification is:

$$u_{ij} = \theta^X X_j + \theta^{SX} SES_i X_j + \theta^{CX} CEE_i X_j + (\eta_i^X + \kappa\nu_i)X_j + \\ f_C(C_j, SES_i, CEE_i) + O_i(SES_i, CEE_i) + \epsilon_{ij}$$

Similarly, define the set of variables $\mathcal{V} = \{Cmptv, Mega, d, C985, Gra, Sci, Cmprh, Fin, Med, Tch\}$. For convenience, use $v \in \mathcal{V}$ for each variable, and also the superscript of parameters to denote the parameter for a particular variable.

The expression above can be rewritten as

$$
\begin{aligned}
u_{ij} = \sum_{v\in\mathcal{V}}\theta^{v}v_j + \sum_{v\in\mathcal{V}}\theta^{Sv}SES_i v_j + \sum_{v\in\mathcal{V}}\theta^{Cv}CEE_i v_j + \sum_{v\in\mathcal{V}}(\sigma^{v}\tilde{\eta}_i^{v} \\
+\kappa^{v}\tilde{\nu}_i) + \theta^{Cmptv2}Cmptv_j^2 + \theta^{Cmptv3}Cmptv_j^3 + \\
\theta^{SCmptv2}SES_i * Cmptv_j^2 + \theta^{CCmptv2}CEE_i * Cmptv_j^2 + O_i + \sigma^{O}\epsilon_i + \theta^{OS} * SES_i + \\
\theta^{OC} * CEE_i + \theta^{OS2} * SES_i^2 + \theta^{OC2} * CEE_i^2 + \epsilon_{ij}
\end{aligned}
$$

where $\tilde{\eta}_i^{v}$, $\tilde{\nu}_i$ are jointly standard normally distributed. Note that there is no superscript in $\tilde{\nu}_i$, reflecting that the parameters κ^{v} mean to measure the magnitude of same-directional or opposite-directional heterogeneous preferences for all the variables in $\mathcal{V}$, whereas $\tilde{\eta}_i^{v}$ represents the heterogeneous preferences that are independent across variables. ϵ_{ij} follows a standard logistic distribution. This is the full version of Specification 1.

For Specification 2, parameter O_i is allowed to be different for the DC Type and the Rational Type. Other parameters are restricted to be the same. Therefore, Specification 2 has one more variable than Specification 1.

For Specification 3, parameters θ^{v}, θ^{Sv}, θ^{Cv}, θ^{Cmptv2}, θ^{Cmptv3}, $\theta^{SCmptv2}$, $\theta^{CCmptv2}$, O_i, θ^{OS}, θ^{OC}, θ^{OS2}, θ^{OC2} are allowed to be different for the DC Type and the Optimal Type. Other parameters are restricted to be the same. Since there are 38 more variables other than O_i, Specification 3 has 38 more variables than Specification 2.

C.2 Counterfactual

To simulate counterfactual, we first estimate applicants' preferences using Specification 3. Given the estimated parameters, we simulate a dataset for all applicants and all colleges 200 times.

To compute the counterfactual, we parametrize the distribution of college cutoffs and assume that the following - with probability p_j, c_j is drawn from the following distribution:

$$c_j \sim N(\mu_{jt}, \sigma_{jt}^2)$$

With probability $1 - p_{j,t}$, the cutoff is c_0, the minimum score requirement of first-tier colleges (in which case there are vacancies after all students have been matched). t refers to the round of iteration - given the resulted parametric distribution in round $t-1$, we calculate the ROL that the Optimal Rule specifies, and then calculate the resulted cutoff for each time of the simulation (so we have 200 cutoffs, each of which results from a particular utility realization). Given the resulted 200 cutoffs, we can calculate the probability that there is no vacancy after matching $\tilde{p}(p_{j,t}, \mu_{j,t}, \sigma^2_{j,t})$, the mean (200×1 vector) and standard deviation (200×1 vector) of the 200 cutoffs for each college j:

$$\tilde{\mu}(p_{j,t}, \mu_{j,t}, \sigma^2_{j,t})$$

$$\tilde{\sigma}^2(p_{j,t}, \mu_{j,t}, \sigma^2_{j,t})$$

The initial distribution of college cutoffs is what we have obtained from Section 3.2.

The updating process is that, for the first 2,000 rounds, the rule is that we replace p with $\tilde{p}$, μ with the predicted mean vector $\tilde{\mu}$. For standard deviation, the 1-norm of the new vector is updated be the same as $\tilde{\sigma}$, but nothing else is changed. In other words, the update of standard deviation only involves scaling up and down all components:

$$p_{j,t+1} = \tilde{p}(p_{j,t}, \mu_{j,t}, \sigma_{j,t})$$

$$\mu_{j,t+1} = \tilde{\mu}(p_{j,t}, \mu_{j,t}, \sigma_{j,t})$$

$$\sigma_{j,t+1} = \frac{|\tilde{\sigma}(p_{j,t}, \mu_{j,t}, \sigma_{j,t})|}{|\sigma_{j,t}|}\sigma_{j,t}$$

$|.|$ is the $L1$-norm of j (scaling up or down based on the norm of predicted standard deviation $\tilde{\sigma}(p_{j,t}, \mu_{j,t}, \sigma_{j,t})$)

For the next and last 1,000 rounds, the updating rule mimics "fictitious play" where all the realization of cutoffs after round 2,000 receives equal weight. In other words, for any $t > 2000$:

$$p_{j,t+1} = \frac{t}{t+1}p_{j,t} + \frac{1}{t+1}\tilde{p}(p_{j,t}, \mu_{j,t}, \sigma_{j,t})$$

$$\mu_{j,t+1} = \frac{t}{t+1}\mu_{j,t} + \frac{1}{t+1}\tilde{\mu}(p_{j,t}, \mu_{j,t}, \sigma_{j,t})$$

$$\sigma_{j,t+1} = \frac{t}{t+1}\sigma_{j,t} + \frac{1}{t+1}\tilde{\sigma}(p_{j,t}, \mu_{j,t}, \sigma_{j,t})$$

The estimated parameters we use for counterfactual are $p_{j,3000}$, $\mu_{j,3000}$, $\sigma_{j,3000}$.

D Mathematical Appendix

D.1 Proof of Proposition 1

To prove the point 1, it is sufficient to consider the scenario when $a^* \neq a'$.

This paragraph will prove $a' \in \{a^*, b^*, c^*, d^*\}$. Suppose, for the sake of contradiction, that $a' \notin \{a^*, b^*, c^*, d^*\}$. By the definition of d^*, we know that for any $x \in \mathcal{C}/\{a^*, b^*, c^*\}$, $p_x u_x < p_{d^*} u_{d^*}$ which implies that $p_{a'} u_{a'} < p_{d^*} u_{d^*}$. On the other hand, the definition of the DC Rule implies that $p_{a'} u_{a'} > p_{d^*} u_{d^*}$, which results in contradiction.

This paragraph will prove $p_{a^*} \leq p_{a'}$. We know from previous paragraph that $a' \in \{a^*, b^*, c^*, d^*\}$. When $a^* \neq a'$, $a' \in \{b^*, c^*, d^*\}$. This implies that $u_{a'} \leq u_{a^*}$. Since $p_{a'} u_{a'} \geq p_{a^*} u_{a^*}$, we have $p_{a'} \geq p_{a^*}$.

To prove Point 2, now let us consider the scenario when $a^* = a'$. The recursive structure allows us to recycle the previous proof to show that $b' \in \{b^*, c^*, d^*\}$. Suppose, for the sake of contradiction, that $b' \notin \{b^*, c^*, d^*\}$. By the definition of d^*, we know that for any $x \in \mathcal{C}/\{a^*, b^*, c^*\}$, $p_x u_x < p_{d^*} u_{d^*}$ which implies that $p_{b'} u_{b'} < p_{d^*} u_{d^*}$. On the other hand, the definition of the DC Rule implies that $p_{b'} u_{b'} > p_{d^*} u_{d^*}$, which results in contradiction.

We can then use similar reasoning to prove $p_{b^*} \leq p_{b'}$. We know from previous paragraph that $b' \in \{b^*, c^*, d^*\}$. When $b^* \neq b'$, $b' \in \{c^*, d^*\}$. This implies that $u_{b'} \leq u_{b^*}$. Since $p_{b'} u_{b'} \geq p_{b^*} u_{b^*}$, we have $p_{b'} \geq p_{b^*}$.

Using the proof iteratively, we can prove point 3 and point 4.

D.2 Proof of Proposition 2

This paragraph will prove that $d^* \in \{a', b', c', d'\}$. Suppose, for the sake of contradiction, that $d^* \notin \{a', b', c', d'\}$. Since the definition of the Optimal Rule stipulates that $p_{d^*} u_{d^*} \geq p_x u_x$ for any $x \in \mathcal{C}/\{a^*, b^*, c^*\}$, and the definition of the DC Rule stipulates that $p_y u_y > p_x u_x$ for any $y \in \{a', b', c', d'\}$ and $x \in \mathcal{C}/\{a', b', c', d'\}$, we have $p_y u_y > p_{d^*} u_{d^*}$, which contradicts the implication of the Optimal Rule, where no more than three colleges other than d^* exceeds the single-college expected utility of d^*.

This paragraph will show directly why the prediction holds. When $max\{p_{a^*}, p_{b^*}, p_{c^*}\} > p_{d^*}$, there exists a non-bottom choice such that the non-bottom choice is of higher probability than the bottom choice. Assume WLOG that $p_{a^*} > p_{d^*}$, since $d^* \in \{a', b', c', d'\}$, and that $u_{a^*} \geq u_{d^*}$, it must be that $a^* \in \{a', b', c', d'\}$ and a^* is ranked higher than d^* under the DC rule. This implies the presence of competitiveness reversal in (a', b', c', d').

D.3 Proof of Observation 1

Suppose the ratio reaches the maximum for the pair of a, b. When the two rules are indistinguishable, we know from Prediction 1 that the two lists are identical. This means that

$$u_a \geq u_b \geq u_c \geq u_d$$

and

$$p_a u_a \geq p_b u_b \geq p_c u_c \geq p_d u_d$$

So for any $x \notin \{a, b, c, d\}$, we have

$$p_x u_x \leq p_b u_b = \frac{p_b}{p_a} p_a u_b \leq \frac{p_b}{p_a} p_a u_a$$

The proof of other pairs can be similarly conducted.

D.4 Proof of Observation 2

When s is far above than cutoffs, we always have

$$\mathcal{L}(s) \approx \frac{1}{1} \approx 0$$

When $\sigma_A = \sigma_B \equiv \sigma$, we have

$$\begin{aligned} \mathcal{L}(s) &= \frac{p_A(s)}{p_B(s)} \\ &= \frac{\Phi(\frac{s-\mu_A}{\sigma})}{\Phi(\frac{s-\mu_B}{\sigma})} \end{aligned}$$

When s is far lower than cutoffs, we have

$$\mathcal{L}(s) \approx \frac{\phi(\frac{s-\mu_A}{\sigma})}{\phi(\frac{s-\mu_B}{\sigma})} \approx 0$$

When $\sigma_A < \sigma_B$ and s is far lower than cutoffs, we have

$$\mathcal{L}(s) \approx \frac{\sigma_A \phi(\frac{s-\mu_A}{\sigma_A})}{\sigma_B \phi(\frac{s-\mu_B}{\sigma_B})} \approx 0$$

When $\sigma_A > \sigma_B$ and s is far below than cutoffs, we have

$$\mathcal{L}(s) \approx \frac{\sigma_A \phi(\frac{s-\mu_A}{\sigma_A})}{\sigma_B \phi(\frac{s-\mu_B}{\sigma_B})} \rightarrow +\infty$$

Obviously $\mathcal{L}(s)$ is continuous. Hence any value in between can be achieved by certain values of s.

E Additional Institutional Details

E.1 College Application

The admission process is stratified according to college quality. Before 2019, the colleges were classified into three tiers of decreasing quality: elite (first-tier), public non-elite (second-tier) and private non-elite (third-tier). The latter two categories merged starting in 2019, but the elite category remains unchanged. In this paper, we focus on the admission to elite colleges, which are argued to play a central role in upward mobility in China because of the tremendous value placed on education and the huge return in labor markets (Jia and Li, 2016).

The share of students who are eligible for 1st-tier colleges is roughly $20 \sim 25\%$ of the exam takers. The eligible students on the science track can choose up to four colleges from among 239 elite colleges. For those who are on the humanities track, the total number of elite colleges is 150. Science track students account for more than 80% of the first-tier applicants.

E.2 Priority Score

The priority score is almost completely determined by the College Entrance Exam (CEE)[25]. The CEE is a nationwide closed-book written exam held once a year on June 7th and 8th, with the rare exception that the exam was postponed to July 7th and 8th in 2020 due to COVID-19. To apply for colleges in an admission cycle, all students must take the CEE of the same cycle. In each province, students on the same track (Humanities or Sciences) will take the same exam.[26] As demonstrated in Figure A1, it will take up to two weeks for Ningxia Provincial Education Authorities to grade students' exams. Students will be notified of their exam score and ranking in Ningxia around the 20th-25th of the month in which the exam takes place.

Exams for both tracks include Chinese, Mathematics[27] and English. For each of these subjects, students get an integer score, with the maximum (best) possible being 150 and the minimum being 0. Additionally, students on the Humanities Track take a comprehensive exam on history, politics, and geography, whereas students on the Sciences Track take another exam on physics, chemistry, and biology. This track-specific exam accounts for 300 points. Thus the total score of the CEE (sum of the scores from the four subjects) is 750 points.

[25]Exceptions include winners of international Olympiad contests, students who win sports scholarships, students with exceptional art talent, students who belong to certain minority ethnic groups, etc. These exceptions are also quantified and added to priority scores on top of the exam scores, and are observable in our data.

[26]See Wang et al. (2023) and Li et al. (2021) for more institutional details about the CEE.

[27]Mathematics for the Humanities Track differs from that for the Science Track.

In case of a tie in total score, ranking will be determined by the score in the comprehensive exam in the respective track, Mathematics, and English, in a lexicographic way.

E.3 Correlation in Admission Events

If the probability of meeting the cutoff of one college is correlated with meeting that of another, our assumptions about independence are defied. In this case, assuming independence of admission probability alone may result in suboptimal portfolio choices (Shorrer, 2019; Rees-Jones et al., 2020). However, in our case students know their priority score and ranking by the time of application, and researchers have assumed independence in similar settings (Larroucau and Rios, 2018, 2020).

Nevertheless, to test the potential presence of pairwise correlation, we conduct an empirical test with the same specification as we used in Section 3.2 on the administrative dataset. The difference is that, in this new empirical exercise, the dependent variable is students' second choices, and the regression is run on those who were not admitted by their first choice:

$$1(\text{Admitted to Second Choices})_i = \alpha_2 + \beta_2 \hat{p}_{ij}$$

If the admission to the first choices does not correlate with the admission probability of the second choices, we would expect $\alpha_2 = 0$ and $\beta_2 = 1$, which is the null hypothesis of this exercise.

The estimation results suggest that our estimates of admission probability remain accurate conditional on the rejection of the first choices, suggesting that in our setting pairwise correlation does not significantly alter admission probability. As shown in Columns 3 and 4 of Table B3, $\hat{\alpha_2} = 0.03$ (SE=0.004) and -0.009 (SE=0.006) for the science and humanity tracks, respectively, whereas $\hat{\beta}_2 = 0.990$ (SE=0.006) and 1.026 (SE=0.011) for the science and humanity tracks, respectively. While the null hypothesis is rejected for the slope under humanity track, the deviation from null is quantitatively small.

E.4 College Preferences vs. Major Preferences

Major preference does not affect assignment of college; essentially the Chinese system is a "college-then-major" system (Chen and Kesten, 2017; Calsamiglia et al., 2020). Major studies typically begin in the second year of bachelor education, and since the 2010s, the Ministry of Education (MoE) of PRC has successfully pushed for a lower barrier in major switching [28]. To assess whether major concerns affect risk taking, we asked students which factor they were most concerned about in college applications. We report the relevant statistical analysis in Table B7. In Columns 1 and 2, we regress the dummy indicating whether students consider major to be their top concern on a normalized SES index. Only 13.7% of the survey takers consider major to be their top concern, and the share is slightly lower among the advantaged students.

Consideration of major could affect students' strategy if students think ahead, and want to outcompete their peers who are admitted to the same college in terms of academic ability. To examine its quantitative impact on risk-taking, we regress the estimated unconditional probability of the first choices (Columns 3-4) and beliefs about being admitted to the first choices (Columns 5-6) on those who consider major to be of top concern. As expected, students tend to be more cautious if they consider major to be of top concern, but its quantitative impact on the probability of first choices is less than 10% compared to those who do not regard major as their top consideration. We calibrate its impact on the full sample average of first choice probabilities by calculating the product of the average impact due to major concerns and the share of students who consider major to be important. Reassuringly, it contributes at most a 1.2% increase in the mean unconditional probability of the first choices.

E.5 Subjective Beliefs

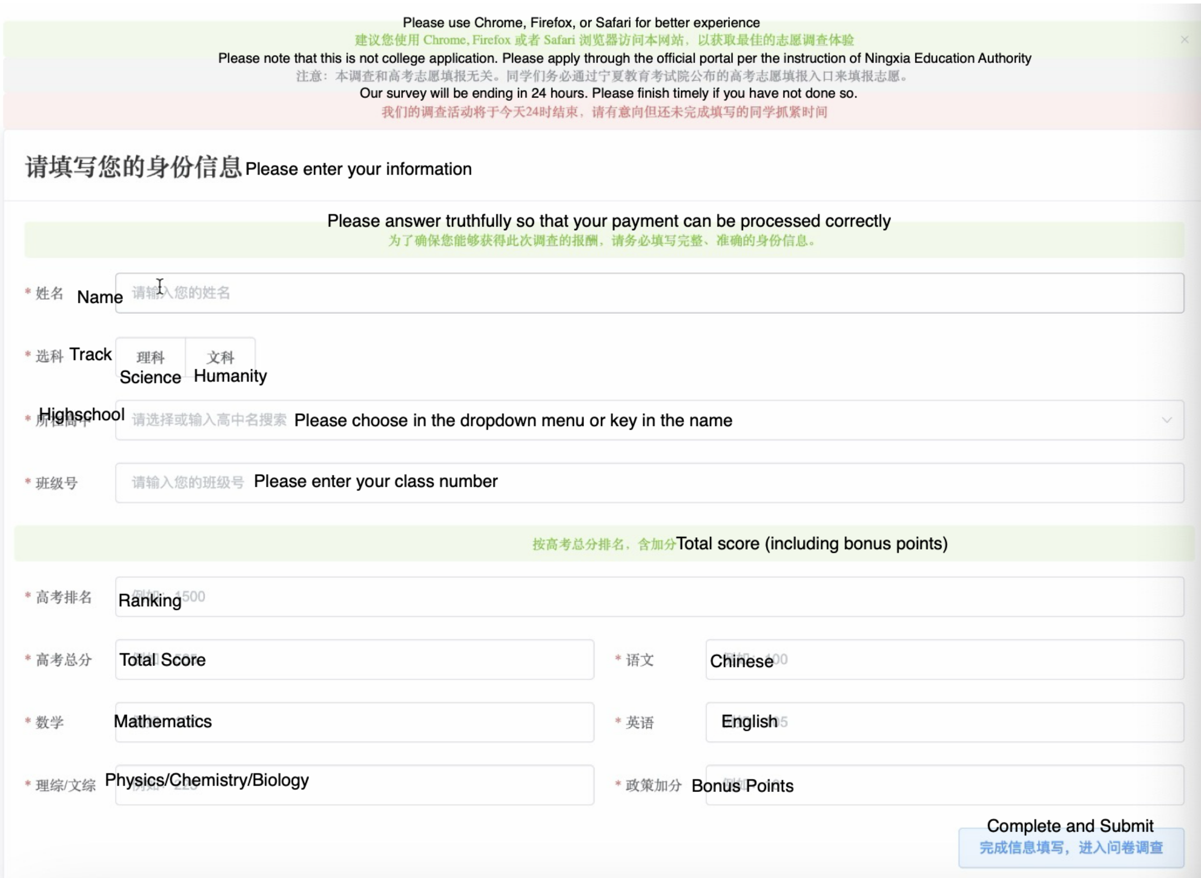
Evidence from Survey A number of recent studies have (Kapor et al., 2020; Arteaga et al., 2021) documented that students may not accurately estimate the probability of admission, especially when the structure of priority scores is more complicated. While in our context the only uncertainty comes from the variation in cutoffs, a one-dimensional object, it is quite unlikely that students' beliefs are perfectly accurate. We elicited students' beliefs regarding the unconditional probability of the four colleges on their lists, and compare the elicited beliefs to the estimated probability that we construct. Subjective beliefs are positively correlated with estimated probability, at a correlation of 0.50. We define belief error as the difference between subjective beliefs and estimated probability, and regress the error on students' socioeconomic index, controlling for other variables including priority score and college preferences. As demonstrated in Table B6, students on average overestimate the chance of admission by 12.1%, and the mean of absolute error is 31.0%.

Despite substantial differences in contexts, the level of these statistics are not far from estimates in Kapor et al. (2020). While socioeconomically advantaged students seem to

[28]The link here is an example. Since this campaign by the MoE, major distinctions have become more coarse and less of a concern to students.

be somewhat more optimistic, 1 SD of increase in SES is associated with less than 1.85% increase in subjective beliefs. The estimated mean of absolute error is also slightly larger among the advantaged (Column 3 and 4), but the magnitude is also small (1 SD of increase in SES is associated with around 1% increase in absolute errors). Taken together, the belief data suggests that belief errors, while substantial, is at best weakly correlated with the demographics.

F Examples of Survey Screenshot


Please use Chrome, Firefox, or Safari for better experience
建议您使用 Chrome, Firefox 或者 Safari 浏览器访问本网站，以获取最佳的志愿调查体验
Please note that this is not college application. Please apply through the official portal per the instruction of Ningxia Education Authority
注意：本调查和高考志愿填报无关。同学们务必通过宁夏教育考试院公布的高考志愿填报入口来填报志愿。
Our survey will be ending in 24 hours. Please finish timely if you have not done so.
我们的调查活动将于今天24时结束，请有意向但还未完成填写的同学抓紧时间

请填写您的身份信息 Please enter your information

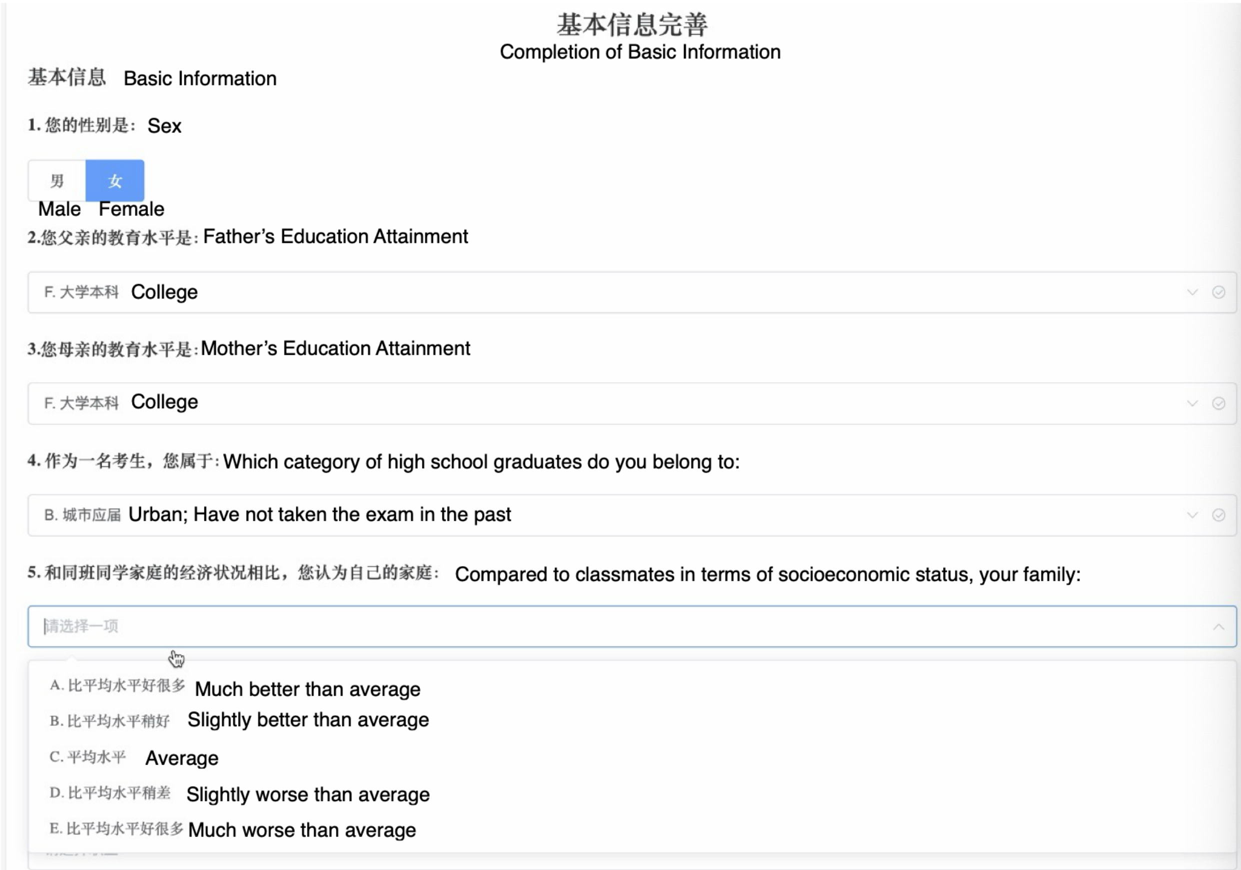
Please answer truthfully so that your payment can be processed correctly
为了确保您能够获得此次调查的报酬，请务必填写完整、准确的身份信息。

* 姓名 Name 请输入您的姓名
* 选科 Track 理科 Science 文科 Humanity
* Highschool 请选择或输入高中名搜索 Please choose in the dropdown menu or key in the name
* 班级号 请输入您的班级号 Please enter your class number

按高考总分排名，含加分Total score (including bonus points)

* 高考排名 Ranking
* 高考总分 Total Score
* 语文 Chinese
* 数学 Mathematics
* 英语 English
* 理综/文综 Physics/Chemistry/Biology
* 政策加分 Bonus Points

Complete and Submit
完成信息填写，进入问卷调查



基本信息完善
Completion of Basic Information

基本信息 Basic Information

1. 您的性别是： Sex

男 女
Male Female

2.您父亲的教育水平是: Father’s Education Attainment

F. 大学本科 College

3.您母亲的教育水平是: Mother’s Education Attainment

F. 大学本科 College

4. 作为一名考生，您属于: Which category of high school graduates do you belong to:

B. 城市应届 Urban; Have not taken the exam in the past

5. 和同班同学家庭的经济状况相比，您认为自己的家庭： Compared to classmates in terms of socioeconomic status, your family:

请选择一项

A. 比平均水平好很多 Much better than average
B. 比平均水平稍好 Slightly better than average
C. 平均水平 Average
D. 比平均水平稍差 Slightly worse than average
E. 比平均水平好很多 Much worse than average

7. 您母亲的工作是： Mother's Occupation

F. 中小学及幼儿园教师 F. Teacher at secondary school, primary school, or kindergarten

8.请选择家庭所在县/市（区）： Location

银川市兴庆区 Yinchuan - Xingqing District

9. 请选择您的出生年月： Birth Year/Month

1996-02 Feb 1996

志愿信息 Application Information

10. 选择三个择校时您最重视的因素（请先选择对您来说最重要的因素） Choose three most important factor for college application, starting from the most important one

第1重要：毕业比例 第2重要：985高校 第3重要：社会声誉

D. 毕业比例 F. 985高校 B. 社会声誉 1st: share of graduation, 2nd: 985 school, 3: reputation

11. 选择三种您最想去的院校类型（请先选择您最满意的院校类型）Please select the three types of college that you want to go most

第1重要：请在下方选择 第2重要：请在下方选择 第3重要：请在下方选择

Please select here 类型，从左到右重要性依次降低

12. 选择三个您最理想的大学所在地理位置（请先选择您最理想的地理位置） Please select the three most preferred locations

第1重要：请在下方选择 第2重要：请在下方选择 第3重要：请在下方选择

Please select here 位置，从左到右重要性依次降低

13. 请从下列选项中选出您比较想在本科学习的专业（请先选择您最想学的专业） Please select the three most preferred majors

第1重要：请在下方选择 第2重要：请在下方选择 第3重要：请在下方选择

Please select here，从左到右重要性依次降低

12. 选择三个您最理想的大学所在地理位置（请先选择您最理想的地理位置）Please select three most preferred locations

第1重要：特大城市（北京，上海，广州，深圳） 第2重要：北方其它主要城市（哈尔滨，沈阳，济南，石家庄，长春） 第3重要：中部主要城市（合肥，郑州，太原，武汉）

2nd: other major cities in Northern China 3rd: major cities in Mid China

A. 特大城市（北京，上海，广州，深圳） E. 北方其它主要城市（哈尔滨，沈阳，济南，石家庄，长春） H. 中部主要城市（合肥，郑州，太原，武汉）

1st: Metropolitan area (Beijing, Shanghai, Guangzhou, Shenzhen)

13. 请从下列选项中选出您比较想在本科学习的专业（请先选择您最想学的专业）Please select three most preferred majors

1st: Econ, Finance, Management (Accounting, Marketing)

第1重要：经济，金融，管理(会计，营销) 第2重要：工学 第3重要：医学

2nd: Engeneering 3rd: Medical

B. 经济，金融，管理(会计，营销) E. 工学 F. 医学

14. 请评估下列信息来源对你了解录取规则，选择志愿院校有多大帮助

Please evaluate how important these factors are for your school selection

* 父母亲戚 Parents/other family members	完全没有帮助 Not at all	有一点帮助 A little	有很大帮助 Quite a lot
* 学校老师 Teachers	完全没有帮助 Not at all	有一点帮助 A little	有很大帮助 Quite a lot
* 同学好友 Classmates	完全没有帮助 Not at all	有一点帮助 A little	有很大帮助 Quite a lot
* 门户网站 Major websites	完全没有帮助 Not at all	有一点帮助 A little	有很大帮助 Quite a lot
* 微信公众号 WeChat public account	完全没有帮助 Not at all	有一点帮助 A little	有很大帮助 Quite a lot
* 招生手册 Admission brochure	完全没有帮助 Not at all	有一点帮助 A little	有很大帮助 Quite a lot
* 报纸/电视节目 Media (newspaer/TV)	完全没有帮助 Not at all	有一点帮助 A little	有很大帮助 Quite a lot

Suppose you are made aware that you are admitted to one college of your choice.

15. 假设在录取过程结束后，您得知自己已被志愿中的某大学录取。请问您是否一定会入读该大学(第一空)；如果填写的是“否”，请问您打算在接下来一年里(第二空)

Please evaluate whether you will matriculate. If not, please state your plan for the next year

第一空 是 否 第二空 在该大学上学

Enter your application for first-tier colleges

填写一批次志愿

Please enter the first-tier colleges that you selected to apply for in the system below,

下面，请您填写您在志愿填报系统中提交的第一批次（非定向）的志愿院校。请完整地填写您填报的志愿院校，以免影响问卷报酬的发放。

Please report your rank-order list completely and ensure correct entry

请至少填写三个学校。只准备填报三个院校的同学不需填写 D 院校；如果您实际只填报的普通一批次院校不足三所，请在B,C院校两空中选择您知道的其它院校。

First-tier college
一本志愿表（本科一批）

Application # 志愿号	院校 School
First-tier choice A 一批次 A 院校	北京大学 Peking University
First-tier choice B 一批次 B 院校	北京交通大学 Beijing Jiaotong University
First-tier choice C 一批次 C 院校	北京邮电大学 Beijing University of Posts and Telecommunications
First-tier choice D 一批次 D 院校	北京外国语大学 Beijing Foreign Studies University

为了更加了解宁夏考生在填报志愿时的想法和困难，我们需要您回答一些与志愿填报有关的问题。 正确的回答越多，您将会获得越多的报酬。

您的回答有助于我们发现考生对录取规则理解的偏差，为高考的改革提供更多线索。因此，我们希望您认真阅读题目，尽力得到准确的答案。

在本情景中，你所获得的额外报酬和韩梅梅对录取院校的满意度相同。比如说，如果她对录取院校的满意度为10，那么你将额外获得10元。

In this scenario, the extra reward you receive is equal to Han Meimei's satisfaction with the admitted university. For example, if her satisfaction with the admitted university is 10, then you will receive an extra 10 Yuan.

韩梅梅没有告诉您高考分数，不过您知道她的分数远远超出一本线。目前，她还没确定一批次A院校该填什么：

Han Meimei didn't tell you her Gaokao score, but you know that her score is well above a certain cutoff. Currently, she needs to decide which university to list as her first-choice A-tier school:

志愿号 Rank-Order List Number	院校 University	满意度 Satisfaction Index
一批次A院校 Tier I - Choice A	?	?
一批次B院校 Tier I - Choice B	兰州大学 Lanzhou University	20
一批次C院校 Tier I - Choice C	宁夏医科大学 Ningxia Medical University	20
一批次D院校 Tier I - Choice D	宁夏大学 Ningxia University	20

She has learned from the admissions office that her score is well above the historical cutoff for Ningxia University. If she is not admitted to her first choice (A-tier A), then she will definitely be admitted to one of Ningxia University, Lanzhou University, or Ningxia Medical University.

她已经从招生办了解到，自己的分数远远超过宁夏大学往年的分数线。如果没有被自己的第一选择（一批次A院校）录取，那么她一定会被宁夏大学或者兰州大学，宁夏医科大学录取。在这种情况下，她的满意度是20，您一定能获得20元。如果她达到一批次院校分数线，但放弃入学资格，她的满意度是0，您将不会获得任何奖励。

In this case, her satisfaction is 20, and you will definitely receive 20 yuan. If she reaches the A-tier score cutoff but gives up enrollment, her satisfaction will be 0, and you will not receive any reward.

经过资料查询，她认为如果被四川大学录取，她的满意度为 **25**，自己的分数达到四川大学分数线的可能性为 **50%**。

After researching, she believes that if admitted to Sichuan University, her satisfaction would be 25, and she has a 50% chance of meeting the cutoff for it.

她估计自己的分数达到厦门大学分数线的可能性为 **25%** 。不过，她还不确定自己对厦门大学的满意程度，需要您帮她做一些计划。

She estimates a 25% chance of reaching the cutoff for Xiamen University. However, she is unsure how satisfied she would be with Xiamen University and needs your help making a plan.

韩梅梅认为她的风险偏好和你的相同，因此请根据你自己对风险的偏好为韩梅梅挑选院校。

Han Meimei believes her risk preference is the same as yours. Therefore, please choose a school for her based on your own risk preference.

请回答以下问题来确认您对情景的理解：

Please answer the following questions to confirm your understanding of the scenario:

1. 如果韩梅梅未被一批次A志愿录取，那么您将获得________

If Han Meimei is not admitted to the A-tier first choice, you will receive

A. 5元	B. 20元	C. 无法确定
5 Yuan	20 Yuan	Unsure

2. 如果韩梅梅被四川大学录取，那么您将获得________

A. 20元	B. 25元	C. 无法确定
20 Yuan	25 Yuan	Unsure

志愿号 Rank-Orderl list Number	院校 University	满意度 Satisfaction Index
一批次A院校 Tier I - Choice A	?	?
一批次B院校 Tier I - Choice B	兰州大学 Lanzhou University	20
一批次C院校 Tier I - Choice C	宁夏医科大学 Ningxia Medical University	20
一批次D院校 Tier I - Choice D	宁夏大学 Ningxia University	20

She has learned from the admissions office that her score is well above the historical cutoff for Ningxia University. If she is not admitted to her first choice (A-tier A), then she will definitely be admitted to one of Ningxia University, Lanzhou University, or Ningxia Medical University.

她已经从招生办了解到，自己的分数远远超过宁夏大学往年的分数线。 如果没有被自己的第一选择（一批次A院校）录取，那么她一定会被宁夏大学或者兰州大学，宁夏医科大学录取。 在这种情况下，她的满意度是20，您一定能获得20元。 如果她达到一批次院校分数线，但放弃入学资格，她的满意度是0，您将不会获得任何奖励。

In this case, her satisfaction is 20, and you will definitely receive 20 yuan. If she reaches the A-tier score cutoff but gives up enrollment, her satisfaction will be 0, and you will not receive any reward.

经过资料查询，她认为如果被四川大学录取，她的满意度为 **25**，自己的分数达到四川大学分数线的可能性为 **50%**。

After researching, she believes that if admitted to Sichuan University, her satisfaction would be 25, and she has a 50% chance of meeting the cutoff for it.

她估计自己的分数达到厦门大学分数线的可能性为 **25%** 。不过，她还不确定自己对厦门大学的满意程度，需要您帮她做一些计划。

She estimates a 25% chance of reaching the cutoff for Xiamen University. However, she is unsure how satisfied she would be with Xiamen University and needs your help making a plan.

韩梅梅认为她的风险偏好和你的相同，因此请根据你自己对风险的偏好为韩梅梅挑选院校。

Han Meimei believes her risk preference is the same as yours. Therefore, please choose a school for her based on your own risk preference.

请您告诉她，在下列七种情况下，她应该选择哪一所大学作为一批次A院校。在下面的每道小题中，每题都有两个选项。其中，选项A的报酬是不变的，选项B的报酬随着题号的增加而上升。因此，随着题号的增加，你至多只能有一次机会从选项A切换到选项B。

Please tell her, in the following seven scenarios, which university she should choose as her A-tier first-choice school. In each of the following questions, there are two options. Option A offers a constant reward, while Option B's reward increases with the question number. Therefore, as the question number increases, you may switch from Option A to Option B at most once.

	A	B
* 1(1):	A.对四川大学的满意度是25，上线概率是50%。 Sichuan University: Satisfaction index is 25, probability of meeting cutoff is 50%.	B.对厦门大学的满意度是30，上线概率25%。 Xiamen University: Satisfaction index is 30, probability of meeting cutoff is 25%.
* 1(2):	A.对四川大学的满意度是25，上线概率是50%。 Sichuan University: Satisfaction index is 25, probability of meeting cutoff is 50%.	B.对厦门大学的满意度是35，上线概率25%。 Xiamen University: Satisfaction index is 35, probability of meeting cutoff is 25%.
* 1(3):	A.对四川大学的满意度是25，上线概率是50%。 Sichuan University: Satisfaction index is 25, probability of meeting cutoff is 50%.	B.对厦门大学的满意度是40，上线概率25%。 Xiamen University: Satisfaction index is 40, probability of meeting cutoff is 25%.
* 1(4):	A.对四川大学的满意度是25，上线概率是50%。 Sichuan University: Satisfaction index is 25, probability of meeting cutoff is 50%.	B.对厦门大学的满意度是45，上线概率25%。 Xiamen University: Satisfaction index is 45, probability of meeting cutoff is 25%.
* 1(5):	A.对四川大学的满意度是25，上线概率是50%。 Sichuan University: Satisfaction index is 25, probability of meeting cutoff is 50%.	B.对厦门大学的满意度是50，上线概率25%。 Xiamen University: Satisfaction index is 50, probability of meeting cutoff is 25%.
* 1(6):	A.对四川大学的满意度是25，上线概率是50%。 Sichuan University: Satisfaction index is 25, probability of meeting cutoff is 50%.	B.对厦门大学的满意度是55，上线概率25%。 Xiamen University: Satisfaction index is 55, probability of meeting cutoff is 25%.
* 1(7):	A.对四川大学的满意度是25，上线概率是50%。 Sichuan University: Satisfaction index is 25, probability of meeting cutoff is 50%.	B.对厦门大学的满意度是60，上线概率25%。 Xiamen University: Satisfaction index is 60, probability of meeting cutoff is 25%.

Risk Preference Game
风险偏好测试游戏

在下面的每道小题中，每题都有两个选项。其中选项A的报酬是不变的，选项B的报酬随着题号的增加而上升。因此，随着题号的增加，你至多只能有一次机会从选项A切换到选项B。In each of the questions below, there are two options. The payment for Option A remains constant across questions, while the payment for Option B increases as you progress down the list.

每一道小题都有相同的可能性被抽中，在该小题中您更偏好的那个选项将作为完成本问卷的额外报酬，因此我们建议您认真完成所有的题目。Therefore, as you move through the questions sequentially, you would switch from Option A to Option B at most once.

* 1(1):	A.50%的概率获得 25元，50%的概率获得20元 Get 25 Yuan with 50% of chance Get 20 Yuan with 50% of chance	B.25%的概率获得 30 元，75%的概率获得20元 Get 30 Yuan with 25% of chance Get 20 Yuan with 75% of chance
* 1(2):	A.50%的概率获得 25元，50%的概率获得20元 Get 25 Yuan with 50% of chance Get 20 Yuan with 50% of chance	B.25%的概率获得 35 元，75%的概率获得20元 Get 35 Yuan with 25% of chance Get 20 Yuan with 75% of chance
* 1(3):	A.50%的概率获得 25元，50%的概率获得20元 Get 25 Yuan with 50% of chance Get 20 Yuan with 50% of chance	B.25%的概率获得 40 元，75%的概率获得20元 Get 40 Yuan with 25% of chance Get 20 Yuan with 75% of chance
* 1(4):	A.50%的概率获得 25元，50%的概率获得20元 Get 25 Yuan with 50% of chance Get 20 Yuan with 50% of chance	B.25%的概率获得 45 元，75%的概率获得20元 Get 45 Yuan with 25% of chance Get 20 Yuan with 75% of chance
* 1(5):	A.50%的概率获得 25元，50%的概率获得20元 Get 25 Yuan with 50% of chance Get 20 Yuan with 50% of chance	B.25%的概率获得 50 元，75%的概率获得20元 Get 50 Yuan with 25% of chance Get 20 Yuan with 75% of chance
* 1(6):	A.50%的概率获得 25元，50%的概率获得20元 Get 25 Yuan with 50% of chance Get 20 Yuan with 50% of chance	B.25%的概率获得 55 元，75%的概率获得20元 Get 55 Yuan with 25% of chance Get 20 Yuan with 75% of chance
* 1(7):	A.50%的概率获得 25元，50%的概率获得20元 Get 25 Yuan with 50% of chance Get 20 Yuan with 50% of chance	B.25%的概率获得 60 元，75%的概率获得20元 Get 60 Yuan with 50% of chance Get 20 Yuan with 75% of chance

在了解了更多情况后，**韩梅梅对B,C,D中院校的满意度降低了**

After learning more information, Hai Meimei's satisfaction with the lower-ranked universities decreases.

志愿号 Application #	院校 University	满意度 Satisfaction Index
一批次A院校 First-Tier Choice A	?	?
一批次B院校 First-Tier Choice B	兰州大学 Lanzhou University	5
一批次C院校 First-Tier Choice C	宁夏医科大学 Ningxia Medical University	5
一批次D院校 First-Tier Choice D	宁夏大学 Ningxia University	5

After researching, she believes that if admitted to Sichuan University, her satisfaction would be 25, and she has a 50% chance of meeting the cutoff for it.

经过资料查询，她认为如果被 四川大学录取，她的满意度为 **25** ，自己的分数达到 四川大学分数线的可能性为 **50%**。

She estimates a 25% chance of reaching the cutoff for Xiamen University. However, she is unsure how satisfied she would be with Xiamen University and needs your help making a plan.

她估计自己的分数达到厦门大学分数线的可能性为 **25%** 。不过，她还不确定自己对厦门大学的满意程度，需要您帮她做一些计划。

Please tell her, in the following seven scenarios, which university she should choose as her A-tier first-choice school. In each of the following questions, there are two options. Option A offers a constant reward, while Option B's reward increases with the question number. Therefore, as the question number increases, you may switch from Option A to Option B at most once.

请您告诉她，在下列七种情况下，她应该选择哪一所大学作为一批次A院校。在下面的每道小题中，每题都有两个选项。其中，选项A的报酬是不变的，选项B的报酬随着题号的增加而上升。因此，随着题号的增加，你至多只能有一次机会从选项A切换到选项B。

	A	B
* 1(1):	**A.对四川大学的满意度是25，上线概率是50%。** Sichuan University: Satisfaction index is 25, probability of meeting cutoff is 50%.	**B.对厦门大学的满意度是30，上线概率25%。** Xiamen University: Satisfaction index is 30, probability of meeting cutoff is 25%.
* 1(2):	**A.对四川大学的满意度是25，上线概率是50%。** Sichuan University: Satisfaction index is 25, probability of meeting cutoff is 50%.	**B.对厦门大学的满意度是35，上线概率25%。** Xiamen University: Satisfaction index is 35, probability of meeting cutoff is 25%.
* 1(3):	**A.对四川大学的满意度是25，上线概率是50%。** Sichuan University: Satisfaction index is 25, probability of meeting cutoff is 50%.	**B.对厦门大学的满意度是40，上线概率25%。** Xiamen University: Satisfaction index is 40, probability of meeting cutoff is 25%.
* 1(4):	**A.对四川大学的满意度是25，上线概率是50%。** Sichuan University: Satisfaction index is 25, probability of meeting cutoff is 50%.	**B.对厦门大学的满意度是45，上线概率25%。** Xiamen University: Satisfaction index is 45, probability of meeting cutoff is 25%.
* 1(5):	**A.对四川大学的满意度是25，上线概率是50%。** Sichuan University: Satisfaction index is 25, probability of meeting cutoff is 50%.	**B.对厦门大学的满意度是50，上线概率25%。** Xiamen University: Satisfaction index is 50, probability of meeting cutoff is 25%.
* 1(6):	**A.对四川大学的满意度是25，上线概率是50%。** Sichuan University: Satisfaction index is 25, probability of meeting cutoff is 50%.	**B.对厦门大学的满意度是55，上线概率25%。** Xiamen University: Satisfaction index is 55, probability of meeting cutoff is 25%.
* 1(7):	**A.对四川大学的满意度是25，上线概率是50%。** Sichuan University: Satisfaction index is 25, probability of meeting cutoff is 50%.	**B.对厦门大学的满意度是60，上线概率25%。** Xiamen University: Satisfaction index is 60, probability of meeting cutoff is 25%.

Scenario: Help Li Hua with his college application

情景：帮助李华报志愿

Suppose Li Hua's score is the same as yours.
李华的分数和您的分数相同。下表列出了他正在考虑的学校以及满意度：
The table below listed the schools he is considering and his satisfaction on a scale of 0~100 if he would be admitted by them

上海交通大学 Shanghai Jiaotong University	50	北京航空航天大学 Beihang University	10
浙江大学 Zhejiang University	35	复旦大学(医学院) Fudan University (Medical School)	30
复旦大学 Fudan University	50	中国科学技术大学 University of Science and Technology of China	30
天津大学 Tianjin University	10	同济大学 Tongji University	10

Now, Li Hua needs your help with his application. In this question, the reward is the same as satisfaction.
李华现在需要您帮助他填报志愿。**在此题中，奖励额度与满意度相同。**

比如说，如果李华被 上海交通大学 录取， 那么我们将给您 **50** 元报酬。如果李华未被任何学校录取，那么您将不能从本题中获得任何报酬。
For example, if Li Hua is admitted by Shanghai Jiaotong University, you will get 50 Yuan. If Li Hua is not admitted by any first-tier university you will get 0 Yuan.
在仔细阅读该情景后，请问答下面的问题：
Please answer the questions:

1. 如果李华被 北京航空航天大学录取，那么您将获得______元奖励
If Li Hua is admitted by Beihang University, then you will get

A. 0 | B. 10 | C. 20

2. 如果李华没有被任何一本志愿录取，那么您将获得______元奖励
If Li Hua is not admitted by any first-tier university, then you will get

A. 0 | B. 5 | C. 10

如果您可以帮助李华选择上述学校的中的四所作为一批次志愿，您会如何选择？

If you could help Li Hua to select four universities from what is listed above as his first-tier rank-order lists, what would you choose?

一本志愿表（本科一批）First-Tier Application

志愿号	院校
First Tier - Choice A 一批次 A 院校	上海交通大学 Shanghai Jiaotong University
First Tier - Choice B 一批次 B 院校	复旦大学(医学院) Fudan University (Medical School)
First Tier - Choice C 一批次 C 院校	复旦大学 Fudan University
First Tier - Choice D 一批次 D 院校	天津大学 Tianjin University

If Li Hua is willing to select only two universities as his choice for first-tier universities, with the other two spots (choice C, D) left blank, which two universities would you choose for choice A and B?

如果李华只愿意选择两所学校作为一批次志愿，C,D院校的位置空着，您会如何选择？

一本志愿表（本科一批）First-Tier Application

志愿号	院校
First Tier - Choice A 一批次 A 院校	选择或搜索院校
First Tier - Choice B 一批次 B 院校	

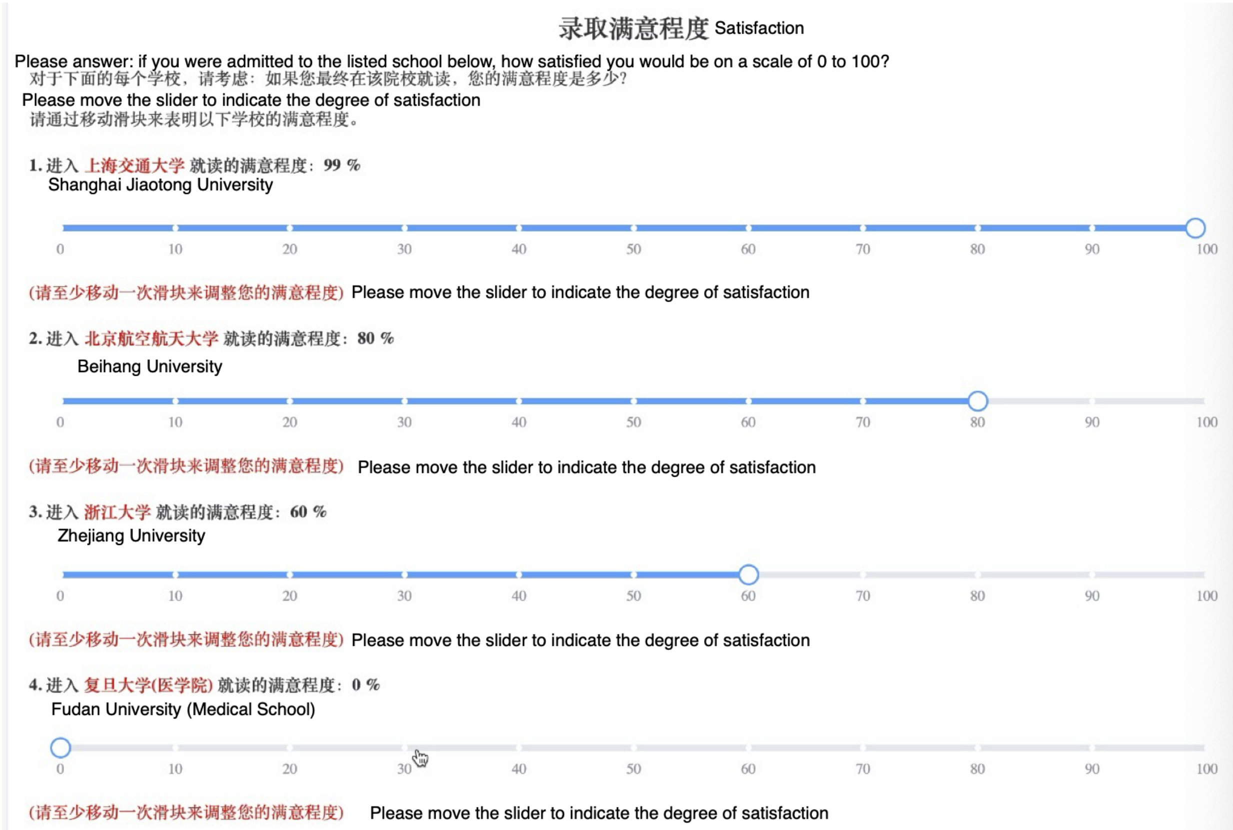
推荐院校名单

- 上海交通大学 Shanghai Jiaotong University
- 北京航空航天大学 Beihang University
- 浙江大学 Zhejiang University
- 复旦大学(医学院) Fudan University (Medical School)
- 复旦大学 Fudan University
- 中国科学技术大学 University of Science and Technology of China
- 天津大学 Tianjin University

如果李华只愿意选

志愿号	
First Tier - Choice A 一批次 A 院校	


志愿号
院校
First-Tier Application
First-Tier Application
一批次A院校
上海交通大学
一批次B院校
复旦大学(医学院)
First Tier - Choice C
一批次C院校
复旦大学 Fudan University
First Tier - Choice D
一批次D院校
天津大学 Tianjin University
If Li Hua is willing to select only two universities as his choice for first-tier universities, with the other two spots (choice C, D) left blank, which two universities would you choose for choice A and B?
如果李华只愿意选择两所学校作为一批次志愿，C,D院校的位置空着，您会如何选择?
一本志愿表（本科一批）First-Tier Application
志愿号
院校
First Tier - Choice A
一批次A院校
上海交通大学 Shanghai Jiaotong University
First Tier - Choice B
一批次B院校
天津大学 Tianjin University
If Li Hua is willing to select only one universities as his choice for first-tier universities, with the other three spots (choice B, C, D) left blank, which two universities would you choose for choice A?
如果李华只愿意选择一所学校，B,C,D院校的位置都空着，您会如何选择?
一本志愿表（本科一批）First-Tier Application
志愿号
院校
First Tier - Choice A
一批次A院校
Please select a university



录取满意程度 Satisfaction

Please answer: if you were admitted to the listed school below, how satisfied you would be on a scale of 0 to 100?
对于下面的每个学校，请考虑：如果您最终在该院校就读，您的满意程度是多少？
Please move the slider to indicate the degree of satisfaction
请通过移动滑块来表明以下学校的满意程度。

1. 进入 上海交通大学 就读的满意程度：**99 %**
Shanghai Jiaotong University

0 10 20 30 40 50 60 70 80 90 100

(请至少移动一次滑块来调整您的满意程度) Please move the slider to indicate the degree of satisfaction

2. 进入 北京航空航天大学 就读的满意程度：**80 %**
Beihang University

0 10 20 30 40 50 60 70 80 90 100

(请至少移动一次滑块来调整您的满意程度) Please move the slider to indicate the degree of satisfaction

3. 进入 浙江大学 就读的满意程度：**60 %**
Zhejiang University

0 10 20 30 40 50 60 70 80 90 100

(请至少移动一次滑块来调整您的满意程度) Please move the slider to indicate the degree of satisfaction

4. 进入 复旦大学(医学院) 就读的满意程度：**0 %**
Fudan University (Medical School)

0 10 20 30 40 50 60 70 80 90 100

(请至少移动一次滑块来调整您的满意程度) Please move the slider to indicate the degree of satisfaction

现在请考虑以下学校录取您的可能性
Please evaluate the chance that your score meets the admission cutoff of the schools listed below

1. 您的分数高于或等于 北京大学 录取分数线的可能性：69% (可能性较大)
The chance that your score is higher than or equal to the admission cutoff of Peking University: 69% (probably)

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

(请移动滑块来调整可能性的大小) (Please move the slider to adjust the entry)

2. 您的分数高于或等于 北京交通大学 录取分数线的可能性：50% (有相当的可能性)
The chance that your score is higher than or equal to the admission cutoff of Beijing Jiaotong University: 50% (likely)

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

(请移动滑块来调整可能性的大小) (Please move the slider to adjust the entry)

3. 您的分数高于或等于 北京邮电大学 录取分数线的可能性：0% (完全不可能)
The chance that your score is higher than or equal to the admission cutoff of Beijing University of Posts and Telecommunications: 0% (impossible)

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

(请移动滑块来调整可能性的大小) (Please move the slider to adjust the entry)

4. 您的分数高于或等于 北京外国语大学 录取分数线的可能性：0% (完全不可能)
The chance that your score is higher than or equal to the admission cutoff of Beijing Foreign Studies University: 0% (impossible)

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%