When to Visit:
Information Acquisition in College Admissions*

Youngwoo Koh† Wooyoung Lim‡

May 3, 2025

Abstract

We theoretically and experimentally study centralized college admissions when both colleges and students lack full information about their potential matching partners; Colleges evaluate students based on their exam scores; Students learn the matching value of each college via costly information acquisition. In a centralized matching via Gale and Shapley's deferred acceptance algorithm, it is incentive-compatible for students to acquire information before submitting their rank-order lists. However, the uncertainty of the final assignment lowers the expected gain from learning, thereby reducing social welfare, compared to a scenario without such uncertainty. Our experiments demonstrate that the welfare loss is greater with more imperfect exam scores. The empirical social welfare obtained in our experimental treatments is consistently lower than the theoretical welfare, and we identify non-equilibrium learning as a main contributor.

Keywords: College admissions, Admission uncertainty, Information acquisition, Timing of learning, Laboratory experiments

***JEL* classification numbers**: C78, C91, D47, D83

1 Introduction

We investigate, both theoretically and experimentally, how matching markets work when both sides of the market lack full information about their potential matching partners in the context of college admissions. Specifically, our study centers on a college admission market

*We are grateful to Dongkyu Chang, Yan Chen, Rustamdjan Hakimov, Yinghua He, Eun Jeong Heo, Doruk Iris, Duk Gyoo Kim, Jinwoo Kim, Siqi Pan, Keeyoung Rhee, and Jie Zheng for their helpful comments and suggestions and to Chun Yeung Knowlton Chan, Hanyu Xiao, and Shiyu (Jake) Zhang for their excellent research assistance. We have benefited from comments by conference participants at the 2021 KER-KAEA International Conference, 2022 Asian Pacific ESA Meeting, the 2022 Conference on Mechanism and Institution Design at NUS, 2022 SKKU International Conference, 2023 SAET conference, 2023 Greater Bay Area Market Design Workshop at University of Macau, and seminar participants at Seoul National University and City University of Hong Kong. This study is supported by a grant from the Research Grants Council of Hong Kong (Grant No. GRF16504523). Koh is also supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2024S1A5A2A03038509).

†Department of Economics, Seoul National University. *Email:* `yw.koh@snu.ac.kr`

‡Department of Economics, HKUST. *Email:* `wooyoung@ust.hk`

where colleges prioritize the caliber of their incoming students but base their decisions on standardized test scores, which are only an approximate indicator of a student's true ability. Students also lack comprehensive knowledge about the distinctive attributes of various colleges, necessitating the expenditure of resources to acquire this information to effectively formulate their preferences.[1] Our analysis, both theoretical and experimental, aims to elucidate the interaction between the imperfectness of exam scores as a measure of student ability and the incentives for students to acquire information.

In practice, it is often observed that market participants are unaware of their own preferences until they gather information about the relevant attributes of potential matching partners. This is particularly true in college admission markets. The process of forming preferences regarding colleges is intricate and time-intensive. For example, many universities host undergraduate information days and campus tours, which are designed to furnish prospective students and their parents with details about their academic programs. Admission discussions, curriculum overviews, and alumni-sharing sessions serve similar purposes. However, attending all such events across various institutions is typically expensive, forcing students to selectively invest in gathering information about specific colleges.[2]

Our model highlights that students' incentives to acquire information in college admission markets are affected not only by the cost of information acquisition but also by the admission uncertainty created by the imperfections of exam scores. We consider an admission system using Gale and Shapley's (1962) deferred acceptance (DA) algorithm. In our stylized model, there are two colleges and a unit mass of students. Each college has a limited capacity, and each student is endowed with an ability which is her private information. Colleges evaluate students based on exam scores, which are positively and imperfectly correlated with students' abilities but are unknown to them. Students have the same prior values for colleges, capturing academic quality or public school ranking, but each student can learn her idiosyncratic preference or suitability of the colleges by incurring costs. As a result, her college rankings can be reversed based on the learning outcomes.

The DA mechanism asks students to submit their preference rankings over the colleges, and so students acquire information before submitting their rank-order lists (ROLs). Students' equilibrium learning decisions are as follows: they choose any college to learn its

[1] Recent contributions to this area of research include studies by Chen and He (2021a,b), Artemov (2021), and Hakimov, Kübler, and Pan (2023), which examine the implications of information acquisition under various matching mechanisms. We will discuss them further in the following literature review section.

[2] Visiting colleges and universities in expensive cities in the US can cost $2,000 for one trip in 2021. See "Set a Budget for College Tours" by Farran Powell in the USNews (https://www.usnews.com/education/best-colleges/articles/2016-07-12/set-a-budget-for-college-tours) and "College Visit Expenses: Don't Overlook These Smaller Costs" by Melissa Brock (https://collegefinance.com/college-admissions/college-visit-expenses-dont-overlook-these-smaller-costs).

suitability as long as the learning cost is not too high and then learn the suitability of the remaining college when the realized value of the suitability of the former college is neither too small nor too large relative to the difference in prior values between the two colleges. Once students submit their ROLs, the allocation is characterized by cutoff scores $\hat{s}_i > \hat{s}_j$ such that students are assigned to the most preferred college when their scores are higher than $\hat{s}_i$, while those with scores in between the two cutoffs are assigned to college j.

When students decide whether to learn the suitability of any college, without knowing the exact exam score, they face uncertainty on admissions because ranking one college over the other does not guarantee assignment to the preferred college. Crucially, this uncertainty depends on the *imperfectness* of exam scores in terms of how accurately they reflect students' true abilities. For example, if exam scores perfectly reflect students' intrinsic abilities, then only high-ability students will be assigned to any college, completely eliminating the admission uncertainty. In contrast, if exam scores are fully random, even high-ability students will encounter admission uncertainty.

To experimentally examine how admission uncertainty affects students' information acquisition incentives and their welfare, we compare DA with a hypothetical benchmark admission system in which students make learning decisions after knowing their admission outcomes. The benchmark is implemented through a simple version of decentralized matching in which students apply to colleges without incurring application costs, colleges admit students based on their scores, and students then choose among the colleges that admitted them.[3] In this benchmark, it is straightforward that students would acquire information after being admitted by both colleges. The idea behind this consideration is that, if students acquire information after knowing where they are admitted, and thereby face no admission uncertainty, then the imperfectness of exam scores should not affect their learning decisions. Moreover, allocations in this benchmark are characterized by cutoff scores: high-score students are admitted by both colleges, while mid-score students are admitted by only one. This allows for a clear comparison of the learning behavior of the students in DA and the benchmark.

Admission uncertainty creates an asymmetry between DA and the benchmark by affecting who learns the suitability (*extensive learning margin*) and how much they learn (*intensive learning margin*). Conditional on learning the suitability of one college, students in DA have weaker incentives to acquire information about the other college, compared to those in the benchmark, precisely because they remain uncertain about their final assignment. This results in a lower intensive learning margin under DA. Moreover, while in the benchmark,

[3] It is not necessary that the benchmark is implemented by such a decentralized matching. We discuss alternative matching mechanisms that yield the same outcome in Remark 2.

only students admitted by both colleges choose to learn about at least one college, students in DA must weigh both the cost of learning and the probability of admission. This makes the extensive learning margin under DA varies with the cost of learning. When the cost is high, only high-ability students choose to learn, leading not only to less intensive learning but also to a smaller share of students acquiring any information compared to the benchmark. When the cost is low, lower-ability students begin to learn as well; however, their decisions have limited impact on higher-ability students, who are more likely to be assigned to their preferred colleges. In both cases, students' overall welfare under DA is lower than under the benchmark.

While the theoretical analysis shows that DA leads to lower welfare than the benchmark, the way imperfect exam scores affect students' learning decisions is not straightforward. At one extreme, when exam scores perfectly reflect students' intrinsic abilities, high-ability students have a strong incentive to acquire information because they can be confident of admission to their preferred colleges. As exam scores become noisier, however, high-ability students are more likely to receive low scores, reducing their chances of admission to top-choice colleges—while the opposite holds for low-ability students, who may benefit from unexpectedly high scores. In this way, admission uncertainty asymmetrically affects students based on ability. To derive optimal learning strategies as predicted by theoretical analysis, students in practice need to form correct beliefs on admission probabilities, which depend on the other students' learning decisions as well as their abilities. This is rather challenging for students in real life, so they may behave sub-optimally, which warrants an empirical analysis.

We propose an experimental design to study how admission uncertainty derived from different timings of learning and the imperfectness of exam scores provide different incentives for students to acquire information. We implement our experiment with four treatments: a treatment with high perfectness of exam scores and a treatment with low perfectness for each of DA and the benchmark system. We develop a simple experimental environment that allows us to observe *whether* each individual subject acquires information about one or both colleges, *how* this is influenced by the imperfectness of exam scores, and *when* the information is acquired.[4]

We find that the observed learning behaviors in our experiment are largely consistent with the theoretical predictions. First, under DA, most students learned before submitting their top-choice colleges, unless their exam scores were so low that they had no chance to be admitted by the "better" college (having a higher admission cutoff). Second, in the

[4]Specifically, we separate individuals' learning decisions from all other decisions including the choice of the top college in DA and the application decision in the benchmark by providing two decision panels that run independently on their screen. See Section 3 for more details.

benchmark, most students learned after being admitted by both colleges, unless their exam scores are so high that they will surely be admitted by any college they apply. Across all treatments, students who had already learned the suitability of one college further learned the suitability of the other college substantially more often when the suitability of the first college was neither too high nor too low, as the theory suggests. However, inconsistent with theory but in line with findings from the recent experimental studies on school choice (e.g., Chen and He, 2021b; Hakimov, Kübler, and Pan, 2023), we observe deviations from the equilibrium learning. Both over-learning and under-learning occurred, though under-learning was more frequent and of greater magnitude.

Our experimental data confirm the key welfare implication of our model. When the exam perfectness is high, the empirical social welfare obtained in DA is not significantly different from that in the benchmark admission system. However, when the exam perfectness is low, the empirical social welfare obtained in DA is significantly smaller than that in the benchmark admission system. Nevertheless, social welfare obtained in each of our experimental treatments was consistently lower than the theoretical welfare level. The observed discrepancy is due not only to the non-equilibrium learning discussed above but also other types of non-equilibrium decisions, including mistakes in the top-choice college submission in DA and in the application decisions in the benchmark, and mistakes in the attendance decisions. We decompose the welfare losses (relative to the equilibrium predictions) and identify that non-equilibrium learning is the main contributor to the observed welfare loss in all treatments.

Related literature.

Previous studies in the matching literature have examined how costly information acquisition affects matching outcomes in college admissions. Chen and He (2021a) theoretically analyze students' incentives under DA and the Boston (Immediate Acceptance) mechanisms, showing that students have an incentive to learn their own cardinal and others' preferences only under the Boston mechanism. Chen and He (2021b) is an experimental companion paper of Chen and He (2021a). Unlike theirs, our paper does not compare different matching mechanisms. It also differs in learning technologies: students in their seting can learn both their own and others' preferences, whereas students in ours can only learn their own. Nevertheless, consistent with their findings, we also find suboptimal learning in the laboratory.

Our work aligns with Artemov (2021) and Hakimov, Kübler, and Pan (2023) in exploring the effects of admission uncertainty on welfare.[5] Artemov (2021) shows that in the random

[5]See also Bade (2015) who shows that in a house allocation problem, serial dictatorship makes agents know their exact choice set when they make learning decisions and proves that it is the unique Pareto-optimal, strategy-proof, and non-bossy mechanism when agents may acquire information on their own preferences.

serial dictatorship (RSD), students gather less information than the social optimum. He also suggests some policies to reduce admission uncertainty, thereby enhancing students' welfare. Our model incorporates RSD and the serial dictatorship (SD) as special cases when the exam scores are fully random and perfectly reveal abilities, respectively, and shows that SD results in higher welfare than RSD as consistent with Artemov (2021).

Hakimov, Kübler, and Pan (2023) compare students' learning incentives under two variants of serial dictatorship: direct SD, where students submit their rank-order lists (ROLs) in advance, and sequential SD, where students choose colleges sequentially in priority order without initially submitting ROLs. They show that the sequential SD improves students' welfare by eliminating admission uncertainty. While our findings are consistent with theirs in showing welfare gain from reducing uncertainty, the underlying mechanisms differ. Their model assumes that colleges are grouped into distinct "tiers," with all students strictly preferring any college in a higher tier and colleges within a tier being ex-ante symmetric. In this setting, the sequential SD makes students acquire *less* information by forcing them to focus on the best available tier at the timing of learning, thereby avoiding wasteful information acquisition. In contrast, our model allows for ex-ante asymmetry across colleges, and students' preferences can be reversed depending on learning outcomes. As a result, in our benchmark model, students are encouraged to acquire *more* information, which improves the quality of student-college matches and leads to higher overall welfare.

Our paper also contributes to a broader literature of search and matching with incomplete information in the context of college admissions.[6] Immorlica, Leshno, Lo, and Lucier (2020) model students' information acquisition as a sequential search problem and introduce "regret-free stability" under costly information acquisition.[7] While they establish the existence of such outcomes, they abstract from the detailed process of individual students' information acquisition. Our work complements theirs by providing a full equilibrium characterization under specific admission mechanisms. Grenet, He, and Kübler (2022) provide empirical evidence from Germany's university admissions system, showing that students are more likely to accept early offers because holding (multiple) early offers prompts them to invest more time in learning about universities. This aligns with our results that students are more likely to acquire information when they are more likely to be admitted by both colleges. Chade, Lewis, and Smith (2014) study students' application strategies when application is costly and colleges observe noisy signals of student ability. In their model, students choose which colleges to apply to, and admission probabilities are determined endogenously based on the

[6]See Chade, Eeckhout, and Smith (2017) for a recent survey of the literature.

[7]They define an outcome to be stable if no student can form a blocking pair with a college or would wish to collect more information, and *regret-free stable* if it is stable and each student has acquired information optimally.

application decisions of all students. While their framework shares with ours the feature that students' strategic decisions affect their admission outcomes, their focus is on application portfolio choices—that is, how students choose the set of colleges to apply—whereas our focus is on the timing and content of learning decisions. Importantly, we incorporate incomplete information on both the student and college sides, which enables us to conduct a novel comparative statics on the perfectness of exam scores within an experimental setting.

Our notion of the perfectness of scores is related to whether admission decisions are made by an imperfect measure of students' abilities. A few recent papers investigate how noisy exam scores as a single measure of students' abilities affect matching outcomes in different centralized mechanisms. Lien, Zheng, and Zhong (2017) compare the Boston mechanism and serial dictatorship and investigate how these mechanisms achieve ex-ante fairness when an admission decision is made based on exam scores from which students' true ability may not be perfectly revealed. Lien, Zheng, and Zhong (2016) bring this comparison to the laboratory, highlighting the importance of timing of preference submission (pre-exam vs. post-exam) created by different mechanisms. Pan (2019) provides evidence from the field and laboratory that pre-exam preference submission in the Boston mechanism cannot fully fix the issue created by a single exam's measurement error.

Our findings on ranking and attendance mistakes in the experiment align with patterns of student behavior under DA documented in recent studies. Regarding ranking mistakes, Chen and Sönmez (2006) find that about 36% of participants misrepresent their preferences in a laboratory setting. More recent studies—Artemov, Che, and He (2021), Hassidima, Romm, and Shorrer (2021), and Shorrer and Sóvágó (2023)—report that 17% to 35% of applicants misrepresent their preferences in college admissions using DA in Australia, Israel, and Hungary, respectively. Attendance mistakes are also related with Narita (2018), who documents that about 7% of NYC high school applicants do not pursue their assigned schools but enter a secondary matching process, often due to newly acquired information about school characteristics or a revised understanding of their own preferences. This parallels our finding that attendance mistakes are closely associated with suboptimal learning.

The remainder of this paper is organized as follows. Section 2 provides our theoretical model and analyzes students' learning behavior and its welfare implications. Section 3 describes the experimental design and presents a set of testable hypotheses, and Section 4 reports our experimental findings. Section 5 concludes. Proofs are relegated to Appendix A. Appendix B presents the extension of our model to three colleges. Additional figures and experimental instructions are provided in Appendices C and D.

2 Theoretical Analysis

In this section, we develop a theoretical model to analyze students' learning behavior under DA. We then introduce a benchmark admission system and examine the welfare implications of admission uncertainty.

2.1 Model

There are two colleges, 1 and 2, and a unit mass of students. Each college $i = 1, 2$ has capacity $k < \frac{1}{2}$ and quality q_i, with $\Delta := q_1 - q_2 \geq 0$ commonly known. Each student is characterized by $(\alpha, \epsilon_1, \epsilon_2)$, where α is intrinsic ability drawn from a distribution F on $[\underline{\alpha}, \overline{\alpha}]$, and ϵ_i is her idiosyncratic preference or "suitability" of college i. The ϵ_i's are independent of each other and of α. A student's value from college i is $v_i = q_i + \epsilon_i$. That is, the student enjoys an extra payoff ϵ_i, in addition to q_i, regardless of her ability.

Although each student's ability is private information, the ϵ_i's are unknown to the student a priori. To learn each ϵ_i, she must pay cost c, with learning occurring sequentially: she first chooses whether to learn one ϵ_i, observes its realization, and then decides whether to learn the other. If learned, each ϵ_i is drawn independently from a distribution G on $[-\delta, \delta]$, where G is continuous, strictly increasing, and symmetric around zero; that is, $G(\epsilon) = 1 - G(-\epsilon)$. Thus, the expected value of college i is $q_i + \epsilon_i$ if learned, and q_i otherwise. We assume $q_2 > \delta$ so attending either college is strictly better than not attending, and $\Delta < 2\delta$ to avoid the trivial case that all students prefer college 1 regardless of the realizations of ϵ_1 and ϵ_2.

Colleges evaluate students based on their scores, which imperfectly reflect their intrinsic ability. Specifically, the score of a student with ability α is given by $s = r\,\alpha + (1 - r)\,\theta$, where $r \in [0, 1)$ is a constant, and $\theta \sim U[-\eta, \eta]$ is a noise term. Students do not observe their own scores, while colleges observe scores but not abilities. Thus, students face uncertainty about how they are evaluated, though higher-ability students are more likely to have higher scores. In this sense, we interpret r as the degree of perfectness of the scores.

We consider centralized admissions using DA in which students submit rank-order lists (ROLs) to a clearinghouse that simulates the following procedure. In the first round, students apply to their top choice, and colleges tentatively accept applicants with the highest scores up to capacity, and reject the rest permanently. In each subsequent round, rejected students apply to their next choice, and colleges re-evaluate all currently admitted students and new applicants based on the scores, again tentatively accepting top students and rejecting the rest. This process continues until there are no more rejections.

The timing is given as follows. At $t = 0$, students choose whether to learn ϵ_1 and/or ϵ_2 or neither. At $t = 1$, students simultaneously submit ROLs. At $t = 2$, the clearinghouse finalizes the assignments. Finally, at $t = 3$, students decide whether to enroll.

2.2 Equilibrium learning behavior

We begin with two remarks. First, since DA makes truthful reporting a weakly dominant strategy for students (Dubins and Freedman, 1981; Roth, 1982), students rank $i \succ j$ if and only if the expected value of college i exceeds that of college j.[8] Second, with a continuum of students, the DA outcome is characterized by cutoff scores $(\hat{s}_1, \hat{s}_2)$ (Azevedo and Leshno, 2016). More precisely, suppose $\hat{s}_1 > \hat{s}_2$ (which will be verified later) and consider a student who submits $1 \succ 2$. If $s \geq \hat{s}_1$, she is accepted by college 1 in the first round and remains there. If $s \in [\hat{s}_2, \hat{s}_1)$, she is rejected by college 1 in the first round but accepted by college 2 in the second round. If $s < \hat{s}_2$, she is rejected by both. Similarly, a student who submits $2 \succ 1$ is accepted by college 2 and is retained whenever $s \geq \hat{s}_2$. Hence, students with $s \geq \hat{s}_1$ are assigned to the college they rank higher, while those with $s \in [\hat{s}_2, \hat{s}_1)$ are assigned to college 2 regardless of their rank orders.

The following theorem characterizes equilibrium learning and ROL submission decisions.

Theorem 1. *There exists a unique equilibrium in which $\hat{s}_1 > \hat{s}_2$ and students' learning and ROL submission decisions are as follows: for each α, there exist $\bar{c}(\alpha)$ and $\bar{\epsilon}(\alpha)$ such that*

(i) for $c \geq \bar{c}(\alpha)$, students do not learn the suitability and submit $1 \succ 2$.

(ii) for $c < \bar{c}(\alpha)$, students learn ϵ_i for some i, and then they submit

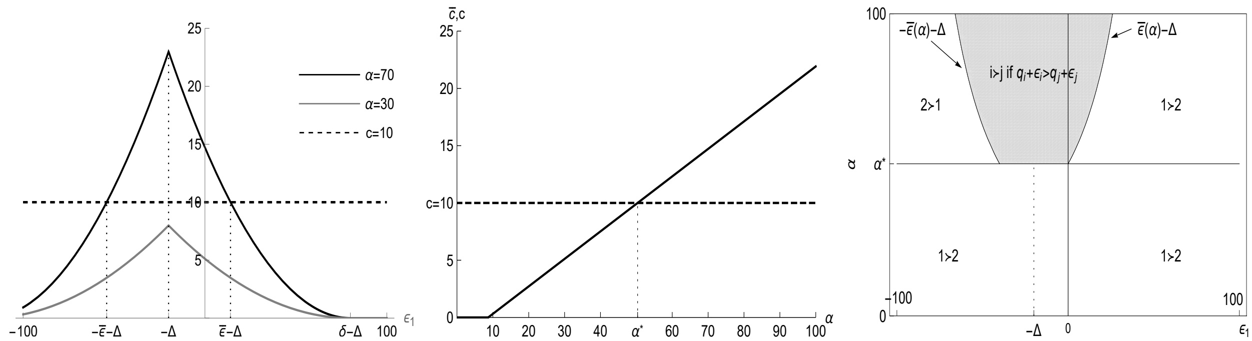
- *(a) $i \succ j$ without learning ϵ_j if $\epsilon_i \geq \bar{\epsilon}(\alpha) + (i - j)\Delta$;*
- *(b) $i \succ (\prec) j$ when $q_i + \epsilon_i > (<) q_j + \epsilon_j$ after learning ϵ_j additionally if $|\epsilon_i + (i - j)\Delta| < \bar{\epsilon}(\alpha)$;*
- *(c) $j \succ i$ without learning ϵ_j if $\epsilon_i \leq -\bar{\epsilon}(\alpha) + (i - j)\Delta$.*

To explain students' learning behavior, we first consider a student with ability α who has learned ϵ_1 and analyze her decision on whether to subsequently learn ϵ_2. We then examine her initial decision to learn ϵ_1. Given ϵ_1, if the student chooses not to learn ϵ_2, then the expected value of colleges are $\mathbb{E}[v_1|\epsilon_1] = q_1 + \epsilon_1$ and $\mathbb{E}[v_2|\epsilon_1] = q_2$. Her expected payoff is

$$u(\epsilon_1; \alpha) = Q_2(\alpha) q_2 + Q_1(\alpha)\big(\Delta + \epsilon_1\big)\mathbb{1}_{\{\epsilon_1 > -\Delta\}},$$

where $Q_i(\alpha) := \text{Prob}(s \geq \hat{s}_i | \alpha)$ is the probability that her score exceeds $\hat{s}_i$, and $\mathbb{1}_{\{\epsilon_1 | \epsilon_1 > -\Delta\}}$ is an indicator function taking 1 if $\epsilon_1 > -\Delta$ and 0 otherwise. Note that if $\epsilon_1 > -\Delta$, the student submits $1 \succ 2$ and is assigned to college 1 with probability $Q_1(\alpha)$, giving the payoff $q_1 + \epsilon_1$,

[8] We assume that if colleges are indifferent to a student, she randomly chooses one to rank higher. Such indifference arises if (i) she learns only ϵ_i and finds $q_i + \epsilon_i = q_j$, or (ii) she learns both ϵ_1 and ϵ_2 and finds $q_1 + \epsilon_1 = q_2 + \epsilon_2$. Both occur with probability zero, so we do not explicitly consider them in what follows.



(a) Learning ϵ_2 after observing ϵ_1 (b) Learning decision for ϵ_1 (c) ROL submissions

Figure 1: **Learning and ROL submission decisions in DA.** Parameters: $\alpha \sim U[0,100]$, $\theta \sim U[-100,100]$, $\epsilon_i \sim U[-100,100]$, $\Delta = 20$, $k = r = 0.4$, $c = 10$, and $(\hat{s}_1, \hat{s}_2) \simeq (45.21, 30.98)$. In panel (a), solid lines represent $u(\epsilon_2|\epsilon_1;\alpha) - u(\epsilon_1;\alpha)$. In panel (b), $\overline{c}(\alpha) = \max\{0.24\alpha - 2.08, 0\}$ is represented by the solid line, and $\alpha^* \simeq 50.35$. In both panels, the dashed line represents c. In panel (c), students in the gray region are those with α who have learned $\epsilon_1 \in (-\overline{\epsilon}(\alpha) - \Delta, \overline{\epsilon}(\alpha) + \Delta)$.

and to college 2 with probability $Q_2(\alpha) - Q_1(\alpha)$, yielding q_2. If $\epsilon_1 < -\Delta$, she submits $2 \succ 1$ and is assigned to college 2 with probability $Q_2(\alpha)$, receiving q_2. Next, suppose that the student also chooses to learn ϵ_2. She then ranks $i \succ j$ if $q_i + \epsilon_i > q_j + \epsilon_j$, and after some algebra, her expected payoff can be written as

$$u(\epsilon_2|\epsilon_1;\alpha) = Q_2(\alpha)q_2 + Q_1(\alpha)\int_{-\delta}^{\epsilon_1+\Delta}(\Delta + \epsilon_1 - \epsilon_2)\,dG(\epsilon_2).$$

Note that the difference $u(\epsilon_2|\epsilon_1;\alpha) - u(\epsilon_1;\alpha)$ captures the expected gain from learning ϵ_2. So, the student chooses to learn it if and only if $u(\epsilon_2|\epsilon_1;\alpha) - u(\epsilon_1;\alpha) > c$, which is equivalent to $\epsilon_1 \in (-\overline{\epsilon}(\alpha) - \Delta, \overline{\epsilon}(\alpha) - \Delta)$ for some threshold $\overline{\epsilon}(\alpha)$. Panel (a) of Figure 1 illustrates this relationship across different ability levels.[9] For a given α, the gain from learning is small when ϵ_1 is either high enough to prefer college 1 or low enough to prefer college 2, even without learning ϵ_2. Accordingly, she does not learn ϵ_2 and submits $1 \succ 2$ if $\epsilon_1 \geq \overline{\epsilon}(\alpha) - \Delta$, or $2 \succ 1$ if $\epsilon_1 \leq -\overline{\epsilon}(\alpha) - \Delta$. For intermediate values, she learns ϵ_2 and ranks the college with the higher value. The figure also shows that the gain from learning increases with the ability α. This is because students with a higher α have a higher chance of admission to college 1, so, the learning of additional information is more valuable. As a result, the threshold $\overline{\epsilon}(\alpha)$ increases with α.

We now turn to students' initial learning decisions—whether to learn ϵ_1. Let

$$V_1(\alpha) := \max\{U(\epsilon_1;\alpha) - c, U(\epsilon_1,\epsilon_2;\alpha) - 2c\}$$

[9]In the figure, $(\hat{s}_1, \hat{s}_2)$, $\overline{c}(\alpha)$, and α^* are equilibrium values determined under the given parameters.

denote the ex-ante expected payoff for a student with α from learning ϵ_1. The first term in the square bracket is the expected payoff from learning only ϵ_1, where $U(\epsilon_1;\alpha) := \mathbb{E}[u(\epsilon_1;\alpha)]$, and the second term is that from learning both ϵ_1 and ϵ_2, where $U(\epsilon_1,\epsilon_2;\alpha) := \mathbb{E}[u(\epsilon_2|\epsilon_1;\alpha)]$. A student chooses to learn ϵ_1 if and only if $V_1(\alpha)$ exceeds the ex-ante expected payoff from submitting $1 \succ 2$ without learning, $V_0(\alpha) := Q_2(\alpha)q_2 + Q_1(\alpha)\Delta$. As shown in panel (b) of Figure 1, there exists a $\bar{c}(\alpha)$ such that $V_1(\alpha) > V_0(\alpha)$ if and only if $c < \bar{c}(\alpha)$, or equivalently, $\alpha > \alpha^*$ where α^* satisfies $\bar{c}(\alpha^*) = c$. Thus, students with $\alpha \leq \alpha^*$ do not learn ϵ_1, while those with $\alpha > \alpha^*$ do—following the learning and ROL submission strategies described earlier. Panel (c) of Figure 1 summarizes these behavior.

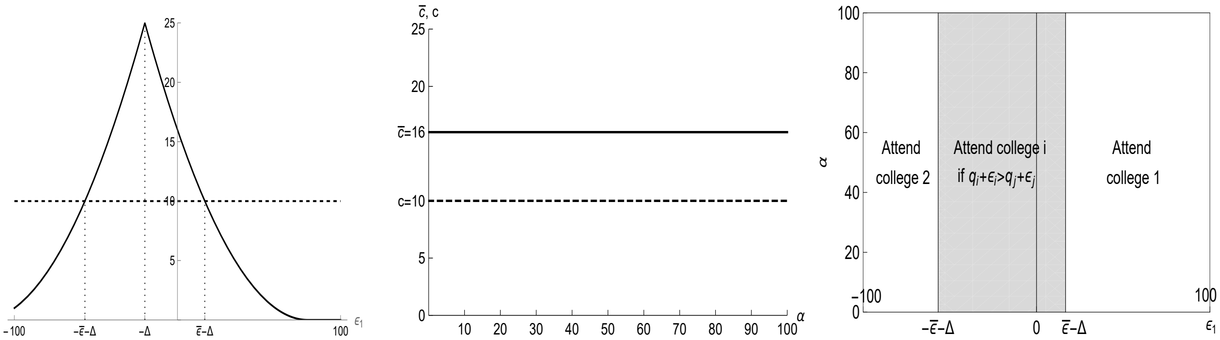
The analysis for the case where students learn ϵ_2 followed by ϵ_1 mirrors the previous case. A student learns ϵ_1 after observing ϵ_2 if $\epsilon_2 \in (-\bar{\epsilon}(\alpha) + \Delta, \bar{\epsilon}(\alpha) + \Delta)$, and learns ϵ_2 initially if $V_2(\alpha) > V_0(\alpha)$, where $V_2(\alpha)$ is defined similar to $V_1(\alpha)$. Importantly, in Appendix A.1, we show that $V_1(\alpha) = V_2(\alpha)$ for each α, meaning that the order of learning between ϵ_1 and ϵ_2 does not matter. To understand this, note that when students learn only one ϵ_i, learning ϵ_1 can change the student's ranking if $\epsilon_1 < -\Delta$ (i.e, $q_1 + \epsilon_1 < q_2$), and learning ϵ_2 can change the ranking if $\epsilon_2 > \Delta$ (i.e., $q_1 < q_2 + \epsilon_2$). By the symmetry of G,

$$\mathrm{Prob}(\epsilon_1 < -\Delta) = G(-\Delta) = 1 - G(\Delta) = \mathrm{Prob}(\epsilon_2 > \Delta),$$

so learning either ϵ_1 or ϵ_2 provides the same information about the colleges' expected values. Similarly, for a given ϵ_i, additional learning of ϵ_j matters only when it alters the rankings. However, the order of learning does not matter again, since

$$\epsilon_1 \gtrless \epsilon_2 - \Delta \Leftrightarrow q_1 + \epsilon_1 \gtrless q_2 + \epsilon_2 \Leftrightarrow \epsilon_2 \lessgtr \epsilon_1 + \Delta.$$

Remark 1. The irrelevance of learning order does not generally extend to the case with more than two colleges. To illustrate, consider three colleges with $q_1 > q_2 > q_3$ and $q_1 - q_2 = q_2 - q_3 \equiv \Delta$. Suppose that the learning cost is high enough, so students can afford to learn only one ϵ_i for $i = 1, 2, 3$. In this case, it is optimal to learn ϵ_2. Intuitively, learning ϵ_2 allows a student's rank order of colleges to be $1 \succ 2 \succ 3$ (if $q_1 > q_2 + \epsilon_2 > q_3$), $2 \succ 1 \succ 3$ (if $q_2 + \epsilon_2 > q_1 > q_3$), or $1 \succ 3 \succ 2$ (if $q_1 > q_3 > q_2 + \epsilon_2$). In contrast, learning ϵ_1 yields the rank orders $1 \succ 2 \succ 3$ ($q_1 + \epsilon_1 > q_2 > q_3$), $2 \succ 1 \succ 3$ ($q_2 > q_1 + \epsilon_1 > q_3$), or $2 \succ 3 \succ 1$ ($q_2 > q_3 > q_1 + \epsilon_1$). While the first two rank orders arise with the same probability under both learning strategies, the third differs: the condition for $1 \succ 3 \succ 2$ (i.e., $\epsilon_2 < -\Delta$) is more likely than that for $2 \succ 3 \succ 1$ (i.e., $\epsilon_1 < -2\Delta$). Thus, by learning ϵ_2, students are more likely to revise their rank orders, and so obtain a higher expected payoff, than learning ϵ_1 (or likely ϵ_3). When students can learn multiple ϵ_i's with a low cost, this suggests a "sequential search" problem with an endogenous order,



(a) Learning ϵ_2 after observing ϵ_1 (b) Learning decision for ϵ_1 (c) Enrollment decision

Figure 2: **Learning and enrollment decisions in the benchmark.** The parameters are the same as those in Figure 1. In equilibrium, $\bar{c} = 16$, $\bar{\epsilon} \simeq 36.57$, and $(\hat{s}_1, \hat{s}_2) \simeq (40, 30.98)$. The solid lines in panels (a) and (b) represent $u(\epsilon_2|\epsilon_1) - u(\epsilon_2)$ and $\bar{c}$, respectively, and the dashed line represents c. In panel (c), the gray region represents students who have learned $\epsilon_1 \in (-\bar{\epsilon} - \Delta, \bar{\epsilon} - \Delta)$.

which complicates the analysis and is beyond the scope of this paper.

Nevertheless, there is a case that allows for a tractable extension with more than two colleges—when all colleges are ex-ante symmetric, i.e., $q_1 = q_2 = q_3 \equiv q$. In this case, the order of learning does not matter, and the learning decision resembles the baseline two-college setting: a student with α chooses to learn any ϵ_i if $c < \bar{c}(\alpha)$ and proceeds to learn ϵ_j if $|\epsilon_i|$ is below a threshold and so on. A key difference from the two-college case is that the expected gain from learning the second ϵ_j, after observing ϵ_i, depends on the sign of ϵ_i. If $\epsilon_i > 0$, the student compares colleges i and j, as college i's value, $q + \epsilon_i$, exceeds that of college k, q. If $\epsilon_i \leq 0$, she instead compares j and k, since college i now offers less than q. Aside from this, the structure of learning remains similar. See Appendix B for a formal analysis.

2.3 Benchmark admission system

We now consider a benchmark admission system in which students make learning decisions *after* admission. Specifically, we consider the following decentralized admissions: at $t = 1$, students apply to any college without cost; at $t = 2$, colleges admit students based on scores; and at $t = 3$, students decide where to enroll, if at all.

In this admission system, it is a weakly dominant strategy for students to apply to both colleges and to delay learning suitability after being admitted by both colleges but before making an enrollment decision at $t = 2.5$. Since students' payoffs from attending colleges are independent of their abilities, this makes their learning decisions independent of abilities. Apart from this, the analysis is analogous to that of DA. Specifically, since $\hat{s}_1 > \hat{s}_2$, students with $s \geq \hat{s}_1$ are admitted by both colleges, those with $s \in [\hat{s}_2, \hat{s}_1)$ are admitted only to college

2, and the rest are rejected. Thus, only the first group of students may benefit from learning, while students in the second group enroll in college 2 without learning. For the first group, learning and enrollment decisions follow the same logic of DA, except that $Q_1(\alpha) = Q_2(\alpha) = 1$ since they are already admitted to both colleges. Figure 2 illustrates these decisions using the same parameters as in Figure 1, assuming that they learn ϵ_1 first. Note that $\overline{c}$ and $\overline{\epsilon}$ in Figure 2 correspond to $\overline{c}(\alpha)$ and $\overline{\epsilon}(\alpha)$ in DA evaluated at $Q_1(\alpha) = 1$. Therefore, we have $\overline{c}(\alpha) \leq \overline{c}$ and $\overline{\epsilon}(\alpha) \leq \overline{\epsilon}$ for all α.

Remark 2. Given that real-life decentralized admissions often involve instability and congestion, one may consider an alternative benchmark such as "sequential serial dictatorship," studied by Hakimov, Kübler, and Pan (2023), in which students sequentially select universities based on their priority order without submitting ROLs. This alternative yields the same matching outcome as our benchmark system because students can choose whether to learn suitability at the point when it is their turn to make a decision. Another relevant alternative is a "real-time" college-proposing DA, in which students are asked in each round of college admissions to accept or reject a college's offer without submitting ROLs upfront. The DA mechanism used in Victory, Australia, is of this kind. In that mechanism, applicants initially submit ROLs before scores are known but can revise them after scores are released (Artemov, Che, and He, 2021). In our setting, such a system would induce learning and ROL revision only by those with scores higher than $\hat{s}_1$, yielding the same outcome the benchmark.

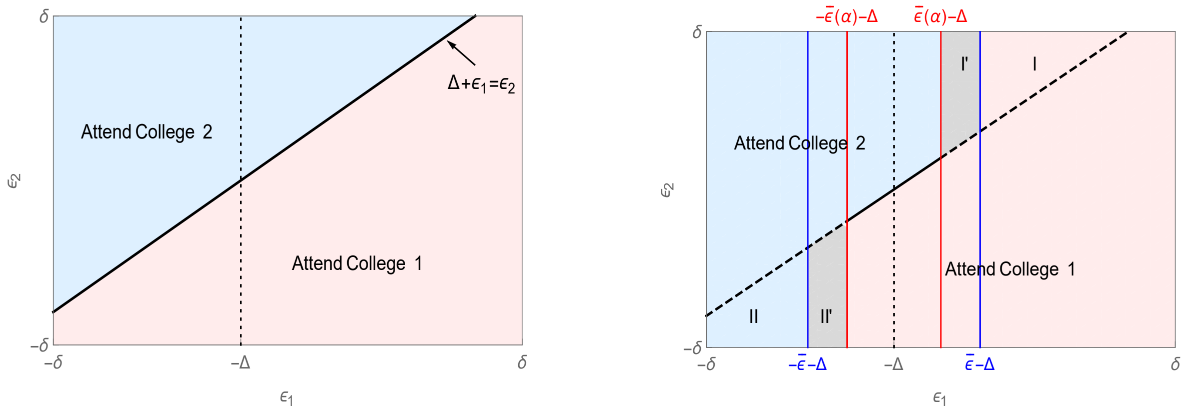
2.4 Welfare

As shown before, students in DA face admission uncertainty, whereas those in the benchmark do not. In this section, we investigate how such uncertainty affects students' learning incentives and their welfare. Formally, define social welfare as follows:

$$SW := (MV_1 + MV_2) - c\,m_L,$$

where MV_i is the aggregate expected value of students attending college i, evaluated using the information they have, and m_L is the mass of students who learn at least one ϵ_i.

Assume that students, if they choose to learn, begin by learning ϵ_1 under both DA and the benchmark. They then proceed to learn ϵ_2 if $\epsilon_1 \in (-\overline{\epsilon}(\alpha) - \Delta, \overline{\epsilon}(\alpha) - \Delta)$ under DA and $\epsilon_1 \in (-\overline{\epsilon} - \Delta, \overline{\epsilon} - \Delta)$ under the benchmark. Since $\overline{\epsilon}(\alpha) \leq \overline{\epsilon}$, it follows that $(-\overline{\epsilon}(\alpha) - \Delta, \overline{\epsilon}(\alpha) - \Delta) \subseteq (-\overline{\epsilon} - \Delta, \overline{\epsilon} - \Delta)$, showing that students are less likely to learn additional information under DA than under the benchmark, conditional on having learned ϵ_1. That is, DA yields a lower *intensive learning margin.* To understand its welfare effect, see Figure 3 that illustrates



(a) First-best allocation

(b) Allocation for a given α

Figure 3: Allocations for those with $s \geq \hat{s}_1$.

allocations of students with scores above $\hat{s}_1$. In the left panel, the optimal allocation under full information assigns students with $\epsilon_1 > \epsilon_2 - \Delta$ to college 1, and those with $\epsilon_1 < \epsilon_2 - \Delta$ to college 2. The right panel illustrates allocations under the benchmark and DA. In the benchmark, students only learn ϵ_2 if $\epsilon_1 \in (-\bar{\epsilon} - \Delta, \bar{\epsilon} - \Delta)$, so those in region I (resp., II) attend college 1 (resp., 2), even though they would have preferred college 2 (resp., 1) if they had learned ϵ_2. This shows an efficiency loss. In DA, since $\bar{\epsilon}(\alpha) \leq \bar{\epsilon}$, students in additional regions I′ and II′ also fail to learn ϵ_2 and are mis-assigned. Thus, DA incurs more welfare loss than the benchmark unless the exam score perfectly reflects students' abilities (i.e., $r = 1$).

Turning to the *extensive learning margin*, recall that $\bar{c}(\alpha) \leq \bar{c}$ for all α and $\bar{c}(\alpha^*) = c$. For any $c < \bar{c}$, students with $\alpha > \alpha^*$ in DA and those admitted by both colleges in the benchmark learn at least one ϵ_i. It is easy to see that α^* increases with c, implying that fewer students choose to acquire any information as c rises, whereas it remains constant in the benchmark as long as $c < \bar{c}$. This suggests that when c is high, DA results in both lower intensity and lower participation in learning, leading to lower welfare. When c is low, DA may feature a larger extensive margin, potentially raising welfare. However, in this case, students with lower abilities begin to learn, but they are less likely to be assigned to their preferred colleges. Consequently, their learning decisions have less influence on students with high abilities, who are more likely to be assigned to their desired colleges and so have a greater impact on welfare.

Let superscripts D and B denote the equilibrium values under DA and the benchmark, respectively. The observations above yield the following results:

Theorem 2. *$SW^D < SW^B$ for any $r \in [0, 1)$, while $SW^D = SW^B$ at $r = 1$. Moreover, SW^B is invariant in r for any $\Delta \geq 0$, and SW^D increases with r whenever $\Delta = 0$.*

The welfare comparison of SW^D across different values of r is complicated. When $r = 1$, no students face admission uncertainty, and only those with sufficiently high abilities acquire information. As r decreases, however, high-ability students become more likely to have low scores, reducing their chance of admission to preferred colleges and thereby lowering the expected gain from learning. Conversely, low-ability students become more likely to receive high scores, increasing their incentives to learn. This asymmetry complicates theoretical predictions about the welfare effects of admission uncertainty.

A tractable case arises when $\Delta = 0$. In this case, we have $\hat{s}_1 = \hat{s}_2 \equiv \hat{s}$, so $Q_1(\alpha) = Q_2(\alpha) \equiv Q(\alpha)$. In the proof of Lemma A6 in Appendix A.2, we show that

$$\frac{dQ(\alpha)}{dr} = \frac{\alpha - \mathbb{E}[\alpha]}{2\eta(1-r)^2}.$$

This highlights the aforementioned asymmetry: as r decreases, students above the average ability face a lower admission probability, while those below the average face a higher one. We further show that the sign of $\frac{dSW^D}{dr}$ coincide with that of

$$\int_{\alpha^*}^{\overline{\alpha}} \frac{dQ(\alpha)}{dr} dF(\alpha) = \frac{1 - F[\alpha^*]}{2\eta(1-r)^2}\big(\mathbb{E}[\alpha|\alpha \geq \alpha^*] - \mathbb{E}[\alpha]\big) \geq 0,$$

which implies that SW^D increases with r. Intuitively, changes in r affect $Q(\alpha)$ directly and also influence α^* (or $\overline{c}(\alpha)$) and $\overline{\epsilon}(\alpha)$ indirectly through $Q(\alpha)$. However, for small changes in r, the indirect effects are of second-order: the marginal type α^* is indifferent between learning and not learning, and the marginal type who learns $\overline{\epsilon}(\alpha)$ is indifferent between learning one ϵ_i or two. As a result, the welfare change mainly comes from the direct effect of r on $Q(\alpha)$, leading to an increase in SW^D.

Although it is analytically intractable to derive such a result when $\Delta > 0$, a similar intuition would hold. Therefore, we experimentally consider an environment in which the monotonicity is preserved and empirically investigate the impact of the perfectness of exam scores using our experimental data.[10]

Remark 3. While Theorem 2 highlights the informational inefficiency of DA, it does not imply that DA always yields lower welfare than decentralized admissions. In our model with a continuum of students and no aggregate uncertainty, DA loses its advantages over the benchmark. With a finite number of students and endogenous effort choices, students may play mixed strategies in applications in decentralized admissions (Hafalir et al., 2018), and DA eliminates colleges' enrollment uncertainty when there is aggregate uncertainty about

[10]We have numerically verified that under the parameters used in Figures 1 and 2 as well as the experiments, SW^D increases in $r \in [0, 1]$.

students' preferences (Che and Koh, 2016). Our model abstracts from these factors to focus on students' information acquisition. Nonetheless, our welfare comparison generalizes existing results. When $r = 0$, DA reduces to random serial dictatorship (RSD), where scores are randomly drawn, and when $r = 1$, it becomes serial dictatorship (SD), where scores fully reflect abilities. Theorem 2 thus confirms that RSD yields lower welfare than SD, consistent with Theorem 5 of Artemov (2021).

3 Experimental Design and Hypotheses

3.1 Experimental design and procedure

Our focus is to study how different timings of information acquisition lead to different social welfare depending on the degree of perfectness of the scores. To this end, we implement our experiment that features a 2×2 treatment design as presented in Table 1. The first treatment variable concerns the *perfectness* of the scores captured by $r \in [0, 1]$. $r = 1$ refers to the case of full perfectness, in which the single determinant of admissions is α, which is known to each student. $r = 0$ refers to the case of full imperfectness, in which the single determinant of admissions is θ, which is completely unknown to each student. We choose $r = 0.9$ and $r = 0.6$ for our treatment design.[11] The second treatment variable concerns whether the admission system is DA or the benchmark (BA). Treatments DH and DL refer to the DA system with high and low perfectness, respectively. Treatments BH and BL refer to the BA system with high and low perfectness, respectively. Treatments DH and DL are collectively called DA treatments, and treatments BH and BL are jointly called BA treatments. Treatments DH and BH are collectively called high-perfectness treatments, and DL and BL are collectively called low-perfectness treatments. The parameters and distributions chosen for our experiments are as follows: $\alpha \sim U[0, 100]$, $\theta \sim U[0, 100]$, $\epsilon_i \sim U[-100, 100]$, $c = 10$, $k = 0.4$, $\Delta = q_1 - q_2 = 170 - 150 = 20$, consistent with Figures 1 and 2, where the supports of the distributions are discretized to involve integer values only.

Our experiment was conducted using oTree (Chen, Schonger, and Wickens, 2016) at the HKUST via Zoom with the real-time online mode. Three sessions were conducted for each treatment. A total of 190 subjects were recruited from the graduate and undergraduate population of the university.[12] When invited, subjects were instructed to find a quiet place to stay for the entire duration of the experiment and join the designated Zoom meeting

[11] The choice of r for our experimental design is guided by the fact that students in DA have no incentives to acquire information if r is too small (below 0.3 in our experimental environment). We thus chose a sufficiently large r to ensure that learning occurs in equilibrium.

[12] The number of participants was 49, 49, 47, and 45 for Treatments DH, BH, DL, and BL, respectively.

Table 1: Experimental treatments

		Score Perfectness	
		High ($r = 0.9$)	Low ($r = 0.6$)
Admission System	Deferred Acceptance (DA)	DH	DL
	Benchmark Admission (BA)	BH	BL

using their own laptop or desktop computer.[13] Turning on their video for the entire course of the experiment was a strict requirement and chatting among subjects was prohibited by the Zoom settings. Each received an electronic copy of the experimental instructions via the chat message in Zoom. To ensure that the information contained in the instructions was public knowledge, the instructions were read aloud via Zoom. We used a between-subject design.

We illustrate the instructions for Treatment DH. The full experimental instructions for Treatment DH and Treatment BH are available in Appendices D.1 and D.2, respectively. There were two colleges, College 1 and College 2. The colleges were simple mechanical admission functions that admitted students as follows. Upon receiving an application, each college admitted a student based on her exam scores (E) and interview scores (I) as well as an exogenously given admission cutoff. E and I were randomly and independently drawn according to the uniform distribution over $\{0, 1, 2, ..., 99, 100\}$ and correspond to α and θ for each student, respectively. Then, the exam score was announced to the student privately while the total score was sent to every college she applied to without being revealed to her.

Note that participants in our experiments only played the role of students while the colleges were not strategic players making a deliberate choice. Admission cutoffs are exogenously given by the theoretical predictions with capacity $k = 0.4$ for each college based on the model presented in the previous section with a continuum of students. It is as if an individual subject in our experiment cannot influence the admission decisions of the colleges and thus take the admission cutoffs as given. By doing so, we abstract away the colleges' strategic decisions. This approach allows us to focus on investigating students' learning decisions and their impact on the welfare generated by each admission system. Without colleges' strategic decisions, the remaining problem becomes a single-person decision problem for each student.

After the exam score (E) was revealed to each subject, she was asked to indicate her top choice between College 1 and College 2. Then the admission procedure began as follows:

1. The admission office sent the application to the college of her top choice.

[13] We recommended they not use their mobile phone or tablet PC to join the experiment due to the potential concern of the presentation quality of the oTree game platform and of unexpected technical issues.

2. The college accepted her application if her total score $T = 0.9 \times E + 0.1 \times I$ was higher than its own admission cutoff given as follows:

	DH	BH	DL	BL
$\hat{s}_1$	36.3	35	45.21	40
$\hat{s}_2$	23		30.98	

3. If her application was accepted by the college of her top choice, the admission process was finalized. Otherwise, the admission office sent her application to the college of her second choice that decided whether to accept her application based on her total score T and the admission cutoff.
4. If her application was accepted by the college of her second choice, the admission process was finalized. Otherwise, she was not admitted by any college, and the process was finalized.

In case she received admission, she was asked to decide whether to pursue a college or not.

The gain a subject obtained from a college depended on how well the college suited her, corresponding to the value of college $q_i + \epsilon_i$ for each $i = 1, 2$. The gain (in tokens) from College 1, denoted by G_1, was randomly and independently chosen from {70, 71, 72, ..., 269, 270}, while each integer in the interval was equally likely. The gain from College 2, denoted by G_2, was randomly and independently chosen from {50, 51, 52, ..., 249, 250}, while each integer in the interval was equally likely. The gain became part of the earnings if and only if a college admitted the subject and the subject decided to pursue it.[14] Otherwise, a subject received the default gain of 50 tokens.

G_1 and G_2 were unknown to a subject at the beginning of each round. Once each round began, the decision screen for each subject contained two panels (left and right): one panel for the application decision and the other panel for the learning decision (see Figures D1 and D2 in Appendix D.1 for the screenshots). The placement (left or right) of the two panels was uniformly randomly chosen for each subject in each round. The learning panel allowed subjects to learn what the exact gain from College 1 (i.e., the value of G_1) was. If the subject decided to learn it, she needed to pay 10 tokens. Then, the subject further decided whether

[14] While we assume that attending college is strictly more beneficial than not attending, our experimental design allowed subjects to choose whether or not to attend a college after being admitted. Some may argue that this design choice was unnecessary. However, research conducted by Narita (2018) using NYC high school matching data suggests that a significant proportion of students may not pursue their immediately available option due to various psychological reasons. Artemov, Che, and He (2021) and Shorrer and Sóvágó (2023) also document that a non-negligible fraction of Australian and Hungarian college applicants adopt unambiguously dominated strategies in strategically straightforward situations. Therefore, allowing subjects to choose whether or not to attend college may provide valuable insights into decision-making processes and the factors that influence them.

to learn what the exact gain from College 2 (i.e., the value of G_2) was.[15] If she decided to learn it, the subject needed to pay an additional 10 tokens. Note that the two panels were always presented side-by-side, and each panel ran independently from the other panel. Thus, it was entirely up to each subject 1) *whether* to learn none/one/both of G_1 and G_2 and 2) *when* to learn them. Subjects could learn none/one/both before or after they were admitted by a college or colleges. The learning cost did not depend on the timing of learning.

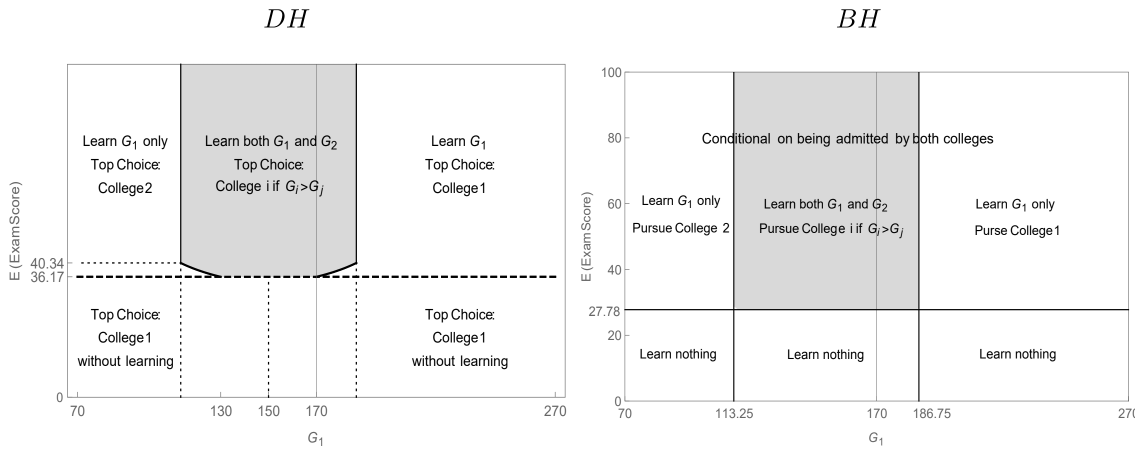
The earnings from each round were the gain from admission minus the total cost of learning paid if a college admitted a subject and the subject decided to pursue it. Otherwise, it was the default 50 tokens minus the total cost of learning. At the end of each round, we provided feedback to each subject on her 1) exam score, 2) interview score, 3) total score, 4) which college(s) admitted her, 5) which college she pursued, 6) learning decisions, 7) G_1 and G_2 regardless of whether she paid to learn none/one/both of them, and 8) the earnings from the round. For the payment, one round out of the 30 rounds was randomly chosen. Including an HKD 40 show-up fee, subjects received, on average, HKD 190 ($\approx$ USD 25). All payments were made electronically via the autopay system of HKUST to the bank account an individual participant provided to the Student Information System (SIS). Each session lasted approximately 1 hour on average.

3.2 Experimental hypotheses

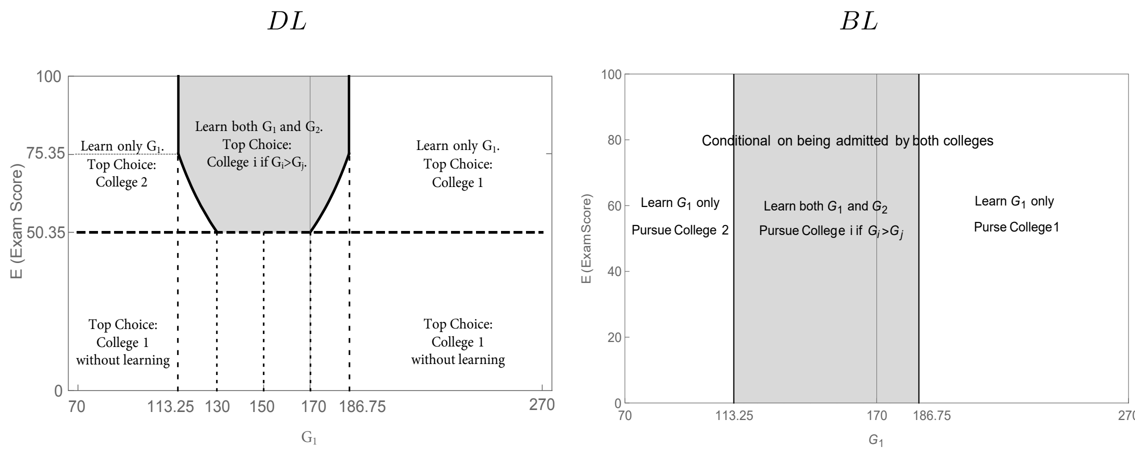
Figure 4 describes the outcomes from the theoretical predictions for the high- and low-perfectness treatments, respectively, where the benchmark cases in both figures consider only those who are admitted by both colleges. The learning decisions, the top choice college in the case of DA and which college to pursue in the case of BA depend on two variables, G_1 (gains from College 1) presented on the horizontal axis and E (exam scores) presented on the vertical axis.

The key difference of DA relative to the benchmark admission system is the timing of learning. In the DA environment, students must learn before submitting their top-choice college, and there is no reason to learn further afterward. The BA offers different incentives to students. On one hand, according to the weakly undominated strategy equilibrium, students have incentives to learn only after they are admitted by both colleges. On the other hand, if the exam score is above 38.9 in BH and 66.7 in BL, students know that both colleges will admit them, regardless of the interview scores, so the timing of learning does not matter. Our first hypothesis summarizes this result.

[15] We fixed the order of learning because it is not our primary objective to test the order-neutrality prediction.



(a) High perfectness treatments



(b) Low perfectness treatments

Figure 4: Outcome comparison

Hypothesis 1 (**Timing of Learning $\mathbf{G_1}$**). *a) In Treatments DH and DL, students learn only before they submit their top choice college. b) In Treatments BH and BL, students whose exam scores are below 38.9 in BH and 66.7 in BL learn only after they are admitted by both colleges.*

The fact that students learn only after they are admitted by both colleges in the BA treatments implies that their first learning decision on G_1 must be independent of the exam score, as long as their exam scores are below 38.9 in BH and 66.7 in BL, which is demonstrated by the right panel of each of Figures 4(a) and 4(b). The area below E = 27.78 labeled as "Learn nothing" in the right panel of Figure 4(a) indicates that students whose exam score is below 27.78 cannot be admitted by both colleges regardless of their interview scores. However, in the DA treatments, the same learning decisions are crucially dependent upon the exam scores, as illustrated by the left panel of each of the two figures. We thus

Table 2: Social cost of pre-application learning

	MV_1	MV_2	TC	SW
DH	81.55	61.48	8.70	138.33
BH	82.30	65.85	9.12	139.03

Social cost of pre-application learning: 0.7

	MV_1	MV_2	TC	SW
DL	77.79	63.64	6.60	134.83
BL	82.30	65.85	9.12	139.03

Social cost of pre-application learning: 4.2

have our second hypothesis, as follows:

Hypothesis 2 (**$\boldsymbol{E}$-Dependence of $\mathbf{G_1}$ Learning**). *a) In treatments DH and DL, students learn G_1 only if the exam scores are above 36.17 in DH and 50.35 in DL. b) In treatments BH and BL, conditional on being admitted by both colleges, whether students learn G_1 does not depend on their exam scores.*

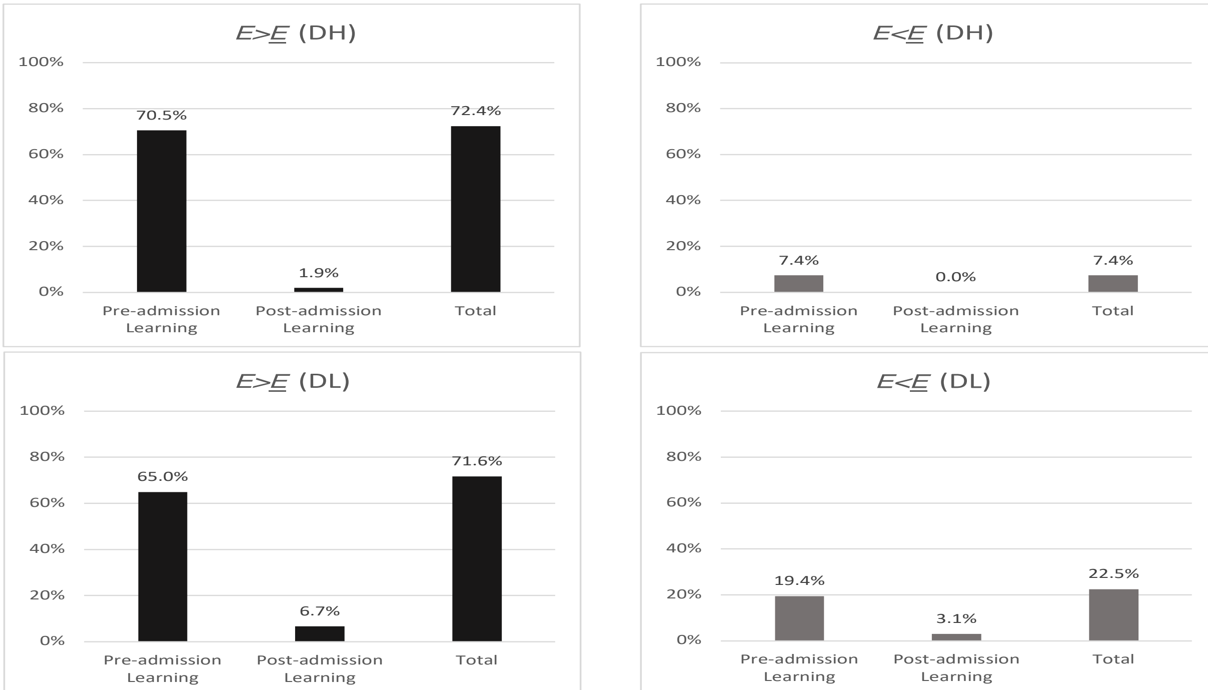
We now shift our attention to the G_2 learning decisions. *Among those who already learned* G_1, whether they *further* learn G_2 depends on the realization of G_1 in a specific way. As illustrated by Figure 4, the decisions are dependent upon whether the realized value of G_1 is in the range $[113.25, 186.75]$ in all four treatment conditions.[16] Thus, we have the following hypothesis.

Hypothesis 3 (**$\mathbf{G_2}$ Learning**). *Among the students who already learned G_1, the proportion of students who further learn G_2 is substantially higher when the realized value of G_1 is in* $[113.25, 186.75]$ *in each treatment.*

Our last hypothesis is about social welfare, our key prediction. Table 2 provides theoretical values of social welfare for each treatment. It illustrates that social welfare in DA is determined by the different learning decisions guided by different degrees of perfectness. As a result, the **social cost of pre-application learning**, defined as the difference in social welfare between the two admission environments, is small (0.7) under the high perfectness while that becomes substantially larger (4.2) under the low perfectness. This result is summarized by our next hypothesis.

Hypothesis 4 (**Social Cost**). *The difference in the average social welfare between treatments DL and BL is substantially larger than that between treatments DH and BH.*

[16]Precisely speaking, this statement is not true because of the two triangular regions below the U-shaped gray areas in the left panels of Figures 4(a) and 4(b). However, we do not specify any testable hypothesis regarding those regions because it is unlikely for us to have sufficient observations that belong to those (small) regions in our data.



■ E refers to the exam score (that corresponds to ability α in the theoretical model). $\underline{E}$ = 36.17 in DH and $\underline{E}$ = 50.35 in DL.

Figure 5: Learning G_1 in treatments DH and DL

4 Experimental Results

We conduct our primary analysis using data aggregated over the last 20 rounds *for each individual*. All qualitative results are robust to the use of data from all 30 rounds or from the last 10 rounds. We begin in Section 4.1 by analyzing the G_1 learning decisions and then move to analyze the G_2 learning behavior. Section 4.2 presents the welfare analysis. In both G_1 and G_2 learning decisions, we identify non-equilibrium decisions, motivating us to have Section 4.3 that is devoted to investigating non-equilibrium decision-making and the welfare consequences. Appendix C presents four scatter diagrams (Figures C1 and C2) that correspond to the theoretical counterparts presented in Figure 4, providing a general picture of the learning decisions observed in the laboratory, and several additional histograms for learning decisions.

4.1 Learning of $\mathbf{G_1}$ and $\mathbf{G_2}$

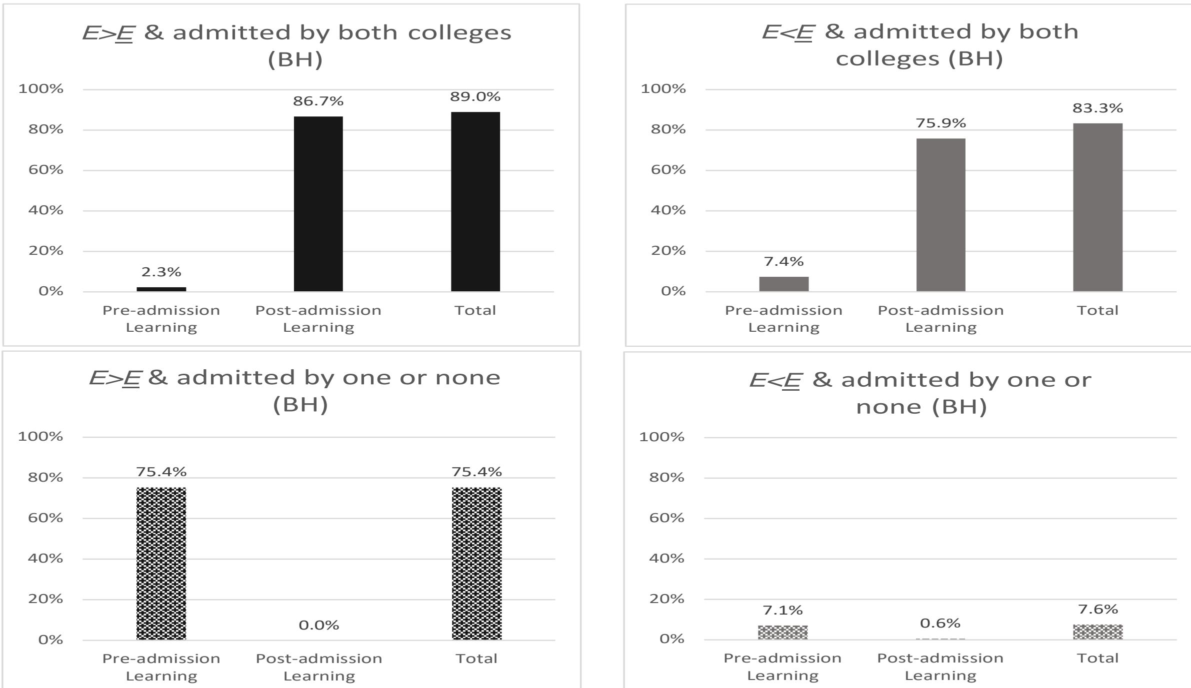
Figure 5 presents the G_1 learning (and timing of learning) decisions of students in treatments DH (two top panels) and DL (two bottom panels). When reporting the results, we divide observations into two categories. The left panels present the learning decisions made by students whose exam scores were above 36.17 in DH and 50.35 in DL: students in this category have incentives to learn before submitting their top choice college. The right panels present the learning decision made by the remaining students: theory predicts that they will not choose to learn anything at any time. To provide a more comprehensive understanding of the subjects' learning decisions, Figure C3 in Appendix C presents two histograms with a bin size of 10, separately for DH and DL. These histograms reaffirm our qualitative conclusion by showing that our findings are not contingent on the binary categories used in our main figures.

Three observations emerge. First, under-learnings are observed in both DH and DL. The proportion of students who learned G_1 in the first category is slightly above 70% but below the theoretical prediction of 100%. Second, non-negligible proportions of those who have no incentives to learn decided to learn as reported in the right panels of Figure 5. In its magnitude, this observed *over-learning* is not as large as that of the *under-learning*. Third, if they learned, students almost always learned before submitting their top choice colleges. The last observation allows us to confirm Hypothesis 1(a).

Result 1 (Timing of learning $\mathbf{G_1}$ in DH and DL). *In treatments DH and DL, the vast majority of students learned G_1 before submitting their top choice colleges.*

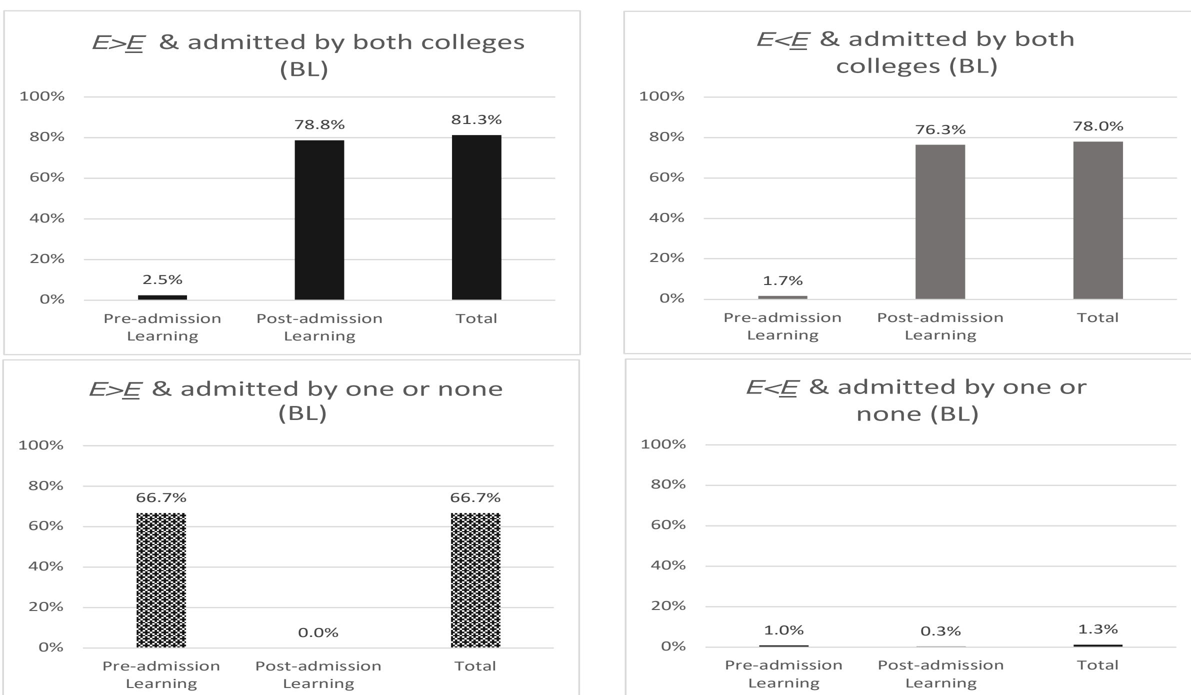
Figure 6 presents G_1 learning (and timing of learning) decisions of students, respectively, in treatments BH and BL. Observations are divided into four categories. In each treatment, the top-left panel presents the learning decisions made by students whose exam scores were above 38.9 in BH and 66.7 in BL such that both colleges admitted them (regardless of the interview scores). The top-right panel presents the learning decisions made by those who had exam scores below 38.9 in BH and 66.7 in BL but were admitted by both colleges ex-post after the interview scores were realized. The bottom-left panel presents the learning decisions made by students whose exam scores were above 38.9 in BH and 66.7 in BL, so they were supposed to be admitted by both colleges, which did not happen due to the fact that they did not apply to both colleges. The bottom-right panel presents the learning decisions made by those who did not get admitted by both colleges with their exam scores below 38.9 in BH and 66.7 in BL. Figures C4 and C5 reported in Appendix C present two histograms each for the learning G_1 decision based on E with the bin size of 10 as well as whether being admitted by both colleges or not. These histograms provide further support for our qualitative conclusion, demonstrating that our findings remain consistent regardless of the binary categories employed in our primary figures.

A few observations are immediately clear in Figure 6. First, similar to the DA treatments, we observe under-learnings (relative to the equilibrium learning) in both BH and BL. Except



■ E refers to the exam score and $\underline{E}$ = 38.9.

(a) Treatment BH



■ E refers to the exam score E and $\underline{E}$ = 66.7.

(b) Treatment BL

Figure 6: Learning of G_1 in treatments BH and BL

for the fourth category, with exam scores below the cutoffs and without multiple admissions, students were supposed to learn G_1 100% of the time. The observed frequencies of learning G_1 are below 100%. Second, the left-bottom panels of Figures 6(a) and 6(b) report that students who received exam scores above 38.9 in BH and 66.7 in BL but did not apply to both colleges always learned G_1 before admission. Knowing that they would be admitted by both colleges, they learned which college suits them better even before admission and applied only to the better one. This behavior is optimal, even though we did not specify it in our theoretical analysis focusing on the weakly dominant strategies. Third, the two upper panels of Figures 6(a) and 6(b) indicate that vast majorities of students who applied to and were admitted by both colleges learned G_1 after they were admitted by both colleges. Last, students with exam scores below 38.9 in BH and 66.7 in BL who did not have multiple admissions rarely learned G_1. The last two observations allow us to confirm Hypothesis 1(b) as follows:

Result 2 (**Timing of learning $\mathbf{G_1}$ in BH and BL**). *The vast majority of students who applied to and were admitted by both colleges learned G_1 after being admitted by both colleges in both BH and BL. Without multiple admissions, the vast majority of students whose exam scores were below 38.9 in BH and 66.7 in BL did not learn G_1.*

Regarding Hypothesis 2(a), the positive proportions (7.4% and 22.5%) of learning observed in treatments DH and DL (the two right panels of Figure 5) from the students with exam scores below the cutoff values are not overwhelmingly large. Overall, the outcome is qualitatively consistent with the theoretical prediction because the vast majority of students who learned G_1 are those who had exam scores above the cutoffs. Comparison of the proportions of learning G_1 between students with exam scores above and below $\underline{E} = 38.9$ in treatment BH (89% vs. 83.3% on average) presented in the two top panels of Figure 6(a) enables us to confirm the first part of Hypothesis 2(b). These two values are not statistically different from each other (two-sided Wilcoxon test, p-value = 0.787). The same conclusion is drawn if we compare the proportions in treatment BL (81.3% vs. 78%) presented in the two upper panels of Figure 6(b). Again, these two values are not statistically different from each other (two-sided Wilcoxon test, p-value = 0.780). The following result summarizes these findings.

Result 3 (***E*-dependence of $\mathbf{G_1}$ learning**). *In treatments DH and DL, the vast majority of students who learned G_1 were those with exam scores above 36.17 in DH and 50.35 in DL. In treatments BH and BL, the G_1 learning decisions made by the students who were admitted by both colleges did not depend on their exam scores.*

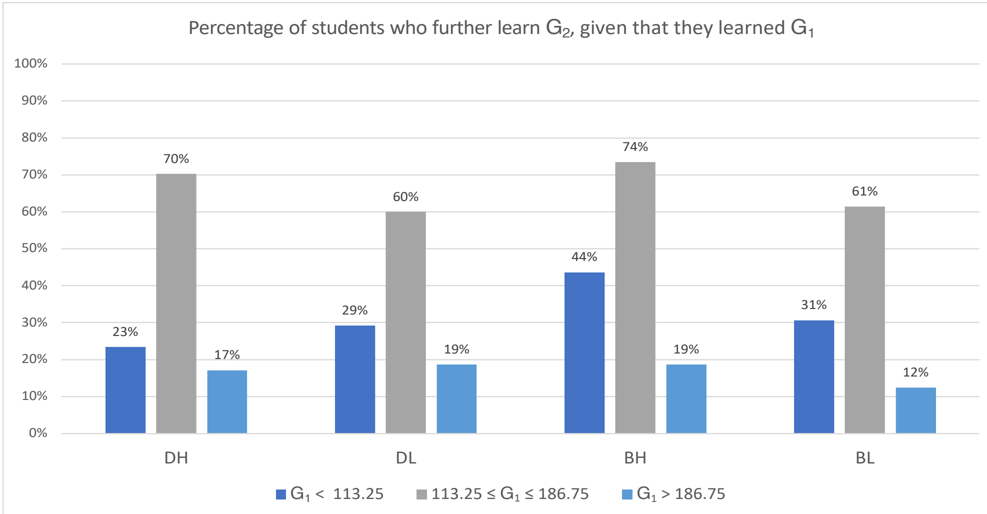


Figure 7: Learning G_2

Figure 7 presents the percentage of students who further learned G_2 given that they already learned G_1. Theory suggests that students who already learned G_1 have an incentive to learn G_2 only if G_1 is in the range of $[113.25, 186.75]$. The one-sided Wilcoxon test reveals that the percentage of students who further learned G_2 is significantly higher (p-values < 0.0001 for all four treatments) when G_1 is in $[113.25, 186.75]$ than when it is not. This observation allows us to confirm Hypothesis 3. However, suboptimal learnings, both under-learning (i.e., the gray bars are below 100% in Figure 7) and over-learning (i.e., the dark and light blue bars are above 0% in Figure 7) are observed across all treatments. Both types of suboptimal learning are substantial in magnitude. This observation motivates us to look into the suboptimal decisions in more detail in Section 4.3. In Appendix C, two sets of histograms, labeled as Figure C6 and Figure C7, present a total of four histograms. These histograms showcase the learning decisions regarding G_2, categorized based on the realization of G_2, and are displayed with a bin size of 10 for each of the four treatments. These additional histograms contribute to the reinforcement of our qualitative conclusion, as they demonstrate the consistency of our findings irrespective of the binary categories used in our primary figures.

Result 4 ($\mathbf{G_2}$ learning). *Among the students who learned G_1, the proportion of students who further learned G_2 was substantially higher when the realized value of G_1 was in $[113.25, 186.75]$ in each treatment. However, a substantial degree of suboptimal learning was observed in all treatments.*

Table 3: Theoretical Social Welfare

DH	MV_1	MV_2	TC	SW
Ex-ante	81.55	61.48	8.70	138.33
Ex-post	82.02	65.86	8.61	**139.27**
$\mathcal{R}$				0.94

DL	MV_1	MV_2	TC	SW
Ex-ante	77.79	63.64	6.60	134.83
Ex-post	77.47	65.79	6.30	**136.96**
$\mathcal{R}$				2.13

BH	MV_1	MV_2	TC	SW
Ex-ante	82.30	65.85	9.12	139.03
Ex-post	81.23	65.65	8.98	**137.90**
$\mathcal{R}$				-1.13

BL	MV_1	MV_2	TC	SW
Ex-ante	82.30	65.85	9.12	139.03
Ex-post	77.50	69.02	9.28	**137.25**
$\mathcal{R}$				-1.78

■ The ex-post social welfare values are calculated based on the optimal strategy of the player but by taking the realizations of G_1 and G_2 from our data (instead of the uniform prior).
■ When $G_1 = G_2$ and a student is admitted by both colleges, we assume that the student attends college 1 so that the corresponding matching value goes to college 1. There was one such tie case each in DH and BH.

Table 4: Empirical Social Welfare

	MV_1	MV_2	TC	SW
DH	81.56	61.32	6.70	136.18
BH	85.73	57.14	8.31	134.56

	MV_1	MV_2	TC	SW
DL	69.56	68.22	6.22	131.56
BL	75.34	65.66	7.08	133.92

■ The empirical social welfare values are calculated by adding the realized values of MV_1 and MV_2 across all subjects and then subtracting the total learning cost C paid.

4.2 Welfare analysis

Table 3 presents the ex-ante social welfare values that theory predicts based on the uniform prior of G_1 and G_2 (i.e., those presented in Table 2), as well as the ex-post social welfare values (presented in boldface) calculated based on the realizations of G_1 and G_2 according to our experimental data.[17] Given the large number of observations we have in our data, the ex-post social welfare values are reasonably close to the ex-ante values, and most of the ordinal welfare rankings are preserved, except that the social welfare value in DH (139.27) is (marginally) larger than that in BH (137.90); apparently, the law of large number does not fully apply.

Table 4 presents the empirical social welfare values calculated using our data. Now we are ready to calculate the empirical social cost (SC) of pre-application learning as follows:

$$SC = \text{Adjusted Empirical Social Welfare in BA} - \text{Adjusted Empirical Social Welfare in DA},$$

[17]When calculating the ex-post social welfare values, we take the realization of G_1 and G_2 from our data and calculate the welfare based on the optimal strategy of the player.

where

$$\text{Adjusted Empirical Social Welfare} = \text{Empirical Social Welfare Value} - \mathcal{R}.$$

$\mathcal{R}$ =(Ex-post Social Welfare − Ex-ante Social Welfare) is a correction term to get rid of the effect of the different ex-post realizations of G_1 and G_2 across treatments. For example, the gap between the ex-post social welfare and the ex-ante social welfare in DH is $(139.27 - 138.33) = 0.94$ while that in BH is $(137.90 - 139.03) = -1.13$. Because both of them originate solely from the realizations of G_1 and G_2, we need to adjust the empirical social welfare by adding the correction term. Then the empirical social costs of pre-application learning for the high perfectness environment and the low perfectness environment are respectively

$$\begin{aligned} SC_H &= (134.56 + 1.13) - (136.18 - 0.94) = 0.45, \\ SC_L &= (133.92 + 1.78) - (131.56 - 2.13) = 6.27. \end{aligned}$$

Recall that the theoretical values for the social cost of pre-application learning provided in Table 2 are 0.7 and 4.2, respectively. The two-sided Mann-Whitney test confirms that the adjusted empirical social welfare in BH adjusted with the correction term is not different from the empirical social welfare in DH (p-value= 0.07314), implying that SC_H is not significantly different from 0. However, the same non-parametric test shows that the empirical social welfare in BL adjusted with the correction term is significantly different from the empirical social welfare in DL (p-value= 0.007593), implying that SC_L is significantly larger than 0. Another noticeable observation is that, in all treatments, the empirical social welfare values are strictly below the ex-post values presented in Table 3.[18] The empirical social welfare being strictly below the ex-post welfare is driven by the non-equilibrium learning decisions reported in the previous two subsections. This result also implies that the higher the perfectness of scores in DA the higher the empirical social welfare.

Result 5 (**Social Welfare**). *In all treatments, empirical social welfare values are strictly below the theoretical levels. The social cost of pre-application learning is significantly larger than zero in the low exam perfectness environment but that is not the case in the high exam perfectness environment.*

[18] The observed (individual-level) average welfare loss ranges between 3.09 and 5.40. These values are equivalent to 2.3%-7.4% of the empirical social welfare values and comparable to 40%-87% of the learning cost paid.

4.3 Non-equilibrium learning and welfare decomposition

Where does the observed discrepancy between the empirical social welfare values and the theoretical ones come from? Apparently, the non-equilibrium learning identified in the previous subsections must be responsible. In this section, we thus investigate non-equilibrium learning more carefully and quantify the welfare loss (relative to the theoretical level) caused by different types of non-equilibrium decisions.

Non-equilibrium learning can be either over-learning or under-learning, where the former (the latter) implies that a student acquired more (less) information than the optimal amount prescribed by the equilibrium. Depending on whether the excessive (missing) learning is on G_1 only, G_2 only, or both, we categorize the non-equilibrium learning as over-learning (under-learning) G_1 only, G_2 only, or both. For example, "over-learning G_1 only" covers the cases in which a student received an offer from one or no college but learned G_1 either in the pre-admission or in the post-admission stage.[19] For each observation classified as suboptimal learning, we calculate the welfare difference between the theoretical value (that the individual could have achieved if he/she were making the optimal learning decision) and the empirical value (coming from the suboptimal learning decision). The calculated welfare differences are aggregated for each category, and the results are reported in Figure 8.[20]

The decomposition of the welfare losses created by different types of suboptimal learning reported in Figure 8 leads to the following observations. First, in all treatments, under-learning was considerably more prevalent than over-learning. Second, among different types of under-learning, under-learning G_1 only was the greatest contributor to a welfare loss. Third, more suboptimal learning was observed in the low-transparency treatments (DL and BL) than in the high-transparency treatments (DH and BH). However, combining all welfare losses from suboptimal learning does not fully account for the observed welfare discrepancy presented in Table 4. For example, the welfare loss from all kinds of suboptimal learning in treatment DL was only 2.57 (= 0.37 + 0.19 − 0.13 + 1.50 + 0.20 + 0.44), which covers less than half of the total welfare loss (5.40). This observation implies that there must be other types of non-equilibrium decisions being made by our subjects.

[19] "Over-learning G_2 only" covers the cases in which 1) a student received offers from both colleges, learned that G_1 is either below 113.25 or above 186.75, but decided to learn G_2 further, and 2) a student learned that G_1 is either below 113.25 or above 186.75 but learned G_2 further then applied to only one college. "Over-learning both" covers the cases in which the total score is below the admission cutoff of College 1 but at any point both G_1 and G_2 are learned. Under-learning is categorized and defined in a consistent manner.

[20] Both over-learning and under-learning could generate positive welfare gain ex-post. For example, in treatment DL, when the exam score is strictly below but sufficiently close to 50.35, the optimal decision is to choose College 1 as the top choice without learning. However, one could make a suboptimal decision to learn both G_1 and G_2. If $G_2 > G_1 + 20$ (the total learning cost paid) then suboptimal learning allows the decision-maker to submit the top choice of College 2 and get admitted, leading to a positive welfare gain.

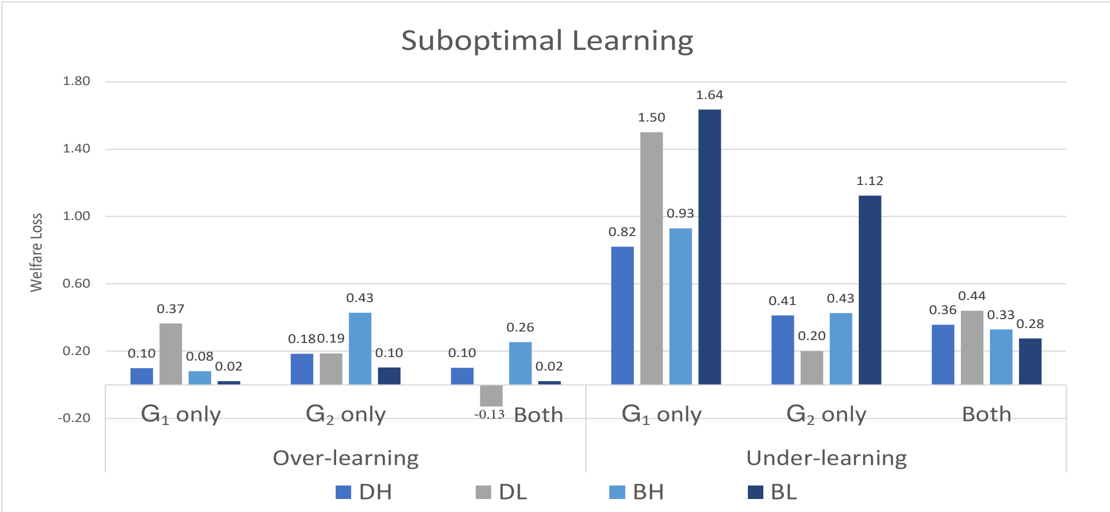


Figure 8: Decomposition of Welfare Loss from Suboptimal Learning

Figure 9 reports the result from the decomposition of the welfare losses created by different types of suboptimal decisions that include not only suboptimal learning but also other types of mistakes. Clearly, suboptimal learning is the greatest contributor to the observed welfare loss, but it alone does not fully account for the entire amount. Two other kinds of suboptimal decisions are made in the ranking reporting (in DH and DL) / application (in BH and BL) decisions and attendance decisions. For example, mistakes in the ranking reporting in treatments DH and DL cover cases in which a student learned both G_i and G_j with $G_i > G_j$ but submitted the top choice college as j and the cases in which a student submitted the top choice college as the ex-ante worse one without learning. The application mistakes in treatments BH and BL cover cases in which a student applied only to College 2 even though he/she was supposed to be admitted by College 1 ex-post if he/she applied to College 1. The attendance mistakes cover cases in which one or more offers were made but a student did not pursue any college. This result is consistent with the empirical findings from the literature including Artemov, Che, and He (2021), Rees-Jones (2018) and Shorrer and Sóvágó (2023) that a non-negligible proportion of applicants in various matching contexts adopted dominated choices.

Result 6 (Non-equilibrium Decisions). *In all treatments, substantial degrees of non-equilibrium decisions are observed. Overall, non-equilibrium learning decisions are the greatest contributor to welfare loss.*

Notably, non-equilibrium *attendance* decisions are responsible for a large proportion of welfare loss in the DA treatments while (almost) no such mistakes are made in the BA treatments. To understand why this occurred only in the DA treatments, we first take a

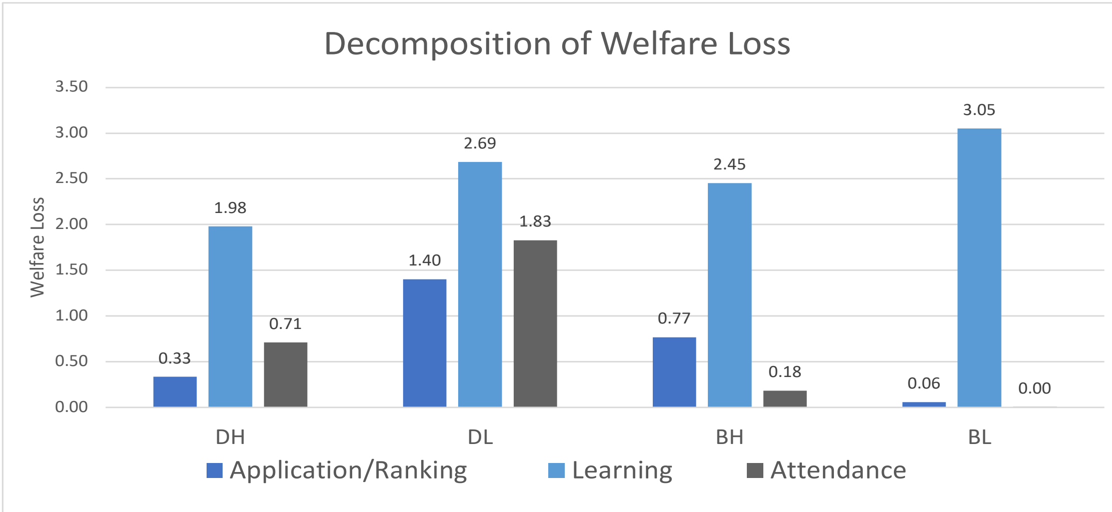


Figure 9: Decomposition of Welfare Loss from Non-equilibrium Decisions

closer look at all 11 observations with attendance mistakes in treatment DL. In all but one case, the subjects chose College 1 (the ex-ante better college) as their top choice college without learning anything. With only one exception, the *post-admission learning* about both colleges or the college admitting the student occurred before the final rejection decision was made. This post-admission learning made students realize that they were admitted either by the college with a lower (than the other one) realized gain (8 cases) or by the college with a higher realized gain but the realized gain itself was not large (95 in one case and 111 in the other case) relative to the default gain of 50. The overall picture of the 5 cases involving attendance mistakes observed in treatment DH is exactly the same. All these observations suggest that the rejection decisions (attendance mistakes) are *associated with suboptimal post-admission learning* and may result from the disappointment students had when they learned they were admitted by the college with a relatively lower realized gain.

In Treatment DL, we also had mistakes in the ranking choice (top choice college) as another important contributor to the welfare loss, as indicated by the dark blue bar in Figure 9. The same kind of ranking choice mistakes was observed in treatment DH, but their welfare consequences were smaller (1.40 vs. 0.33). The observed suboptimal behavior is driven mainly by *pessimism*.[21] In treatment DL, 37 cases out of 39 in total occurred when students had no incentive to learn at all because their exam scores were below the cutoff 50.35.[22] Without learning, students in all 37 cases pessimistically chose College 2,

[21] The mistakes in the ranking choice observed in the DA treatments cannot be regarded as a behavior to distort the system and take strategic advantage over other students (see, e.g., Rees-Jones, 2017) because there are no strategic interactions among students in our environment.

[22] In the other 2 cases, students learned both G_1 and G_2 before submitting the topic choice, but they submitted the college with a lower realized value.

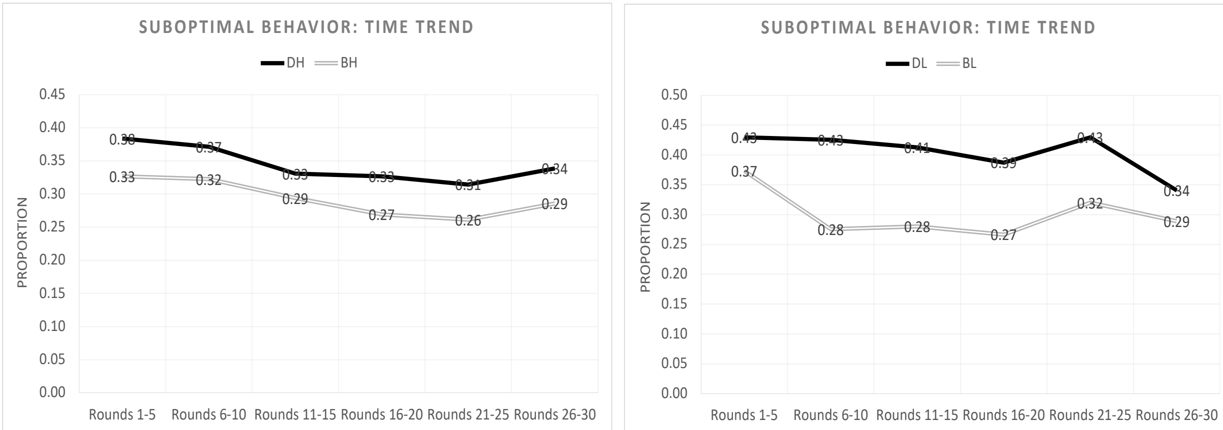


Figure 10: Welfare Loss - Time Trend

the ex-ante inferior college, as their top choice. However, the final total scores were above the admission cutoff for College, 1 so they would have been admitted to the ex-post better college if they had chosen College 1 as their top choice. In treatment DH, we had only 11 such cases, and the difference in their frequencies (39 in DL vs. 11 in DH) stemmed from the different degrees of perfectness of scores in these two treatments: students with low exam scores were more likely to be admitted by the college with a higher admission cutoff in treatment DL. As a result, pessimism led to a real mistake more often with lower perfectness.

Do people learn to make fewer mistakes over time? Figure 10 presents the proportions of suboptimal decision-making. While the trend decreases over time across all treatments, the decline is modest except in two cases: the first 10 rounds in BL and the last 10 rounds in DL. In all cases, the proportions of suboptimal behavior remain above 25%, suggesting that learning occurs only to a limited extent. It is also evident that more frequent suboptimal behaviors are observed in the DA treatments compared to the BA treatments, with a larger difference noted in the low exam perfectness environment. This result explains why the empirical social cost of pre-application learning in our data is not as substantial as the theoretical value in general, and is even insignificantly different from zero in the low exam perfectness environment.[23]

Why do people make non-equilibrium and thus suboptimal decisions? Although our experiment is not designed to address this question directly, we observe that subjects were more likely to make a suboptimal decision that led to a less substantial payoff loss. The average payoff losses from each suboptimal learning, ranking/application choice, and attendance de-

[23] Figure C8 in Appendix C presents the time trend of the welfare loss. It indicates that the time trend of the welfare loss is more volatile in the high exam perfectness environment than in the low exam perfectness environment.

cision are 8.9×10^{-3}, 79×10^{-3}, and 121×10^{-3}, while the corresponding treatment-average frequencies are 290, 13.75, and 4.5, respectively. This observation is in line with *payoff-dependent* mistakes, one of the most conventional ways to model mistake behavior in game theory, including Myerson (1978), Blume (1993), and McKelvey and Palfrey (1995).

5 Concluding Remarks

We theoretically and experimentally investigate the college admission system via DA when both colleges and students lack full information about the others. In our theoretical analysis, we characterized students learning and enrollment decisions. We identified the efficiency loss of DA induced by pre-application learning by comparing it with the benchmark system involving post-admission learning. Consistent with theory, we found in the laboratory that most students in DA acquired information only before submitting their top choice colleges, while those in the benchmark acquired information after being admitted by both colleges. We also observed substantial degrees of suboptimal learning as well as other suboptimal decisions that are responsible for the observed welfare loss.

A main contribution of our study is to empirically provide clear comparative statics results regarding how the imperfectness of exam scores influences students' incentives to acquire information. This finding carries important policy implications: when students lack comprehensive knowledge about the distinctive attributes of colleges, investing in the refinement of the examination system to diminish admission uncertainty becomes beneficial. Conversely, when students can effectively form their preferences with minimal information, a higher tolerance for imprecision in exam scores is permissible.

Our findings suggest that the degree of uncertainty students face during information acquisition should be an important consideration in the design of college admission system. From a designer's point of view, whether or not students acquire information with and without admission uncertainty may not be just a by-product of the choice of an admission system but rather a separate, independent choice variable. Although it is not straightforward for the designer to induce a particular degree of uncertainty, it is not outright impossible either. For instance, Hakimov, Kübler, and Pan (2023) show in their experimental setting that providing historical cutoff scores in the direct serial dictatorship improves students' welfare. Artemov (2021) proposes several policies, including disclosure of priorities, that improve welfare when students' information acquisition matters in the random serial dictatorship. Although it is beyond the scope of this paper to design a particular mechanism, our analyses suggest that it is important to understand how the uncertainty that students face is translated into the informational (dis)advantage of different admission mechanisms.

References

Artemov, Georgy. 2021. “Assignment mechanisms: Common preferences and information acquisition.” *Journal of Economic Theory* 198:105370.

Artemov, Georgy, Yeon-Koo Che, and Yinghua He. 2021. “Strategic Mistakes: Implications for Market Design Research.” Manuscript.

Azevedo, Eduardo M. and Jacob D. Leshno. 2016. “A Supply and Demand Framework for Two-Sided Matching Markets.” *Journal of Political Economy* 124:1235–1268.

Bade, Sophie. 2015. “Serial Dictatorship: The Unique Optimal Allocation Ruel When Information is Endogenous.” *Theoretical Economics* :385–410.

Blume, Lawrence E. 1993. “The statistical mechanics of strategic interaction.” *Games and Economic Behavior* 5:387–424.

Chade, Hector, Jan Eeckhout, and Lones Smith. 2017. “Sorting through Search and Matching Models in Economics.” *Journal of Economic Literature* 55:493–554.

Chade, Hector, Gregory Lewis, and Lones Smith. 2014. “Portfolios and the College Admissions Problem.” *Review of Economic Studies* 81:971–1002.

Che, Yeon-Koo and Youngwoo Koh. 2016. “Decentralized College Admissions.” *Journal of Political Economy* 124:1295–1338.

Chen, Daniel L, Martin Schonger, and Chris Wickens. 2016. “oTree–An open-source platform for laboratory, online, and field experiments.” *Journal of Behavioral and Experimental Finance* 9:88–97.

Chen, Yan and Yinghua He. 2021a. “Information acquisition and provision in school choice: a theoretical investigation.” *Economic Theory* :1–35.

———. 2021b. “Information acquisition and provision in school choice: an experimental study.” *Journal of Economic Theory* 197:105345.

Chen, Yan and Tayfun Sönmez. 2006. “School Choice: An Experimental Study.” *Journal of Economic Theory* 127:202–231.

Dubins, L. E. and D. A. Freedman. 1981. “Machiavelli and the Gale-Shapley Algorithm.” *American Mathematical Monthly* 88:485–494.

Gale, David and Lloyd Shapley. 1962. "College Admissions and the Stability of Marriage." *American Mathematical Monthly* 69:9–15.

Grenet, Julien, YingHua He, and Dorothea Kübler. 2022. "Preference Discovery in University Admissions: The Case for Dynamic Multi-offer Mechanisms." *Journal of Political Economy* 130:1427–1476.

Hafalir, Isa E, Rustamdjan Hakimov, Dorothea Kübler, and Morimitsu Kurino. 2018. "College admissions with entrance exams: Centralized versus decentralized." *Journal of Economic Theory* 176:886–934.

Hakimov, Rustamdjan, Dorothea Kübler, and Siqi Pan. 2023. "Costly Information Acquisition in Centralized Matching Markets." *Quantitative Economics* 14:1447–1490.

Hassidima, Avinatan, Assaf Romm, and Ran I. Shorrer. 2021. "The Limits of Incentives in Economic Matching Procedures." *Management Science* 67:951–963.

Immorlica, Nicole, Jacob Leshno, Irene Lo, and Brendan Lucier. 2020. "Information acquisition in matching markets: The role of price discovery." *Available at SSRN* .

Lien, Jaimie W, Jie Zheng, and Xiaohan Zhong. 2016. "Preference submission timing in school choice matching: testing fairness and efficiency in the laboratory." *Experimental Economics* 19:116–150.

———. 2017. "Ex-ante fairness in the Boston and serial dictatorship mechanisms under pre-exam and post-exam preference submission." *Games and Economic Behavior* 101:98–120.

McKelvey, Richard D and Thomas R Palfrey. 1995. "Quantal response equilibria for normal form games." *Games and Economic Behavior* 10:6–38.

Myerson, Roger B. 1978. "Refinements of the Nash equilibrium concept." *International Journal of Game Theory* 7:73–80.

Narita, Yusuke. 2018. "Match or mismatch? Learning and inertia in school choice." *Learning and Inertia in School Choice (June 18, 2018)* .

Pan, Siqi. 2019. "The instability of matching with overconfident agents." *Games and Economic Behavior* 113:396–415.

Rees-Jones, Alex. 2017. "Mistaken play in the deferred acceptance algorithm: Implications for positive assortative matching." *American Economic Review* 107:225–229.

———. 2018. “Suboptimal behavior in strategy-proof mechanisms: Evidence from the residency match.” *Games and Economic Behavior* 108:317–330.

Roth, Alvin E. 1982. “The Economics of Matching: Stability and Incentives.” *Mathematics of Operations Research* 7:617–628.

Shorrer, Ran I. and Sándor Sóvágó. 2023. “Dominated choices in a strategically simple college admissions environment.” *Journal of Political Economy Microeconomics* 1:781–807.

Online Appendix

In this appendix, we provide omitted proofs, discussions on more than two-college case, additional figures and tables, and sample instructions of the experiment.

A Omitted Proofs

A.1 Proof of Section 2.2

For given $\hat{s}_1 > \hat{s}_2$, we establish Lemmas A1 to A3 that characterize students' learning behaviors. We then show that $\hat{s}_1 > \hat{s}_2$ in equilibrium.

Lemma A1. *Suppose $\hat{s}_1 > \hat{s}_2$. For each α, the following results hold.*

(i) $u(\epsilon_j|\epsilon_i;\alpha) - u(\epsilon_j;\alpha) = 0$ if $\epsilon_1 \geq \delta - \Delta$ when $i = 1, j = 2$ or $\epsilon_2 \leq \Delta - \delta$ when $i = 2, j = 1$. Otherwise,

$$u(\epsilon_j|\epsilon_i;\alpha) - u(\epsilon_i) = Q_1(\alpha)\int_{|\epsilon_i+(j-i)\Delta|}^{\delta}(1 - G(\epsilon_j))d\epsilon_j,$$

which is strictly increasing (resp., decreasing) in ϵ_i for $\epsilon_i < (j-i)\Delta$ (resp., $\epsilon_i > (j-i)\Delta$). Moreover, $u(\epsilon_j|\epsilon_i;\alpha) - u(\epsilon_i;\alpha)$ is increasing in α.

(ii) For a given c, there exist $\bar{\epsilon}(\alpha)$ such that $u(\epsilon_j|\epsilon_i;\alpha) - u(\epsilon_j;\alpha) > c$ if $|\epsilon_i + (j-i)\Delta| < \bar{\epsilon}(\alpha)$, whenever $u(\epsilon_j|\epsilon_i;\alpha) - u(\epsilon_j;\alpha) > 0$. Moreover $\bar{\epsilon}(\alpha)$ is increasing in α.

Proof. (i). Suppose, first, that $i = 1$ and $j = 2$. Then, we have $u(\epsilon_1;\alpha) = Q_2(\alpha)q_2 + Q_1(\alpha)(\Delta + \epsilon_1)$ if $\epsilon_1 \geq -\Delta$ and $u(\epsilon_1;\alpha) = Q_2(\alpha)q_2$ if $\epsilon_1 < -\Delta$. Next, observe that

$$\begin{aligned}
u(\epsilon_2|\epsilon_1;\alpha) &= \int_{-\delta}^{\Delta+\epsilon_1}\left\{Q_1(\alpha)(q_1+\epsilon_1) + [(Q_2(\alpha) - Q_1(\alpha)](q_2+\epsilon_2)\right\}dG(\epsilon_2) \\
&\quad + \int_{\Delta+\epsilon_1}^{\delta} Q_2(\alpha)(q_2+\epsilon_2)dG(\epsilon_2) \\
&= Q_2(\alpha)q_2 + Q_1(\alpha)\int_{-\delta}^{\epsilon_1+\Delta}(\Delta+\epsilon_1-\epsilon_2)dG(\epsilon_j) \\
&= \begin{cases} Q_2(\alpha)q_2 + Q_1(\alpha)(\Delta+\epsilon_1) & \text{if } \epsilon_1 \geq \delta - \Delta, \\ Q_2(\alpha)q_2 + Q_1(\alpha)\int_{-(\Delta+\epsilon_1)}^{\delta}(1-G(\epsilon_2))d\epsilon_2 & \text{if } \epsilon_1 < \delta - \Delta, \end{cases}
\end{aligned} \tag{A.1}$$

where the last equality holds since for $\epsilon_1 \geq \delta - \Delta$,

$$\int_{-\delta}^{\epsilon_1+\Delta}(\Delta+\epsilon_1-\epsilon_2)dG(\epsilon_2) = \begin{cases} \int_{-\delta}^{\delta}(\Delta+\epsilon_1-\epsilon_2)dG(\epsilon_2) = \Delta+\epsilon_1 & \text{for } \epsilon_1 \geq \delta-\Delta, \\ \int_{-\delta}^{\epsilon_1+\Delta}G(\epsilon_2)d\epsilon_2 = \int_{-(\epsilon_1+\Delta)}^{\delta}(1-G(\epsilon_2))d\epsilon_2 & \text{for } \epsilon_1 < \delta-\Delta, \end{cases}$$

using the integration by parts and the symmetry of G for the case that $\epsilon_1 < \delta - \Delta$. Combining them together, we have that for $\Delta < \delta$,

$$u(\epsilon_2|\epsilon_1;\alpha) - u(\epsilon_1;\alpha) = \begin{cases} 0 & \text{if } \epsilon_1 \geq \delta - \Delta, \\ Q_1(\alpha)\left[-(\Delta+\epsilon_1) + \int_{-\epsilon_1+\Delta}^{\delta}(1-G(\epsilon_2))d\epsilon_2\right] & \text{if } \epsilon_1 \in [-\Delta, \delta-\Delta), \\ Q_1(\alpha)\int_{-(\Delta+\epsilon_1)}^{\delta}(1-G(\epsilon_2))d\epsilon_2 & \text{if } \epsilon_1 < -\Delta, \end{cases}$$

and for $\Delta \geq \delta$,

$$u(\epsilon_2|\epsilon_1;\alpha) - u(\epsilon_1;\alpha) = \begin{cases} 0 & \text{if } \epsilon_1 \geq \delta - \Delta, \\ Q_1(\alpha)\left[-(\Delta+\epsilon_1) + \int_{-(\epsilon_1+\Delta)}^{\delta}(1-G(\epsilon_2))d\epsilon_2\right] & \text{if } \epsilon_1 \in [-\delta, \delta-\Delta). \end{cases}$$

Thus, $u(\epsilon_2|\epsilon_1;\alpha) = u(\epsilon_1;\alpha)$ for $\epsilon_1 \geq \delta - \Delta$. Next, we establish the following result.

Claim A1. *For any* $\bar{\epsilon} \in [-\delta, \delta]$,

$$-\bar{\epsilon} + \int_{-\bar{\epsilon}}^{\delta}(1-G(\epsilon))d\epsilon = \int_{\bar{\epsilon}}^{\delta}(1-G(\epsilon))d\epsilon > 0.$$

Proof . For $\bar{\epsilon} \in [0, \delta]$, observe that

$$\begin{aligned} -\bar{\epsilon} + \int_{-\bar{\epsilon}}^{\delta}(1-G(\epsilon))d\epsilon &= -\bar{\epsilon} + \int_{-\bar{\epsilon}}^{0}(1-G(\epsilon))d\epsilon + \int_{0}^{\delta}(1-G(\epsilon))d\epsilon \\ &= -\int_{-\bar{\epsilon}}^{0}G(\epsilon)d\epsilon + \int_{0}^{\delta}(1-G(\epsilon))d\epsilon \\ &= -\int_{0}^{\bar{\epsilon}}(1-G(\epsilon))d\epsilon + \int_{0}^{\delta}(1-G(\epsilon))d\epsilon = \int_{\bar{\epsilon}}^{\delta}(1-G(\epsilon))d\epsilon > 0. \end{aligned}$$

The proof for the case with $\bar{\epsilon} \in [-\delta, 0)$ is similar, and we omit it. □

Note that for $\epsilon_1 < \delta - \Delta$, let $\bar{\epsilon} = \epsilon_1 + \Delta$. Then, the desired result follows from Claim A1. Next, consider the case that $i = 2$ and $j - 1$. Then, it is easy to see that $u(\epsilon_2;\alpha)$ is $Q_2(\alpha)(q_2+\epsilon_2)$ if $\epsilon_2 > \Delta$ and is $Q_2(\alpha)(q_2+\epsilon_2) + Q_1(\alpha)(\Delta - \epsilon_2)$ if $\epsilon_2 \leq \Delta$. We also have that

$$\begin{aligned} u(\epsilon_1|\epsilon_2;\alpha) &= Q_2(\alpha)(q_2+\epsilon_2) + Q_1(\alpha)\int_{\epsilon_2-\Delta}^{\delta}(\Delta+\epsilon_1-\epsilon_2)dG(\epsilon_1) \\ &= \begin{cases} Q_2(\alpha)(q_2+\epsilon_2) + Q_1(\alpha)\int_{\epsilon_2-\Delta}^{\delta}(1-G(\epsilon_1))d\epsilon_1 & \text{if } \epsilon_2 > \Delta - \delta, \\ Q_2(\alpha)(q_2+\epsilon_2) + Q_1(\alpha)(\Delta-\epsilon_2) & \text{if } \epsilon_2 \leq \Delta - \delta. \end{cases} \end{aligned} \tag{A.2}$$

The remaining proof is analogous, and so we omit it.

(ii). Suppose $u(\epsilon_j|\epsilon_i;\alpha) - u(\epsilon_j;\alpha) > 0$. Define $\bar{\epsilon}(\alpha)$ such that

$$Q_1(\alpha)\int_{\bar{\epsilon}}^{\delta}(1 - G(\epsilon))d\epsilon = c.$$

Note that since $Q_1(\alpha)$ is increasing in α, so is $\bar{\epsilon}(\alpha)$. Otherwise, the left-hand side becomes smaller than c. ■

Lemma A2. *Suppose $\hat{s}_1 > \hat{s}_2$. For each α, the following results hold.*

(i) $U(\epsilon_i;\alpha) = U(\epsilon_j;\alpha) \geq V_0(\alpha)$, where the inequality is strict if $\Delta < \delta$.

(ii) $U(\epsilon_i,\epsilon_j;\alpha) = U(\epsilon_j,\epsilon_i;\alpha) > U(\epsilon_i;\alpha)$.

Proof. (i). From (2.2),

$$U(\epsilon_1;\alpha) = \int_{-\delta}^{\delta} u(\epsilon_1;\alpha)dG(\epsilon_1) = \begin{cases} Q_2(\alpha)q_2 + Q_1(\alpha)\int_{-\delta}^{\delta}(\Delta + \epsilon_1)dG(\epsilon_1) & \text{if } \Delta \geq \delta, \\ Q_2(\alpha)q_2 + Q_1(\alpha)\int_{-\Delta}^{\delta}(\Delta + \epsilon_1)dG(\epsilon_1) & \text{if } \Delta < \delta, \end{cases}$$

and similarly,

$$U(\epsilon_2;\alpha) = \int_{-\delta}^{\delta} u(\epsilon_2;\alpha)dG(\epsilon_2) = \begin{cases} Q_2(\alpha)q_2 + Q_1(\alpha)\int_{-\delta}^{\delta}(\Delta - \epsilon_2)dG(\epsilon_2) & \text{if } \Delta \geq \delta, \\ Q_2(\alpha)q_2 + Q_1(\alpha)\int_{-\delta}^{\Delta}(\Delta - \epsilon_2)dG(\epsilon_2) & \text{if } \Delta < \delta. \end{cases}$$

Hence, $U(\epsilon_1;\alpha) = U(\epsilon_2;\alpha) = Q_2(\alpha)q_2 + Q_1(\alpha)\Delta = V_0(\alpha)$ for $\Delta \geq \delta$. Suppose $\Delta < \delta$. Note that

$$\int_{-\delta}^{\Delta}(\Delta - \epsilon)dG(\epsilon) = \int_{-\delta}^{\Delta} G(\epsilon)d\epsilon = \int_{-\delta}^{\Delta}(1 - G(-\epsilon))d\epsilon = \int_{-\Delta}^{\delta}(1 - G(\epsilon))d\epsilon = \int_{-\Delta}^{\delta}(\Delta + \epsilon)dG(\epsilon),$$

hence, we have

$$\begin{aligned} U(\epsilon_1;\alpha) = U(\epsilon_2;\alpha) &= Q_2(\alpha)q_2 + Q_1(\alpha)\int_{-\Delta}^{\delta}(1 - G(\epsilon))d\epsilon \\ = Q_2(\alpha)q_2 + Q_1(\alpha)\left[\Delta + \int_{\Delta}^{\delta}(1 - G(\epsilon))d\epsilon\right] &= V_0(\alpha) + Q_1(\alpha)\int_{\Delta}^{\delta}(1 - G(\epsilon))d\epsilon. \end{aligned} \tag{A.3}$$

(ii). Observe that $U(\epsilon_1,\epsilon_2;\alpha) = \int_{-\delta}^{\delta} u(\epsilon_2|\epsilon_1;\alpha)dG(\epsilon_1)$ and so from (A.1),

$$U(\epsilon_1,\epsilon_2;\alpha) = Q_2(\alpha)q_2 + Q_1(\alpha)\left[\int_{\delta-\Delta}^{\delta}(\Delta + \epsilon_1)dG(\epsilon_1) + \int_{-\delta}^{\delta-\Delta}\int_{-(\Delta+\epsilon_1)}^{\delta}(1 - G(\epsilon_2))d\epsilon_2 dG(\epsilon_1)\right],$$

and similarly, $U(\epsilon_2,\epsilon_1;\alpha) = \int_{-\delta}^{\delta} u(\epsilon_1|\epsilon_2;\alpha)dG(\epsilon_2)$ and so from (A.2),

$$U(\epsilon_2,\epsilon_1;\alpha) = Q_2(\alpha)q_2 + Q_1(\alpha)\left[\int_{-\delta}^{\Delta-\delta}(\Delta - \epsilon_2)dG(\epsilon_2) + \int_{\Delta-\delta}^{\delta}\int_{-(\Delta-\epsilon_2)}^{\delta}(1 - G(\epsilon_1))d\epsilon_1 dG(\epsilon_2)\right].$$

Note that

$$\int_{\delta-\Delta}^{\delta}(\Delta+\epsilon_1)dG(\epsilon_1) = \delta G(\Delta-\delta) + \int_{-\delta}^{\Delta-\delta} G(\epsilon)d\epsilon = \int_{-\delta}^{\Delta-\delta}(\Delta-\epsilon_2)dG(\epsilon_2),$$

using the integration by parts. Thus, the first terms in the square-bracket of $U(\epsilon_1,\epsilon_2;\alpha)$ and $U(\epsilon_2,\epsilon_1;\alpha)$ are identical. Next, observe also that

$$\begin{aligned}
&\int_{-\delta}^{\delta-\Delta}\int_{-(\Delta+\epsilon_1)}^{\delta}(1-G(\epsilon_2))d\epsilon_2\,dG(\epsilon_1) = \int_{-\delta}^{\delta-\Delta}\int_{-(\Delta+\epsilon_1)}^{\delta}G(-\epsilon_2)d\epsilon_2\,dG(\epsilon_1)\\
&= \int_{-\delta}^{\delta-\Delta}\int_{-\delta}^{\Delta+\epsilon_1}G(t)dt\,dG(\epsilon_1) = \int_{\Delta-\delta}^{\delta}\int_{-\delta}^{\Delta-s}G(t)dt\,dG(s)\\
&= \int_{\Delta-\delta}^{\delta}\int_{-\delta}^{\Delta-s}(1-G(-t))dt\,dG(s) = \int_{\Delta-\delta}^{\delta}\int_{-\Delta+\epsilon_2}^{\delta}(1-G(\epsilon_1))d\epsilon_1\,dG(\epsilon_2).
\end{aligned}$$

This shows that the last terms in the square-bracket of $U(\epsilon_1,\epsilon_2;\alpha)$ and $U(\epsilon_2,\epsilon_1;\alpha)$ are identical. Arranging terms yields that

$$U(\epsilon_1,\epsilon_2;\alpha) = U(\epsilon_2,\epsilon_1;\alpha) = V_0(\alpha) + Q_1(\alpha)\int_{-\delta}^{\delta-\Delta}(1-G(\Delta+\epsilon))G(\epsilon)d\epsilon. \tag{A.4}$$

Thus, $U(\epsilon_i,\epsilon_j;\alpha) > U(\epsilon_i;\alpha) = V_0(\alpha)$ for $\Delta \geq \delta$. To see $U(\epsilon_i,\epsilon_j;\alpha) > U(\epsilon_i;\alpha)$ for $\Delta < \delta$, observe that from (A.3) and (A.4),

$$\begin{aligned}
&U(\epsilon_i,\epsilon_j;\alpha) - U(\epsilon_i;\alpha) = Q_1(\alpha)\left[\int_{-\delta}^{\delta-\Delta}(1-G(\Delta+\epsilon))G(\epsilon)d\epsilon - \int_{\Delta}^{\delta}(1-G(\epsilon))d\epsilon\right] \qquad \text{(A.5)}\\
&= Q_1(\alpha)\left[\int_{-\Delta}^{\delta-\Delta}(1-G(\Delta+\epsilon))G(\epsilon)d\epsilon + \int_{-\delta}^{-\Delta}(1-G(\Delta+\epsilon))G(\epsilon)d\epsilon - \int_{-\delta}^{-\Delta}G(\epsilon)d\epsilon\right]\\
&= Q_1(\alpha)\left[\int_{-\Delta}^{\delta-\Delta}(1-G(\Delta+\epsilon))G(\epsilon)d\epsilon - \int_{-\delta}^{-\Delta}G(\Delta+\epsilon)G(\epsilon)d\epsilon\right]\\
&= Q_1(\alpha)\left[\int_{-\Delta}^{\delta-\Delta}(1-G(\Delta+\epsilon))G(\epsilon)d\epsilon - \int_{-\delta}^{-\Delta}(1-G(-\Delta-\epsilon))(1-G(-\epsilon))d\epsilon\right]\\
&= Q_1(\alpha)\left[\int_{-\Delta}^{\delta-\Delta}(1-G(\Delta+\epsilon))G(\epsilon)d\epsilon - \int_{0}^{\delta-\Delta}(1-G(t))(1-G(\Delta+t))dt\right]\\
&= Q_1(\alpha)\left[\int_{-\Delta}^{0}(1-G(\Delta+\epsilon))G(\epsilon)d\epsilon + \int_{0}^{\delta-\Delta}(1-G(\Delta+\epsilon))(2G(\epsilon)-1)d\epsilon\right] > 0,
\end{aligned}$$

where the fifth equality follows from change of variable ($t = -\Delta-\epsilon$), and the last equality holds since that $G(\epsilon) > \frac{1}{2}$ for any $\epsilon > 0$ by the symmetry of G. ■

Lemma A3. *Suppose $\hat{s}_1 > \hat{s}_2$. For each α, the following results hold.*

(i) There exists $\overline{c}(\alpha)$ such that $V(\alpha) > V_0(\alpha)$ whenever $c < \overline{c}(\alpha)$.

(ii) There exists α^* such that $c < \overline{c}(\alpha)$ if and only if $\alpha > \alpha^*$.

Proof. *(i).* Fix any α and suppose that $\Delta \geq \delta$. In this case, $U(\epsilon_i;\alpha) = V_0(\alpha)$ and from (A.4),

$$U(\epsilon_i,\epsilon_j;\alpha) - U(\epsilon_i;\alpha) = Q_1(\alpha)\int_{-\delta}^{\delta-\Delta}(1-G(\epsilon))d\epsilon =: \hat{c}(\alpha).$$

Therefore, if $c \geq \hat{c}(\alpha)$, $V(\alpha) = U(\epsilon_i;\alpha) - c < V_0(\alpha)$; and if $c < \hat{c}(\alpha)$, $V(\alpha) = U(\epsilon_i,\epsilon_j;\alpha) - 2c$. In the latter case, $V(\alpha) > V_0(\alpha)$ if and only if $c < \frac{\hat{c}(\alpha)}{2}$ from (A.4). Hence, we let $\overline{c}(\alpha) = \frac{\hat{c}(\alpha)}{2}$.

Next, suppose that $\Delta < \delta$. Define

$$\check{c}(\alpha) := U(\epsilon_i,\epsilon_j;\alpha) - U(\epsilon_i;\alpha) = Q_1(\alpha)\left[\int_{-\delta}^{\delta-\Delta}(1-G(\Delta+\epsilon))G(\epsilon)d\epsilon - \int_{\Delta}^{\delta}(1-G(\epsilon))d\epsilon\right],$$
$$\tilde{c}(\alpha) := U(\epsilon_i;\alpha) - V_0(\alpha) = Q_1(\alpha)\int_{\Delta}^{\delta}(1-G(\epsilon))d\epsilon,$$

from (A.5) and (A.3), respectively. Note that $U(\epsilon_i,\epsilon_j) - 2c > U(\epsilon_i) - c$ if and only if $c < \check{c}(\alpha)$, and $u(\epsilon_i;\alpha) - c > V_0(\alpha)$ if and only if $c < \tilde{c}(\alpha)$. Next, let

$$W(\Delta;\alpha) := \check{c}(\alpha) - \tilde{c}(\alpha) = Q_1(\alpha)\left[\int_{-\delta}^{\delta-\Delta}(1-G(\Delta+\epsilon))G(\epsilon)d\epsilon - 2\int_{\Delta}^{\delta}(1-G(\epsilon))d\epsilon\right]$$

Observe that $W(\Delta;\alpha)$ is strictly increasing in Δ and $W(0;\alpha) < 0 < W(\delta;\alpha)$.[24] Hence, there exists a unique $\Delta^o \in (0,\delta)$ such that $W(\Delta^o;\alpha) = 0$.

- For $\Delta \in [\Delta^o,\delta)$, it holds that $\tilde{c}(\alpha) < \check{c}(\alpha)$. Note that for $c \geq \check{c}(\alpha)$, $U(\epsilon_i,\epsilon_j;a) - 2c \leq U(\epsilon_i;\alpha) - c \leq V_0$, where the last inequality holds since $U(\epsilon_i;\alpha) - c \leq V_0$ for $c \geq \tilde{c}(\alpha)$ and $\tilde{c}(\alpha) \leq \check{c}(\alpha) \leq c$. Thus, no student with α learns in this case. Next, for $c < \check{c}(\alpha)$, $U(\epsilon_i,\epsilon_j;a) - 2c > U(\epsilon_i;\alpha) - c$. In this case, $U(\epsilon_i,\epsilon_j;\alpha) - 2c > V_0(\alpha)$ if $c < \frac{\hat{c}(\alpha)}{2}$ and $U(\epsilon_i,;\alpha) - c > V_0(\alpha)$ if $c < \tilde{c}(\alpha)$. Hence, let $\overline{c}(\alpha) = \max\{\frac{\hat{c}(\alpha)}{2}, \tilde{c}(\alpha)\}$.
- For $\Delta < \Delta^0$, it holds that $\check{c}(\alpha) < \tilde{c}(\alpha)$. Note that for $c \geq \tilde{c}(\alpha)$, $U(\epsilon_i,\epsilon_j;\alpha) - 2c < U(\epsilon_i;\alpha) - c \leq V_0(\alpha)$, where the first inequality holds since $c > \check{c}(\alpha)$. Hence, no student learns. For $c < \tilde{c}(\alpha)$, $V(\alpha) \geq U(\epsilon_i;\alpha) - c > V_0(\alpha)$. Hence, we let $\overline{c}(\alpha) = \tilde{c}(\alpha)$.

In sum, $\overline{c}(\alpha)$ is given as follows:

$$\overline{c}(\alpha) := \begin{cases} \frac{\hat{c}(\alpha)}{2} & \text{for } \Delta \geq \delta, \\ \max\left\{\frac{\hat{c}(\alpha)}{2}, \tilde{c}(\alpha)\right\} & \text{for } \Delta^o \leq \Delta < \delta, \\ \tilde{c}(\alpha) & \text{for } \Delta < \Delta^o. \end{cases}$$

[24] $W(0) = Q_1(\alpha)\left[\int_{-\delta}^{\delta}(1-G(\epsilon))G(\epsilon)d\epsilon - 2\int_0^{\delta}(1-G(\epsilon))d\epsilon\right] = Q_1(\alpha)\left[2\int_0^{\delta}(1-G(\epsilon))G(\epsilon)d\epsilon - 2\int_0^{\delta}(1-G(\epsilon))d\epsilon\right] < 0$ by the symmetry of G, and $W(\delta;\alpha) = Q_1(\alpha)\int_{-\delta}^{0}(1-G(\delta+\epsilon))G(\epsilon)d\epsilon > 0$.

(ii). Note that since $Q_1(\alpha)$ is increasing in α, so are $\check{c}(\alpha)$ and $\tilde{c}(\alpha)$, further implying that $\overline{c}(\alpha)$ is increasing in α. For a given c, let $\alpha^* := \inf\{\alpha | \overline{c}(\alpha) \geq c\}$. It is clear that α^* is increasing in c and $c < \overline{c}(\alpha)$ if and only if $\alpha > \alpha^*$. ■

Lemmas A1 to A3 so far are based on fixed $\hat{s}_1 > \hat{s}_2$. In equilibrium, $(\hat{s}_1, \hat{s}_2)$ must be chosen to make the mass of students assigned to each college equal to its capacity k. We now show that there is a unique equilibrium in which $\hat{s}_1 > \hat{s}_2$.

Proof of Theorem 1. Let $\mathfrak{m}_i$ denote the mass of students assigned to each college i. Assuming that all students with $\alpha > \alpha^*$ learn ϵ_1 first, $\mathfrak{m}_1$ and $\mathfrak{m}_2$ are given as follows:

$$\mathfrak{m}_1 = \int_{\underline{\alpha}}^{\alpha^*} Q_1(\alpha) dF(\alpha) + m_{12} \quad \text{and} \quad \mathfrak{m}_2 = \int_{\underline{\alpha}}^{\overline{\alpha}} [Q_2(\alpha) - Q_1(\alpha)] dF(\alpha) + m_{21}, \tag{A.6}$$

where

$$m_{12} \equiv \int_{\alpha^*}^{\overline{\alpha}} Q_1(\alpha) \Big[\text{Prob}(\epsilon_1 \geq \overline{\epsilon}(\alpha) - \Delta) + \text{Prob}(|\epsilon_1 + \Delta| < \overline{\epsilon}(\alpha), \epsilon_2 \leq \epsilon_1 + \Delta) \Big] dF(\alpha),$$
$$m_{21} \equiv \int_{\alpha^*}^{\overline{\alpha}} Q_1(\alpha) \Big[\text{Prob}(\epsilon_1 \leq -\overline{\epsilon}(\alpha) - \Delta) + \text{Prob}(|\epsilon_1 + \Delta| < \overline{\epsilon}(\alpha), \epsilon_2 > \epsilon_1 + \Delta) \Big] dF(\alpha).$$

In what follows, we first show that there is a unique pair $(\hat{s}_1, \hat{s}_2)$ satisfying $\mathfrak{m}_1 = k = \mathfrak{m}_2$, and then show that such a pair must satisfy $\hat{s}_1 > \hat{s}_2$. The proof consists of several steps.

Step 1. *For any given* $(\hat{s}_1, \hat{s}_2)$, $m_{12} > m_{21}$

Proof . For any α, let $\overline{Q}(\alpha) \equiv \text{Prob}(s \geq \hat{s}_1, s \geq \hat{s}_2 | \alpha)$. Rewrite m_{12} and m_{21} as follows:

$$m_{12} = \int_{\alpha^*}^{\overline{\alpha}} \overline{Q}(\alpha) \left[(1 - G(\overline{\epsilon}(\alpha) - \Delta)) + \int_{-\overline{\epsilon}(\alpha)-\Delta}^{\overline{\epsilon}(\alpha)-\Delta} G(\epsilon_1 + \Delta) dG(\epsilon_1) \right] dF(\alpha), \tag{A.7}$$
$$m_{21} = \int_{\alpha^*}^{\overline{\alpha}} \overline{Q}(\alpha) \left[G(-\overline{\epsilon}(\alpha) - \Delta) + \int_{-\overline{\epsilon}(\alpha)-\Delta}^{\overline{\epsilon}(\alpha)-\Delta} (1 - G(\epsilon_1 + \Delta)) dG(\epsilon_1) \right] dF(\alpha). \tag{A.8}$$

Observe that for any given α, $1 - G(\overline{\epsilon}(\alpha) - \Delta) = G(-\overline{\epsilon}(\alpha) + \Delta) > G(-\overline{\epsilon}(\alpha) - \Delta)$ and

$$\begin{aligned} \int_{-\overline{\epsilon}(\alpha)-\Delta}^{\overline{\epsilon}(\alpha)-\Delta} (1 - G(\epsilon_1 + \Delta)) dG(\epsilon_1) &= \int_{-\overline{\epsilon}(\alpha)-\Delta}^{\overline{\epsilon}(\alpha)-\Delta} G(-\epsilon_1 - \Delta) dG(\epsilon_1) \\ &= \int_{-\overline{\epsilon}(\alpha)}^{\overline{\epsilon}(\alpha)} G(t) dG(t) = \int_{-\overline{\epsilon}(\alpha)-\Delta}^{\overline{\epsilon}(\alpha)-\Delta} G(\epsilon_1 + \Delta) dG(\epsilon). \end{aligned}$$

Thus, we have $m_{12} > m_{21}$. □

Step 2. $\hat{s}_1 > \hat{s}_2$.

Proof . Suppose $\hat{s}_1 = \hat{s}_2$. Then, $Q_1(\alpha) = Q_2(\alpha)$, so $\mathfrak{m}_1 = \int_{\underline{\alpha}}^{\alpha^*} Q_1(\alpha)dF(\alpha) + m_{12} > \mathfrak{m}_2 = m_{21}$, a contradiction. Next, suppose $\hat{s}_1 < \hat{s}_2$. Then, we have

$$\mathfrak{m}_1 = \int_{\underline{\alpha}}^{\alpha^*} Q_1(\alpha)dF(\alpha) + m_{12} > m_{21} > \mathfrak{m}_2 = \int_{\underline{\alpha}}^{\alpha^*} (Q_2(\alpha) - Q_1(\alpha))dF(\alpha) + m_{21},$$

where the second equality holds since $Q_2(\alpha) < Q_1(\alpha)$ for each α due to that $\hat{s}_1 < \hat{s}_2$, yielding a contradiction, again. Thus, we must have $\hat{s}_1 > \hat{s}_2$. □

Step 3. *There is a unique pair* $(\hat{s}_1, \hat{s}_2)$ *such that* $\mathfrak{m}_1 = k = \mathfrak{m}_2$.

Proof . First, note that since $s = r\alpha + (1-r)\theta$, $\alpha \in [\underline{\alpha}, \overline{\alpha}]$ and $\theta \in [-\eta, \eta]$, we have $s \in [\underline{s}, \overline{s}]$ where $\underline{s} \equiv r\underline{\alpha} - (1-r)\eta < \overline{s} \equiv r\overline{\alpha} + (1-r)\eta$. Consider $\mathfrak{m}_1$. Since $Q_1(\alpha)$ is decreasing in $\hat{s}_1$ and so is $\mathfrak{m}_1$. Moreover, $\hat{s}_1 < \overline{s}$ since otherwise $Q_1(\alpha) = 0$ for all α and so $\mathfrak{m}_1 = 0$, and $\hat{s}_1 > \underline{s}$ since otherwise $Q_1(\alpha) = 1$ for all α and so $\hat{s}_1 = \underline{s}_1 \leq \hat{s}_2$, a contradiction to Step 2. Since $Q_1(\alpha)$ is strictly decreasing in $\hat{s}_1$ for $\hat{s}_1 \in (\underline{s}, \overline{s})$, it follows that there is a unique $\hat{s}_1$ satisfying $\mathfrak{m}_1 = k$.

Next, consider $\mathfrak{m}_2$. Observe that for the fixed $\hat{s}_1$ defined above, it is clear that $\hat{s}_2 \in (\underline{s}, \hat{s}_1)$. Note also that if $\hat{s}_2 \leq \underline{s}$, then $\mathfrak{m}_2 = 1 - \mathfrak{m}_1 = 1 - k > k$, where the last equality follows from the definition of $\hat{s}_1$ and the inequality holds since $k < \frac{1}{2}$. Similarly, if $\hat{s}_2 \geq \hat{s}_1$, then $\mathfrak{m}_2 \leq m_{21} < m_{12} < \mathfrak{m}_1 = k$, where the first inequality holds since $Q_2(\alpha) \leq Q_1(\alpha)$, and the second inequality follows from Step 1. Since $\mathfrak{m}_2$ is strictly decreasing in $\hat{s}_2$ for $\hat{s}_2 \in (\underline{s}, \hat{s}_1)$, the desired result follows. □

Step 4. *There is no equilibrium with* $\hat{s}_1 \leq \hat{s}_2$.

Proof . Suppose to the contrary that there is such an equilibrium. Observe that in this case, students with $s \geq \hat{s}_2$ will be assigned to whichever college they rank first in their ROL, and those with $s \in [\hat{s}_1, \hat{s}_2)$ (if $\hat{s}_1 < \hat{s}_2$) are assigned to college 1 regardless of their ROLs. A straightforward analysis yields that

$$u(\epsilon_i; \alpha) = Q_1(\alpha)\mathbb{E}[v_1|I] + Q_2(\alpha)(\mathbb{E}[v_2|I] - \mathbb{E}[v_1|I])\mathbf{1}_A,$$

where $\mathbb{E}[v_1|I] = q_1 + \epsilon_1$, $\mathbb{E}[v_2|I] = q_2$ and $A = \{\epsilon_1 | \epsilon_1 < -\Delta\}$ if $\sigma(\alpha) = (1, \emptyset)$; or $\mathbb{E}[v_1|I] = q_1$, $\mathbb{E}[v_2|I] = q_2 + \epsilon_2$ and $A = \{\epsilon_2 | \epsilon_2 > \Delta\}$ if $\sigma(\alpha) = (2, \emptyset)$. Similarly,

$$u(\epsilon_j|\epsilon_i; \alpha) = Q_1(\alpha)\mathbb{E}[v_1|I] + Q_2(\alpha)\int_A (-\Delta + \epsilon_2 - \epsilon_1)dG(\epsilon_{\sigma_2}),$$

where $\mathbb{E}[v_1|I] = q_1 + \epsilon_1$ and $A = \{\epsilon_2 | \epsilon_2 > \epsilon_1 + \Delta\}$ if $i = 1$, $j = 2$; or $\mathbb{E}[v_1] = q_1$ and $A = \{\epsilon_1 | \epsilon_1 < \epsilon_2 - \Delta\}$ if $i = 2$, $j = 1$. Using them, it is easy to see that for any ϵ_1,

$$u(\epsilon_2|\epsilon_1; \alpha) - u(\epsilon_1; \alpha) = Q_2(\alpha)\int_{|\epsilon_1 + \Delta|}^{\delta} (1 - G(\epsilon_2))d\epsilon_2,$$

since $u(\epsilon_1;\alpha) = Q_1(\alpha)(q_1+\epsilon_1)$ if $\epsilon_1 \geq -\Delta$; or $u(\epsilon_1;\alpha) = Q_1(\alpha)(q_1+\epsilon_1) + Q_2(\alpha)(-\epsilon_1-\Delta)$ if $\epsilon_1 < -\delta$, and

$$u(\epsilon_2|\epsilon_1;\alpha) = Q_1(\alpha)(q_1+\epsilon_1) + Q_2(\alpha)\int_{\epsilon_1+\Delta}^{\delta}(1-G(\epsilon_2))d\epsilon_2.$$

Similarly, we also have that for any ϵ_2,

$$u(\epsilon_1|\epsilon_2;\alpha) - u(\epsilon_2;a) = Q_2(\alpha)\int_{|\epsilon_2-\Delta|}^{\delta}(1-G(\epsilon_2))d\epsilon_2.$$

Therefore, the results from Lemmas A1 to A3 follow (with $V_0(\alpha) = Q_1(\alpha)q_1 - Q_2(\alpha)\Delta$), which in turn imply that

$$\begin{aligned}
m_{12} &= \int_{\alpha^*}^{\overline{\alpha}} Q_2(\alpha)\Big[\text{Prob}(\epsilon_1 \geq \overline{\epsilon}(\alpha)-\Delta) + \text{Prob}(|\epsilon_1+\Delta| < \overline{\epsilon}(\alpha), \epsilon_2 \leq \epsilon_1+\Delta)\Big]dF(\alpha) \\
m_{21} &= \int_{\alpha^*}^{\overline{\alpha}} Q_2(\alpha)\Big[\text{Prob}(\epsilon_1 \leq -\overline{\epsilon}(\alpha)-\Delta) + \text{Prob}(|\epsilon_1+\Delta| < \overline{\epsilon}(\alpha), \epsilon_2 > \epsilon_1+\Delta)\Big]dF(a),
\end{aligned}$$

assuming that all students with $\alpha \geq \alpha^*$ learn ϵ_1 first. Note that m_{ij} captures the mass of students who submit ROL $i \succ j$ among those with $\alpha \geq \alpha^*$. Thus, the mass of students assigned to each college $i = 1, 2$ is given by

$$\mathfrak{m}_1 = \int_{\underline{\alpha}}^{\alpha^*} Q_1(\alpha)dF(\alpha) + \int_{\alpha^*}^{\overline{\alpha}}(Q_1(\alpha)-Q_2(\alpha))dF(\alpha) + m_{12} \quad \text{and} \quad \mathfrak{m}_2 = m_{21},$$

where the first term in the RHS of $\mathfrak{m}_1$ is the mass of students who submit ROL $1 \succ 2$ without learning the suits (since $q_1 > q_2$) among those assigned to college 1, and the second term those with $s \in [\hat{s}_1, \hat{s}_2)$ and so assigned to college 1 regardless of their ROLs. Note that $m_{12} > m_{21}$ by the same argument in Step 1, which implies that $\mathfrak{m}_1 > \mathfrak{m}_2$, a contradiction. □

A.2 Proof of Theorem 2

Consider the benchmark first and suppose that all students with $s \geq \hat{s}_1^D$ learn ϵ_1 first whenever $c < \overline{c}$. Then, $SW^B = MV^B - TC^B$. Note that $MV^B = MV_1^B + MV_2^B$ and

$$\begin{aligned}
MV_1^B &= \int_{\underline{a}}^{\overline{\alpha}} Q_1^B(\alpha)\left[\int_{\overline{\epsilon}-\Delta}^{\delta}(q_1+\epsilon_1)dG(\epsilon_1) + \int_{-\overline{\epsilon}-\Delta}^{\overline{\epsilon}-\Delta}\left(\int_{-\delta}^{\epsilon_1+\Delta}(q_1+\epsilon_1)dG(\epsilon_2)\right)dG(\epsilon_1)\right]dF(\alpha), \\
MV_2^B &= \int_{\underline{a}}^{\overline{\alpha}} Q_1^B(\alpha)\left[\int_{-\delta}^{-\overline{\epsilon}-\Delta} q_2 dG(\epsilon_1) + \int_{-\overline{\epsilon}-\Delta}^{\overline{\epsilon}-\Delta}\left(\int_{\epsilon_1+\Delta}^{\delta}(q_2+\epsilon_2)dG(\epsilon_2)\right)dG(\epsilon_1)\right]dF(\alpha) \\
&\quad + \int_{\underline{\alpha}}^{\overline{\alpha}} q_2(Q_2^B(\alpha) - Q_1^B(\alpha))dF(\alpha).
\end{aligned}$$

Thus, we have

$$\begin{aligned}MV^B = \int_{\underline{\alpha}}^{\overline{\alpha}} Q_1^B(\alpha)\Bigg[&\int_{\overline{\epsilon}-\Delta}^{\delta} (\Delta+\epsilon_1)dG(\epsilon_1) + \int_{-\delta}^{\overline{\epsilon}-\Delta} (q_2-q_2)dG(\epsilon_1) \\ &+ \int_{-\overline{\epsilon}-\Delta}^{\overline{\epsilon}-\Delta}\left(\int_{-\delta}^{\epsilon_1+\Delta}(\Delta+\epsilon_1)dG(\epsilon_2) + \int_{\epsilon_1+\Delta}^{\delta}(q_2+\epsilon_2-q_2)dG(\epsilon_2)\right)dG(\epsilon_1)\Bigg]dF(\alpha) \\ &+ \int_{\underline{\alpha}}^{\overline{\alpha}} q_2 Q_2^B(\alpha)dF(\alpha).\end{aligned}$$

The total learning cost is given by

$$TC^B = c\,m_L^B = \int_{\underline{\alpha}}^{\overline{\alpha}} Q_1^B(\alpha)\left[\int_{-\delta}^{-\overline{\epsilon}-\Delta} c\,dG(\epsilon_1) + \int_{\overline{\epsilon}-\Delta}^{\delta} c\,dG(\epsilon_1) + \int_{-\overline{\epsilon}-\Delta}^{\overline{\epsilon}-\Delta} 2c\,dG(\epsilon_2)\right]dF(\alpha).$$

Next, consider DA and suppose that students with $\alpha > \alpha^*$ learn ϵ_1 first. Then, $SW^D = MV^D - TC^D$. Note that

$$\begin{aligned}MV_1^D &= \int_{\alpha^*}^{\overline{\alpha}} Q_1^D(\alpha)\left[\int_{\overline{\epsilon}(\alpha)-\Delta}^{\delta}(q_1+\epsilon_1)dG(\epsilon_1) + \int_{-\overline{\epsilon}(\alpha)-\Delta}^{\overline{\epsilon}(\alpha)-\Delta}\int_{-\delta}^{\epsilon_1+\Delta}(q_1+\epsilon_1)dG(\epsilon_2)dG(\epsilon_1)\right]dF(\alpha) \\ &\quad + \int_{\underline{\alpha}}^{\alpha^*} q_1 Q_1^D(\alpha)dF(\alpha), \\ MV_2^D &= \int_{\alpha^*}^{\overline{\alpha}} Q_1^D(\alpha)\left[\int_{-\delta}^{-\overline{\epsilon}(\alpha)-\Delta} q_2 dG(\epsilon_1) + \int_{-\overline{\epsilon}(\alpha)-\Delta}^{\overline{\epsilon}(\alpha)-\Delta}\int_{\epsilon_1+\Delta}^{\delta}(q_2+\epsilon_2)dG(\epsilon_2)dG(\epsilon_1)\right]dF(\alpha) \\ &\quad + \int_{\underline{\alpha}}^{\overline{\alpha}} q_2(Q_2^D(\alpha) - Q_1^D(\alpha))dF(\alpha)\end{aligned}$$

and so $MV^D = MV_1^D + MV_2^D$ is

$$\begin{aligned}MV^D = \int_{\alpha^*}^{\overline{\alpha}} Q_1^D(\alpha)\Bigg[&\int_{\overline{\epsilon}(\alpha)-\Delta}^{\delta}(\Delta+\epsilon_1)dG(\epsilon_1) \\ &+ \int_{-\overline{\epsilon}(\alpha)-\Delta}^{\overline{\epsilon}(\alpha)-\Delta}\left(\int_{-\delta}^{\epsilon_1+\Delta}(\Delta+\epsilon_1)dG(\epsilon_2) + \int_{\epsilon_1+\Delta}^{\delta}\epsilon_2 dG(\epsilon_2)\right)dG(\epsilon_1)\Bigg]dF(\alpha) \\ &+ \int_{\underline{\alpha}}^{\overline{\alpha}} q_2 Q_2^D(\alpha)dF(\alpha) + \int_{\underline{\alpha}}^{\alpha^*}\Delta\, Q_1^D(\alpha)dF(\alpha),\end{aligned} \tag{A.9}$$

and TC^D is written as

$$TC^D = c\,m_L^D = \int_{\alpha^*}^{\overline{\alpha}}\left[\int_{-\delta}^{-\overline{\epsilon}(\alpha)-\Delta} c\,dG(\epsilon_1) + \int_{\overline{\epsilon}(\alpha)-\Delta}^{\delta} c\,dG(\epsilon_1) + \int_{-\overline{\epsilon}(\alpha)-\Delta}^{\overline{\epsilon}(\alpha)-\Delta} 2c\,dG(\epsilon_2)\right]dF(\alpha). \tag{A.10}$$

We now establishes a series of lemmas that prove Theorem 2.

Lemma A4. $SW^B > SW^D$ *for any* r.

Proof. Note that

$$
\begin{aligned}
SW^B =& MV^B - TC^B \\
&= \int_{\underline{\alpha}}^{\overline{\alpha}} Q_1^B(\alpha)\Bigg[\int_{\underline{\epsilon}-\Delta}^{\delta}(\Delta+\epsilon_1-c)dG(\epsilon_1)+\int_{-\delta}^{\overline{\epsilon}-\Delta}(-c)dG(\epsilon_1) \\
&\qquad + \int_{-\overline{\epsilon}-\Delta}^{\overline{\epsilon}-\Delta}\left(\int_{-\delta}^{\epsilon_1+\Delta}(\Delta+\epsilon_1-2c)dG(\epsilon_2)+\int_{\epsilon_1+\Delta}^{\delta}(\epsilon_2-2c)dG(\epsilon_2)\right)dG(\epsilon_1)\Bigg]dF(\alpha) \\
&\quad + \int_{\underline{a}}^{\overline{a}} Q_2^B(\alpha)q_2 dF(\alpha).
\end{aligned}
\tag{A.11}
$$

and

$$
\begin{aligned}
SW^D &= MV^D - TC^D \\
&\leq MV^D - \int_{\alpha^*}^{\overline{\alpha}} Q_1^D(\alpha)\Bigg[\int_{-\delta}^{-\overline{\epsilon}(\alpha)-\Delta} c\,dG(\epsilon_1)+\int_{\overline{\epsilon}-\Delta}^{\delta} c\,dG(\epsilon_1)+\int_{-\overline{\epsilon}-\Delta}^{\overline{\epsilon}-\Delta}2c\,dG(\epsilon_2)\Bigg]dF(\alpha) \\
&= \int_{\alpha^*}^{\overline{\alpha}} Q_1^D(\alpha)\Bigg[\int_{\overline{\epsilon}(\alpha)-\Delta}^{\delta}(\Delta+\epsilon_1-c)dG(\epsilon_1)+\int_{-\delta}^{\overline{\epsilon}(\alpha)-\Delta}(-c)dG(\epsilon_1)\Bigg]dF(\alpha) \\
&\qquad + \int_{-\overline{\epsilon}(\alpha)-\Delta}^{\overline{\epsilon}(\alpha)-\Delta}\left(\int_{-\delta}^{\epsilon_1+\Delta}(\Delta+\epsilon_1-2c)dG(\epsilon_2)+\int_{\epsilon_1+\Delta}^{\delta}(\epsilon_2-2c)dG(\epsilon_2)\right)dG(\epsilon_1)\Bigg]dF(\alpha) \\
&\quad + \int_{\underline{\alpha}}^{\overline{\alpha}} Q_2^D(\alpha)q_2\,dF(\alpha) + \int_{\underline{\alpha}}^{\alpha^*} Q_1^D(\alpha)\Delta\,dF(\alpha) \\
&\equiv \overline{SW}^D
\end{aligned}
\tag{A.12}
$$

Since $SW^B - SW^D \geq SW^B - \overline{SW}^D$, we show $SW^B > \overline{SW}^D$ in what follows. Note that

$$
\begin{aligned}
& SW^B - \overline{SW}^D \\
&= \int_{\underline{\alpha}}^{\overline{\alpha}} Q_1^B(\alpha)\Bigg[\int_{\underline{\epsilon}-\Delta}^{\delta}(\Delta+\epsilon_1-c)dG(\epsilon_1)+\int_{-\delta}^{\overline{\epsilon}-\Delta}(-c)dG(\epsilon_1) \\
&\qquad + \int_{-\overline{\epsilon}-\Delta}^{\overline{\epsilon}-\Delta}\left(\int_{-\delta}^{\epsilon_1+\Delta}(\Delta+\epsilon_1-2c)dG(\epsilon_2)+\int_{\epsilon_1+\Delta}^{\delta}(\epsilon_2-2c)dG(\epsilon_2)\right)dG(\epsilon_1)\Bigg]dF(\alpha) \\
&\quad - \int_{\alpha^*}^{\overline{\alpha}} Q_1^D(\alpha)\Bigg[\int_{\overline{\epsilon}(\alpha)-\Delta}^{\delta}(\Delta+\epsilon_1-c)dG(\epsilon_1)+\int_{-\delta}^{\overline{\epsilon}(\alpha)-\Delta}(-c)dG(\epsilon_1)\Bigg]dF(\alpha) \\
&\qquad + \int_{-\overline{\epsilon}(\alpha)-\Delta}^{\overline{\epsilon}(\alpha)-\Delta}\left(\int_{-\delta}^{\epsilon_1+\Delta}(\Delta+\epsilon_1-2c)dG(\epsilon_2)+\int_{\epsilon_1+\Delta}^{\delta}(\epsilon_2-2c)dG(\epsilon_2)\right)dG(\epsilon_1)\Bigg]dF(\alpha) \\
&\quad - \int_{\underline{\alpha}}^{\alpha^*} \Delta\, Q_1^D(\alpha)dF(\alpha)
\end{aligned}
$$

$$= \int_{\underline{\alpha}}^{\overline{\alpha}} Q_1^B(\alpha)\Big[\int_{\underline{\epsilon}-\Delta}^{\delta}(\epsilon_1 - c)dG(\epsilon_1) + \int_{-\delta}^{\overline{\epsilon}-\Delta}(-c)dG(\epsilon_1)$$

$$+ \int_{-\overline{\epsilon}-\Delta}^{\overline{\epsilon}-\Delta}\Big(\int_{-\delta}^{\epsilon_1+\Delta}(\epsilon_1 - 2c)dG(\epsilon_2) + \int_{\epsilon_1+\Delta}^{\delta}(\epsilon_2 - 2c)dG(\epsilon_2)\Big)dG(\epsilon_1)\Big]dF(\alpha)$$

$$- \int_{\alpha^*}^{\overline{\alpha}} Q_1^D(\alpha)\Big[\int_{\overline{\epsilon}(\alpha)-\Delta}^{\delta}(\epsilon_1 - c)dG(\epsilon_1) + \int_{-\delta}^{\overline{\epsilon}(\alpha)-\Delta}(-c)dG(\epsilon_1)\Big]dF(\alpha)$$

$$+ \int_{-\overline{\epsilon}(\alpha)-\Delta}^{\overline{\epsilon}(\alpha)-\Delta}\Big(\int_{-\delta}^{\epsilon_1+\Delta}(\epsilon_1 - 2c)dG(\epsilon_2) + \int_{\epsilon_1+\Delta}^{\delta}(\epsilon_2 - 2c)dG(\epsilon_2)\Big)dG(\epsilon_1)\Big]dF(\alpha).$$

The first equality follows from the fact that

$$\int_{\underline{\alpha}}^{\overline{\alpha}} Q_2^B(\alpha)dF(\alpha) = 2\kappa = \int_{\underline{\alpha}}^{\overline{\alpha}} Q_2^D(\alpha)dF(\alpha),$$

and to understand the second equality, observe that

$$k = \mathfrak{m}_1^D = \int_{\alpha^*}^{\overline{\alpha}} Q_1^D(\alpha)\Big[\int_{\overline{\epsilon}(\alpha)-\Delta}^{\delta} 1dG(\epsilon_1) + \int_{-\overline{\epsilon}(\alpha)-\Delta}^{\overline{\epsilon}(\alpha)-\Delta}\Big(\int_{-\delta}^{\epsilon_1+\Delta} 1dG(\epsilon_2)\Big)dG(\epsilon_2)\Big]dF(\alpha) + \int_{\underline{\alpha}}^{\alpha*} Q_1^D(\alpha)dF(\alpha), \tag{A.13}$$

so, we have

$$\int_{\underline{\alpha}}^{\alpha^*} \Delta Q_1^D(\alpha)dF(\alpha)$$

$$= \Delta\,\kappa - \int_{\alpha^*}^{\overline{\alpha}} Q_1^D(\alpha)\Big[\int_{\overline{\epsilon}(\alpha)-\Delta}^{\delta} \Delta dG(\epsilon_1) + \int_{-\overline{\epsilon}(\alpha)-\Delta}^{\overline{\epsilon}(\alpha)-\Delta}\Big(\int_{-\delta}^{\epsilon_1+\Delta} \Delta dG(\epsilon_2)\Big)dG(\epsilon_1)\Big]dF(\alpha)$$

$$= \int_{\underline{\alpha}}^{\overline{\alpha}} Q_1^B(\alpha)\Big[\int_{\overline{\epsilon}-\Delta}^{\delta} \Delta dG(\epsilon_1) + \int_{-\overline{\epsilon}-\Delta}^{\overline{\epsilon}-\Delta}\Big(\int_{-\delta}^{\epsilon_1+\Delta} \Delta dG(\epsilon_2)\Big)dG(\epsilon_1)\Big]dF(\alpha)$$

$$- \int_{\alpha^*}^{\overline{\alpha}} Q_1^D(\alpha)\Big[\int_{\overline{\epsilon}(\alpha)-\Delta}^{\delta} \Delta dG(\epsilon_1) + \int_{-\overline{\epsilon}(\alpha)-\Delta}^{\overline{\epsilon}(\alpha)-\Delta}\Big(\int_{-\delta}^{\epsilon_1+\Delta} \Delta dG(\epsilon_2)\Big)dG(\epsilon_1)\Big]dF(\alpha),$$

where the last equality follows from the capacity constraint of college 1 in the benchmark, that is,

$$\mathfrak{m}_1^B = \int_{\underline{\alpha}}^{\overline{\alpha}} Q_1^B(\alpha)\Big[\int_{\overline{\epsilon}-\Delta}^{\delta} 1dG(\epsilon_1) + \int_{-\overline{\epsilon}-\Delta}^{\overline{\epsilon}-\Delta}\Big(\int_{-\delta}^{\epsilon_1+\Delta} 1dG(\epsilon_2)\Big)dG(\epsilon_1)\Big]dF(\alpha) = k. \tag{A.14}$$

Claim A2. $\int_{\underline{\alpha}}^{\overline{\alpha}} Q_1^B(\alpha)dF(\alpha) \geq \int_{\alpha^*}^{\overline{\alpha}} Q_1^D(\alpha)dF(\alpha)$.

Proof . Consider the terms in the square bracket of $\mathfrak{m}_1^B$ and $\mathfrak{m}_1^D$ in (A.14) and (A.13), respectively. Observe that for any fixed α,

$$\begin{aligned}
&\left[\int_{\overline{\epsilon}(\alpha)-\Delta}^{\delta} 1dG(\epsilon_1) + \int_{-\overline{\epsilon}(\alpha)-\Delta}^{\overline{\epsilon}(\alpha)-\Delta}\left(\int_{-\delta}^{\epsilon_1+\Delta} 1dG(\epsilon_2)\right)dG(\epsilon_2)\right] \\
&-\left[\int_{\overline{\epsilon}-\Delta}^{\delta} 1dG(\epsilon_1) + \int_{-\overline{\epsilon}-\Delta}^{\overline{\epsilon}-\Delta}\left(\int_{-\delta}^{\epsilon_1+\Delta} 1dG(\epsilon_2)\right)dG(\epsilon_1)\right] \\
&= \int_{\overline{\epsilon}(\alpha)-\Delta}^{\overline{\epsilon}-\Delta} 1dG(\epsilon_1) - \int_{\overline{\epsilon}(\alpha)-\Delta}^{\overline{\epsilon}-\Delta} G(\epsilon_1+\Delta)dG(\epsilon_1) - \int_{-\overline{\epsilon}-\Delta}^{-\overline{\epsilon}(\alpha)-\Delta} G(\epsilon_1+\Delta)dG(\epsilon_1) \\
&= \int_{\overline{\epsilon}(\alpha)-\Delta}^{\overline{\epsilon}-\Delta} (1-G(\epsilon_1+\Delta))dG(\epsilon_1) - \int_{-\overline{\epsilon}-\Delta}^{-\overline{\epsilon}(\alpha)-\Delta} G(\epsilon_1+\Delta)dG(\epsilon_1) \\
&= \int_{\overline{\epsilon}(\alpha)-\Delta}^{\overline{\epsilon}-\Delta} G(-\epsilon_1-\Delta)dG(\epsilon_1) - \int_{-\overline{\epsilon}-\Delta}^{-\overline{\epsilon}(\alpha)-\Delta} G(\epsilon_1+\Delta)dG(\epsilon_1) = 0,
\end{aligned}$$

where the last equality holds since

$$\int_{\overline{\epsilon}(\alpha)-\Delta}^{\overline{\epsilon}-\Delta} G(-\epsilon_1-\Delta)dG(\epsilon_1) = \int_{-\overline{\epsilon}}^{-\overline{\epsilon}(\alpha)} G(t)dG(t-\Delta) = \int_{-\overline{\epsilon}-\Delta}^{-\overline{\epsilon}(\alpha)-\Delta} G(\eta+\Delta)dG(\eta),$$

form a sequence of change of variables $t = -\epsilon_1 - \Delta$ and $\eta = t - \Delta$. Thus, we have

$$\mathfrak{m}_1^D = \int_{\alpha^*}^{\overline{\alpha}} Q_1^D(\alpha)dF(\alpha)\left[\int_{\overline{\epsilon}-\Delta}^{\delta} 1dG(\epsilon_1)+\int_{-\overline{\epsilon}-\Delta}^{\overline{\epsilon}-\Delta}\left(\int_{-\delta}^{\epsilon_1+\Delta} 1dG(\epsilon_2)\right)dG(\epsilon_1)\right]+\int_{\underline{\alpha}}^{\alpha*} Q_1^C(\alpha)dF(\alpha).$$

and from the fact that $\mathfrak{m}_1^B = k = \mathfrak{m}_1^D$, we further have

$$\begin{aligned}
&\left(\int_{\underline{\alpha}}^{\overline{\alpha}} Q_1^B(\alpha)dF(\alpha) - \int_{\alpha^*}^{\overline{\alpha}} Q_1^D(\alpha)dF(\alpha)\right)\left[\int_{\overline{\epsilon}-\Delta}^{\delta} 1dG(\epsilon_1) + \int_{-\overline{\epsilon}-\Delta}^{\overline{\epsilon}-\Delta}\left(\int_{-\delta}^{\epsilon_1+\Delta} 1dG(\epsilon_2)\right)dG(\epsilon_1)\right] \\
&= \int_{\underline{\alpha}}^{\alpha^*} Q_1^D(\alpha)dF(\alpha) \geq 0,
\end{aligned}$$

which yields the desired result.□

Now, using Claim A2, we have

$$\begin{aligned}
&SW^B - \overline{SW} \\
&\geq \int_{\alpha^*}^{\overline{\alpha}} Q_1^D(\alpha)\left[\int_{\underline{\epsilon}-\Delta}^{\delta}(\epsilon_1-c)dG(\epsilon_1) + \int_{-\delta}^{\overline{\epsilon}-\Delta}(-c)dG(\epsilon_1)\right. \\
&\qquad\left. + \int_{-\overline{\epsilon}-\Delta}^{\overline{\epsilon}-\Delta}\left(\int_{-\delta}^{\epsilon_1+\Delta}(\epsilon_1-2c)dG(\epsilon_2) + \int_{\epsilon_1+\Delta}^{\delta}(\epsilon_2-2c)dG(\epsilon_2)\right)dG(\epsilon_1)\right]dF(\alpha) \\
&\quad - \int_{\alpha^*}^{\overline{\alpha}} Q_1^D(\alpha)\left[\int_{\overline{\epsilon}(\alpha)-\Delta}^{\delta}(\epsilon_1-c)dG(\epsilon_1) + \int_{-\delta}^{\overline{\epsilon}(\alpha)-\Delta}(-c)dG(\epsilon_1)\right]dF(\alpha)
\end{aligned}$$

$$+\int_{-\overline{\epsilon}(\alpha)-\Delta}^{\overline{\epsilon}(\alpha)-\Delta}\left(\int_{-\delta}^{\epsilon_1+\Delta}(\epsilon_1-2c)dG(\epsilon_2)+\int_{\epsilon_1+\Delta}^{\delta}(\epsilon_2-2c)dG(\epsilon_2)\right)dG(\epsilon_1)\Bigg]dF(\alpha)$$

$$=\int_{\alpha^*}^{\overline{\alpha}}Q_1^D(\alpha)\Bigg[\int_{\overline{\epsilon}(\alpha)-\Delta}^{\overline{\epsilon}-\Delta}\left(\int_{-\delta}^{\epsilon_1+\Delta}(\epsilon_1-2c)dG(\epsilon_2)+\int_{\epsilon_1+\Delta}^{\delta}(\epsilon_2-2c)dG(\epsilon_2)-(\epsilon_1-c)\right)dG(\epsilon_1)$$

$$+\int_{-\overline{\epsilon}-\Delta}^{-\overline{\epsilon}(\alpha)-\Delta}\left(\int_{-\delta}^{\epsilon_1+\Delta}(\epsilon_1-2c)dG(\epsilon_2)+\int_{\epsilon_1+\Delta}^{\delta}(\epsilon_2-2c)dG(\epsilon_2)-(-c)\right)dG(\epsilon_1)\Bigg]dF(\alpha)$$

$$=\int_{\alpha^*}^{\overline{\alpha}}Q_1^D(\alpha)\Bigg[\int_{\overline{\epsilon}(\alpha)-\Delta}^{\overline{\epsilon}-\Delta}\left(\int_{\epsilon_1+\Delta}^{\delta}(\epsilon_2-(\epsilon_1+\Delta))dG(\epsilon_2)-c\right)dG(\epsilon_1)$$

$$+\int_{-\overline{\epsilon}-\Delta}^{-\overline{\epsilon}(\alpha)-\Delta}\left(\int_{-\delta}^{\epsilon_1+\Delta}((\epsilon_1+\Delta)-\epsilon_2)dG(\epsilon_2)-c\right)dG(\epsilon_1)\Bigg]dF(\alpha)>0.$$

To see the last inequality, observe that

$$\int_{\overline{\epsilon}(\alpha)-\Delta}^{\overline{\epsilon}-\Delta}\left(\int_{\epsilon_1+\Delta}^{\delta}(\epsilon_2-(\epsilon_1+\Delta))dG(\epsilon_2)-c\right)dG(\epsilon_1)$$

$$=\int_{\overline{\epsilon}(\alpha)-\Delta}^{\overline{\epsilon}-\Delta}\left(\int_{\epsilon_1+\Delta}^{\delta}\epsilon_2 dG(\epsilon_2)-(\epsilon_1+\Delta)(1-G(\epsilon_1+\Delta))-c\right)dG(\epsilon_1)$$

$$=\int_{\overline{\epsilon}(\alpha)-\Delta}^{\overline{\epsilon}-\Delta}\left(\delta-(\epsilon_1+\Delta)G(\epsilon_1+\Delta)-\int_{\epsilon_1+\Delta}^{\delta}G(\epsilon_2)d\epsilon_2-(\epsilon_1+\Delta)(1-G(\epsilon_1+\Delta))-c\right)dG(\epsilon_1)$$

$$=\int_{\overline{\epsilon}(\alpha)-\Delta}^{\overline{\epsilon}-\Delta}\left(\delta-(\epsilon_1+\Delta)-\int_{\epsilon_1+\Delta}^{\delta}G(\epsilon_2)d\epsilon_2-\int_{\overline{\epsilon}}^{\delta}(1-G(\epsilon))d\epsilon\right)dG(\epsilon_1)$$

$$=\int_{\overline{\epsilon}(\alpha)-\Delta}^{\overline{\epsilon}-\Delta}\left(\overline{\epsilon}-(\epsilon_1+\Delta)-\int_{\epsilon_1+\Delta}^{\overline{\epsilon}}G(\epsilon_2)d\epsilon_2\right)dG(\epsilon_1)=\int_{\overline{\epsilon}(\alpha)-\Delta}^{\overline{\epsilon}-\Delta}\left(\int_{\epsilon_1+\Delta}^{\overline{\epsilon}}(1-G(\epsilon_2))d\epsilon_2\right)dG(\epsilon_1)>0$$

where the second equality follows from the integration by parts, and the third equality follows from the definition of $\overline{\epsilon}$. Similarly, we also have

$$\int_{-\overline{\epsilon}-\Delta}^{\overline{\epsilon}(\alpha)-\Delta}\left(\int_{-\delta}^{\overline{\epsilon}(\alpha)-\Delta}((\epsilon_1+\Delta)-\epsilon_2)dG(\epsilon_2)-c\right)dG(\epsilon_1)>0.$$

Therefore, we have $SW^B > SW^D$ for any $r\in[0,1)$. ■

Lemma A5. *SW^B is invariant in r, and $SW^D = SW^B$ at $r=1$.*

Proof. To show that SW^B is invariant in r, it suffices to show that $\int_{\underline{\alpha}}^{\overline{\alpha}}Q_i^B(\alpha)dF(\alpha)$ is invariant in r for all $i=1,2$. This is clear from (A.14), $\int_{\underline{\alpha}}^{\overline{\alpha}}Q_1^B(\alpha)dF(\alpha)$ is a constant. Using this and $\mathfrak{m}_2^B=k$, it also follows that $\int_{\underline{\alpha}}^{\overline{\alpha}}Q_2^B(\alpha)dF(\alpha)$ does not depend on r.

Next, consider the case that $r=1$ so that $s=\alpha$ for each α. In the benchmark, there are $\hat{\alpha}_1^B>\hat{\alpha}_2^B$ such that $Q_1^B(\alpha)=\mathbb{1}_{\{\alpha\geq\hat{\alpha}_1^B\}}$ and $Q_2^B(\alpha)=\mathbb{1}_{\{\alpha\geq\hat{\alpha}_2^B\}}$. Except for this, students' learning decisions are the same as before. Similarly, in DA, there are $\hat{\alpha}_1^D>\hat{\alpha}_2^D$ such that $Q_1^D(\alpha)=\mathbb{1}_{\{\alpha\geq\hat{\alpha}_1^D\}}$ and $Q_2^D(\alpha)=\mathbb{1}_{\{\alpha\geq\hat{\alpha}_2^D\}}$, and students' learning decisions are the same as before. Thus,

$\overline{c}(\alpha) = Q_1^D(\alpha)\overline{c}$ is the same as $\overline{c}$ for $\alpha \geq \hat{a}_1^D$ and zero otherwise, and $Q_1^D(\alpha)\int_{\overline{\epsilon}}^{\delta}(1-G(\epsilon))d\epsilon$ is these same as $\int_{\overline{\epsilon}}^{\delta}(1-G(\epsilon))d\epsilon$ and zero otherwise. Therefore, $\hat{\alpha}_i^B = \hat{\alpha}_i^D$ for all $i = 1,2$ and, consequently, $MV_i^B = MV_i^D$ and $TC^B = TC^D$. ■

Lemma A6. *Suppose* $\Delta = 0$. *Then,* SW^D *increases with* r.

Proof. The proof consists of several steps: we first show that $\hat{s}_1 = \hat{s}_2$ in equilibrium and then show that SW^D increases with r.

Step 1. $\hat{s}_1 = \hat{s}_2$.

Proof . Suppose $\hat{s}_i > \hat{s}_j$. Then, students with $s \geq \hat{s}_i$ can attend whichever college they rank higher, while those with $s \in [\hat{s}_j, \hat{s}_i)$ will be assigned to college j. For the former group, the mass of students who prefer college 1 over college 2 (and 2 over 1) is given by (A.7) and (A.8), with $\Delta = 0$. That is,

$$m_{12} = \int_{\alpha^*}^{\overline{\alpha}} Q_i(\alpha)\left[1 - G(\overline{\epsilon}(\alpha)) + \int_{-\overline{\epsilon}(\alpha)}^{\overline{\epsilon}(\alpha)} G(\epsilon_1)dG(\epsilon_1)\right]dF(\alpha),$$
$$m_{21} = \int_{\alpha^*}^{\overline{\alpha}} Q_i(\alpha)\left[G(-\overline{\epsilon}(\alpha)) + \int_{-\overline{\epsilon}(\alpha)}^{\overline{\epsilon}(\alpha)} (1 - G(\epsilon_1))dG(\epsilon_1)\right]dF(\alpha).$$

For a given α, the symmetry of $G(\cdot)$ implies that $1 - G(\overline{\epsilon}(\alpha)) = G(-\overline{\epsilon}(\alpha))$ and

$$\int_{-\overline{\epsilon}(\alpha)}^{\overline{\epsilon}(\alpha)} (1 - G(\epsilon_1))dG(\epsilon_1) = \int_{-\overline{\epsilon}(\alpha)}^{\overline{\epsilon}(\alpha)} G(-\epsilon_1)dG(\epsilon_1) = \int_{-\overline{\epsilon}(\alpha)}^{\overline{\epsilon}(\alpha)} G(\epsilon_1)dG(\epsilon_1).$$

Therefore, $m_{12} = m_{21}$, which leads to a contradiction: either college i does not fill its capacity, or college j exceeds its capacity. Hence, it must be that $\hat{s}_1 = \hat{s}_2$. □

Step 2. $\frac{dQ(\alpha)}{dr} = \frac{\alpha - \mathbb{E}[\alpha]}{2\eta(1-r)^2}$.

Proof . Denote by $\hat{s} := \hat{s}_1 = \hat{s}_2$, and $Q_1(\alpha) = Q_2(\alpha) =: Q(\alpha) = \text{Prob}(s \geq \hat{s}|\alpha)$. Since $\theta \in [-\eta, \eta]$ follows the uniform distribution, we have $Q(\alpha) = 1 - \frac{1}{2\eta}\left(\frac{\hat{s} - r\alpha}{1-r} + \eta\right)$. Note that $\hat{s}$ is determined by the colleges' joint capacity constraint:

$$\int_{\underline{\alpha}}^{\overline{\alpha}} Q(\alpha)dF(\alpha) = 2k \iff \frac{1}{2\eta}\int_{\underline{\alpha}}^{\overline{\alpha}}\left(\frac{\hat{s} - r\alpha}{1-r} + \eta\right)dF(\alpha) = \frac{1}{2\eta}\left(\frac{\hat{s}}{1-r} - \frac{r}{1-r}\mathbb{E}[\alpha] + \eta\right) = 1 - 2k,$$

so $\hat{s} = r\mathbb{E}[\alpha] + \eta(1-r)(1-4k)$. Substituting this into $Q(\alpha)$ above, we have

$$Q(\alpha) = 1 - \frac{\frac{r(\mathbb{E}[\alpha]-\alpha)}{1-r} + \eta(1-4k) + \eta}{2\eta} \Rightarrow \frac{dQ(\alpha)}{dr} = \frac{\alpha - \mathbb{E}[\alpha]}{2\eta(1-r)^2}.$$

□

Step 3. *SW^D increases in r.*

Proof . From (A.9) and (A.10), we have

$$MV^D = \int_{\alpha^*}^{\overline{\alpha}} Q(\alpha)\Bigg[\underbrace{\int_{\overline{\epsilon}(\alpha)}^{\delta} \epsilon_1 dG(\epsilon_1) + \int_{-\overline{\epsilon}(\alpha)}^{\overline{\epsilon}(\alpha)}\left(\int_{-\delta}^{\epsilon_1} \epsilon_1 dG(\epsilon_2) + \int_{\epsilon_1}^{\delta} \epsilon_2 dG(\epsilon_2)\right) dG(\epsilon_1)}_{=:\mathcal{E}}\Bigg]dF(\alpha)$$

$$+ \underbrace{\int_{\underline{\alpha}}^{\overline{\alpha}} qQ(\alpha)dF(\alpha),}_{=2kq}$$

$$TC^D = \int_{\alpha^*}^{\overline{\alpha}}\left[\int_{-\delta}^{-\overline{\epsilon}(\alpha)} cdG(\epsilon_1) + \int_{\overline{\epsilon}(\alpha)}^{\delta} cdG(\epsilon_1) + \int_{-\overline{\epsilon}(\alpha)}^{\overline{\epsilon}(\alpha)} 2cdG(\epsilon_2)\right]dF(\alpha),$$

where $\overline{\epsilon}(\alpha)$ satisfies $Q(\alpha)\int_{\overline{\epsilon}}^{\delta}(1-G(\epsilon))d\epsilon = c$, and α^* satisfies $\overline{c}(\alpha^*) = c$ with $\overline{c}(\alpha)$ being given by $\overline{c}(\alpha) := Q(\alpha)\int_0^{\delta}(1-G(\epsilon))d\epsilon$. Note that at $\alpha = \alpha^*$, $\overline{\epsilon}(\alpha)$ satisfies

$$Q(\alpha^*)\int_{\overline{\epsilon}}^{\delta}(1-G(\epsilon)) = c = Q(\alpha^*)\int_0^{\delta}(1-G(\epsilon))d\epsilon,$$

where the second equality follows from the definition of α^*. Hence, $\overline{\epsilon}(\alpha^*) = 0$.

Now, observe that

$$\begin{aligned}\frac{dMV^D}{dr} &= -Q(\alpha^*)\left[\int_{\overline{\epsilon}(\alpha^*)}^{\delta} \epsilon_1 dG(\epsilon_1) + \int_{-\overline{\epsilon}(\alpha^*)}^{\overline{\epsilon}(\alpha^*)}\left(\int_{-\delta}^{\epsilon_1} \epsilon_1 dG(\epsilon_2) + \int_{\epsilon_1}^{\delta} \epsilon_2 dG(\epsilon_2)\right) dG(\epsilon_1)\right] f(\alpha^*)\frac{d\alpha^*}{dr}\\
&\quad + \int_{\alpha^*}^{\overline{\alpha}}\left(\frac{dQ(\alpha)}{dr}\mathcal{E} + Q(\alpha)\frac{d\mathcal{E}}{dr}\right)dF(\alpha)\\
&= -Q(\alpha^*)\int_0^{\delta}\epsilon_1 dG(\epsilon_1) f(\alpha^*)\frac{d\alpha^*}{dr} + \int_{\alpha^*}^{\overline{\alpha}}\left(\frac{dQ(\alpha)}{dr}\mathcal{E} + Q(\alpha)\frac{d\mathcal{E}}{dr}\right)dF(\alpha)\end{aligned}$$

where the last equality holds since $\overline{\epsilon}(\alpha^*) = 0$. Note also that $\frac{d\mathcal{E}}{dr}$ is given by

$$\begin{aligned}\frac{d\mathcal{E}}{dr} &= -\overline{\epsilon}(\alpha)g(\overline{\epsilon}(\alpha))\frac{d\overline{\epsilon}(\alpha)}{dr} + \left(\int_{-\delta}^{\overline{\epsilon}(\alpha)}\overline{\epsilon}(\alpha)dG(\epsilon_2) + \int_{\overline{\epsilon}(\alpha)}^{\delta}\epsilon_2 dG(\epsilon_2)\right)g(\overline{\epsilon}(\alpha))\frac{d\overline{\epsilon}(\alpha)}{dr}\\
&\quad + \left(\int_{-\delta}^{-\overline{\epsilon}(\alpha)}(-\overline{\epsilon}(\alpha))dG(\epsilon_2) + \int_{-\overline{\epsilon}(\alpha)}^{\delta}\epsilon_2 dG(\epsilon_2)\right)g(-\overline{\epsilon}(\alpha))\frac{d\overline{\epsilon}(\alpha)}{dr}\\
&= g(\overline{\epsilon}(\alpha))\frac{d\overline{\epsilon}(\alpha)}{dr}\left(-\overline{\epsilon}(\alpha) + \overline{\epsilon}(\alpha)G(\overline{\epsilon}(\alpha)) + \int_{\overline{\epsilon}(\alpha)}^{\delta}\epsilon_2 dG(\epsilon_2)\right)\\
&\quad + g(-\overline{\epsilon}(\alpha))\frac{d\overline{\epsilon}(\alpha)}{dr}\left(-\overline{\epsilon}(\alpha)G(-\overline{\epsilon}(\alpha)) + \int_{-\overline{\epsilon}(\alpha)}^{\delta}\epsilon_2 dG(\epsilon_2)\right)\\
&= \left[g(\overline{\epsilon}(\alpha))\left(-\overline{\epsilon}(\alpha) + \delta - \int_{\overline{\epsilon}(\alpha)}^{\delta}G(\epsilon_2)d\epsilon_2\right) + g(-\overline{\epsilon}(\alpha))\left(\delta - \int_{-\overline{\epsilon}(\alpha)}^{\delta}G(\epsilon_2)d\epsilon_2\right)\right]\frac{d\overline{\epsilon}(\alpha)}{dr}\end{aligned}$$

$$= \left[g(\bar\epsilon(\alpha))\left(\int_{\bar\epsilon(\alpha)}^{\delta}(1-G(\epsilon))d\epsilon\right) + g(-\bar\epsilon(\alpha))\left(-\bar\epsilon(\alpha) + \int_{-\bar\epsilon(\alpha)}^{\delta}(1-G(\epsilon))d\epsilon_2\right)\right]\frac{d\bar\epsilon(\alpha)}{dr}$$
$$= \left[g(\bar\epsilon(\alpha) + g(-\bar\epsilon(\alpha))\right]\int_{\bar\epsilon(\alpha)}^{\delta}(1-G(\epsilon))d\epsilon\frac{d\bar\epsilon(\alpha)}{dr},$$

where the last equality holds since $-\bar\epsilon + \int_{-\bar\epsilon}^{\delta}(1-G(\epsilon))d\epsilon = \int_{\bar\epsilon}^{\delta}(1-G(\epsilon))d\epsilon$ by Claim A1.

Next, observe also that

$$\frac{dTC}{dr} = -\left\{c + c\left[G(\bar\epsilon(\alpha^*)) - G(-\bar\epsilon(\alpha^*))\right]\right\}f(\alpha^*)\frac{d\alpha^*}{dr} + \int_{\alpha^*}^{\overline{\alpha}} c\left[g(\bar\epsilon(\alpha)) + cg(-\bar\epsilon(\alpha))\right]\frac{d\bar\epsilon(\alpha)}{d\alpha}dF(\alpha)$$
$$= -cf(\alpha^*)\frac{d\alpha^*}{dr} + \int_{\alpha^*}^{\overline{\alpha}} c\left[g(\bar\epsilon(\alpha)) + cg(-\bar\epsilon(\alpha))\right]\frac{d\bar\epsilon(\alpha)}{dr}dF(\alpha),$$

where the last equality holds since $\bar\epsilon(\alpha^*) = 0$ so $G(\bar\epsilon(\alpha^*)) = G(-\bar\epsilon(\alpha^*)) = G(0)$.

Therefore, after arranging terms, it follows that

$$\frac{dSW^D}{dr} = \frac{dMV^D}{dr} - \frac{dTC^D}{dr}$$
$$= \int_{\alpha^*}^{\overline{\alpha}}\frac{dQ(\alpha)}{dr}\mathcal{E}dF(\alpha) + \int_{\alpha^*}^{\overline{\alpha}}\left\{Q(\alpha)\frac{d\mathcal{E}}{dr} - c\left[g(\bar\epsilon(\alpha)) + cg(-\bar\epsilon(\alpha))\right]\frac{d\bar\epsilon(\alpha)}{dr}\right\}dF(\alpha)$$
$$+ f(\alpha^*)\frac{d\alpha^*}{dr}\left[-Q(\alpha^*)\int_0^{\delta}\epsilon dG(\epsilon) + c\right]$$
$$= \int_{\alpha^*}^{\overline{\alpha}}\frac{dQ(\alpha)}{dr}\mathcal{E}dF(\alpha),$$

where the last equality holds since

$$\int_{\alpha^*}^{\overline{\alpha}}\left\{Q(\alpha)\frac{d\mathcal{E}}{dr} - c\left[g(\bar\epsilon(\alpha)) + cg(-\bar\epsilon(\alpha))\right]\frac{d\bar\epsilon(\alpha)}{d\alpha}\right\}dF(\alpha)$$
$$= \int_{\alpha^*}^{\overline{\alpha}}\left[Q(\alpha)\int_{\bar\epsilon(\alpha)}^{\delta}(1-G(\epsilon))d\epsilon - c\right]g(\bar\epsilon(\alpha))\frac{d\bar\epsilon(\alpha)}{dr}dF(\alpha)$$
$$+ \int_{\alpha^*}^{\overline{\alpha}}\left[Q(\alpha)\left(-\bar\epsilon(\alpha) + \int_{-\bar\epsilon(\alpha)}^{\delta}(1-G(\epsilon))d\epsilon\right) - c\right]g(-\bar\epsilon(\alpha))\frac{d\bar\epsilon(\alpha)}{dr}dF(\alpha) = 0$$

by definition of $\bar\epsilon(\alpha)$ (recall that $-\bar\epsilon + \int_{-\bar\epsilon}^{\delta}(1-G(\epsilon))d\epsilon = \int_{\bar\epsilon}^{\delta}(1-G(\epsilon))d\epsilon = c$), and

$$Q(\alpha^*)\int_0^{\delta}\epsilon dG(\epsilon) = Q(\alpha^*)\int_0^{\delta}(1-G(\epsilon))d\epsilon = c$$

by definition of α^*. Next, observe that $\mathcal{E}$ is increasing in α, since

$$\frac{d\mathcal{E}}{d\alpha} = \left[g(\bar\epsilon(\alpha) + g(-\bar\epsilon(\alpha))\right]\int_{\bar\epsilon(\alpha)}^{\delta}(1-G(\epsilon))d\epsilon\frac{d\bar\epsilon(\alpha)}{d\alpha}$$

and $\bar{\epsilon}(\alpha)$ is increasing in α. Hence, we have

$$\frac{dSW^D}{dr} = \int_{\alpha^*}^{\overline{\alpha}} \frac{dQ(\alpha)}{dr}\mathcal{E}dF(\alpha) \geq \frac{1}{2\eta(1-r)^2}\int_{\alpha^*}^{\overline{\alpha}}(\alpha - \mathbb{E}[\alpha])dF(\alpha)\int_{\alpha^*}^{\overline{\alpha}}\mathcal{E}dF(\alpha) \geq 0,$$

where the first inequality follows from the fact that the covariance of increasing functions of a random variable is positive, and the last inequality holds since

$$\begin{aligned}\int_{\alpha^*}^{\overline{\alpha}}(\alpha - \mathbb{E}[\alpha])dF(\alpha) &= \int_{\alpha^*}^{\overline{\alpha}}\alpha dF(\alpha) - \mathbb{E}[\alpha](1-F(\alpha^*)) = (1-F(\alpha^*))\left(\frac{\int_{\alpha^*}^{\overline{\alpha}}\alpha dF(\alpha)}{1-F(\alpha^*)} - \mathbb{E}[\alpha]\right)\\ &= (1-F(\alpha^*))\big(\mathbb{E}[\alpha|\alpha\geq\alpha^*] - \mathbb{E}[\alpha]\big) \geq 0.\end{aligned}$$

Thus, SW^D increases with r. □ ■

B Three Colleges with $\Delta = 0$ in Remark 1.

Consider the case of three colleges with $\Delta = 0$, that is, $q_1 = q_2 = q_3 \equiv q$. In this setting, we must have $\hat{s}_1 = \hat{s}_2 = \hat{s}_3 \equiv \hat{s}$ and we let $Q(\alpha) := \text{Prob}(s \geq \hat{s}|\alpha)$.

To see students' learning decisions, consider a student who has already learned both ϵ_i and ϵ_j, with $\epsilon_i \geq \epsilon_j$. Note that the student's learning decision in this stage is the same as that under the baseline model since she only compares ϵ_i and ϵ_k. Specifically, if she does not learn ϵ_k, then she will rank college i highest if $\epsilon_i \geq 0$, and otherwise prefer either j or k. Her expected payoff in this case is $u(\epsilon_i;\alpha) = Q(\alpha)(q+\epsilon_i)$ if $\epsilon_i \geq 0$, and $u(\epsilon_i;\alpha) = Q(\alpha)q$ if $\epsilon_i < 0$. Now, suppose she also learns ϵ_k. Then, she will rank i above k if $\epsilon_i \geq \epsilon_k$, and k above i otherwise. Her expected payoff is then

$$\begin{aligned}u(\epsilon_k|\epsilon_i \geq \epsilon_j;\alpha) &= Q(\alpha)\big\{\text{Prob}(\epsilon_i \geq \epsilon_k)(q+\epsilon_i) + \text{Prob}(\epsilon_i < \epsilon_k)\mathbb{E}[q+\epsilon_k|\epsilon_i<\epsilon_k]\big\}\\ &= Q(\alpha)\left[q + \epsilon_i + \int_{\epsilon_i}^{\delta}(1-G(\epsilon_k))d\epsilon_k\right] \qquad \text{(B.1)}\end{aligned}$$

Thus, the gain from learning ϵ_k is

$$u(\epsilon_k|\epsilon_i \geq \epsilon_j;\alpha) - u(\epsilon_i;\alpha) = \begin{cases}\int_{\epsilon_i}^{\delta}(1-G(\epsilon_k))d\epsilon_l & \text{if } \epsilon_i > 0\\ \epsilon_i + \int_{\epsilon_i}^{\delta}(1-G(\epsilon_k))d\epsilon_k & \text{if } \epsilon_i \leq 0.\end{cases}$$

Using Claim A1 and the same logic as in the baseline model, we conclude that the student learns ϵ_k in addition to $\epsilon_i > \epsilon_j$ if and only if $|\epsilon_i| < \bar{\epsilon}(\alpha)$, where $\bar{\epsilon}(\alpha)$ satisfies

$$Q(\alpha)\int_{\bar{\epsilon}}^{\delta}(1-G(\epsilon))d\epsilon = c.$$

Next, consider a student who has learned ϵ_i and decides whether to learn ϵ_j alone or both ϵ_j and ϵ_k. We analyze the cases $\epsilon_i > 0$ and $\epsilon_i \leq 0$ separately.

- Suppose $\epsilon_i > 0$. If the student does not learn ϵ_j (and hence not ϵ_k), her expected payoff is $u(\epsilon_i > 0;\alpha) = Q(\alpha)(q+\epsilon_i)$. If she learns ϵ_j only, she ranks college i highest if $\epsilon_i \geq \epsilon_j$, and ranks j highest otherwise. Her expected payoff becomes

$$\begin{aligned} u(\epsilon_j|\epsilon_i > 0;\alpha) &= Q(\alpha)\big\{\text{Prob}(\epsilon_i \geq \epsilon_j)(q+\epsilon_i) + \text{Prob}(\epsilon_i < \epsilon_j)\mathbb{E}[q+\epsilon_j|\epsilon_i < \epsilon_j]\big\} \\ &= Q(\alpha)\left[q+\epsilon_i+\int_{\epsilon_i}^{\delta}(1-G(\epsilon_j))d\epsilon_j\right] \end{aligned}$$

If she also learns ϵ_k, she ranks the college with the highest ϵ. So, her expected payoff is

$$\begin{aligned} u(\epsilon_k|\epsilon_i,\epsilon_j;\alpha) &= Q(\alpha)\big\{\text{Prob}(\epsilon_i \geq \epsilon_j)\mathbb{E}[u(\epsilon_k|\epsilon_i \geq \epsilon_j;\alpha)] + \text{Prob}(\epsilon_i < \epsilon_j)\mathbb{E}[u(\epsilon_k|\epsilon_i < \epsilon_j;\alpha)]\big\} \\ &= Q(\alpha)\left\{q+\epsilon_i+\int_{\epsilon_i}^{\delta}(1-G(\epsilon_j)^2)d\epsilon_j\right\}, \end{aligned}$$

where the last equality follows from substituting (B.1) into $u(\epsilon_k|\epsilon_i \geq \epsilon_j;\alpha)$ and $u(\epsilon_k|\epsilon_i < \epsilon_j;\alpha)$ and rearranging terms. The gains from learning ϵ_j alone and from learning ϵ_k after ϵ_j are given by

$$\begin{aligned} u(\epsilon_j|\epsilon_i \geq 0;\alpha) - u(\epsilon_i \geq 0;\alpha) &= Q(\alpha)\int_{\epsilon_i}^{\delta}(1-G(\epsilon))d\epsilon, \\ u(\epsilon_k|\epsilon_i,\epsilon_j;\alpha) - u(\epsilon_j|\epsilon_i \geq 0;\alpha) &= Q(\alpha)\int_{\epsilon_i}^{\delta}(1-G(\epsilon_j))G(\epsilon_j)d\epsilon_j. \end{aligned}$$

Define $\hat{\epsilon}(\alpha)$ such that

$$Q(\alpha)\int_{\hat{\epsilon}}^{\delta}(1-G(\epsilon_j))G(\epsilon_j)d\epsilon_j = c.$$

Observe that $\hat{\epsilon}(\alpha) < \bar{\epsilon}(\alpha)$, since otherwise

$$\int_{\bar{\epsilon}}^{\delta}(1-G(\epsilon_i))d\epsilon_i = \int_{\hat{\epsilon}}^{\delta}(1-G(\epsilon_i))G(\epsilon_i))d\epsilon_i < \int_{\hat{\epsilon}}^{\delta}(1-G(\epsilon_i))d\epsilon_i \leq \int_{\bar{\epsilon}}^{\delta}(1-G(\epsilon_i)d\epsilon_i,$$

where the first equality follows the definitions of $\bar{\epsilon}(\alpha)$ and $\hat{\epsilon}(\alpha)$. This yields a contradiction. Hence, the value of learning ϵ_j given $\epsilon_i \geq 0$ is

$$\max\{u(\epsilon_j|\epsilon_i \geq 0;\alpha) - c, u(\epsilon_k|\epsilon_i,\epsilon_j;\alpha) - 2c\} = \begin{cases} u(\epsilon_j|\epsilon_i \geq 0,\alpha) - c & \text{if } \epsilon_i > \hat{\epsilon}(\alpha), \\ u(\epsilon_k|\epsilon_i,\epsilon_j;\alpha) - 2c & \text{if } \epsilon_i < \hat{\epsilon}(\alpha). \end{cases}$$

This exceeds $u(\epsilon_i \geq 0;\alpha)$ if and only if $\epsilon_i < \overline{\epsilon}(\alpha)$, since $\hat{\epsilon}(\alpha) < \overline{\epsilon}(\alpha)$.

- Suppose $\epsilon_i \leq 0$. If the student does not learn ϵ_j, she ranks either college j or k above i, and her expected payoff is $u(\epsilon_i \leq 0;\alpha) = Q(\alpha)q$. If she learns ϵ_j but not ϵ_k, she ranks college k highest if $\epsilon_j < 0$ and j highest if $\epsilon_j > 0$. Her expected payoff is

$$\begin{aligned} u(\epsilon_j|\epsilon_i \leq 0;\alpha) &= Q(\alpha)\big\{\text{Prob}(\epsilon_j \leq 0)q + \text{Prob}(\epsilon_j > 0 \geq \epsilon_i)\mathbb{E}[q + \epsilon_j|\epsilon_j > 0]\big\} \\ &= Q(\alpha)\left[q + \int_0^{\delta}(1 - G(\epsilon_j))d\epsilon_j\right] \end{aligned}$$

If she also learns ϵ_k, she compares all three ϵ's, and her expected payoff is

$$u(\epsilon_k|\epsilon_i,\epsilon_j) = Q(\alpha)\left[q + \epsilon_i + \int_{\epsilon_i}^{\delta}(1 - G(\epsilon_j)^2))d\epsilon_j\right].$$

The expected gains from learn ϵ_j alone and from learning ϵ_k after ϵ_j are respectively

$$\begin{aligned} u(\epsilon_j|\epsilon_i \leq 0;\alpha) - u(\epsilon_i \leq 0;\alpha) &= Q(\alpha)\int_0^{\delta}(1 - G(\epsilon_j))d\epsilon_j =: \overline{c}(\alpha), \\ u(\epsilon_k|\epsilon_i,\epsilon_j;\alpha) - u(\epsilon_j|\epsilon_i \leq 0;\alpha) &= Q(\alpha)\left[\epsilon_i + \int_{\epsilon_i}^{\delta}(1 - G(\epsilon_j))G(\epsilon_j)d\epsilon_j - \int_0^{\delta}(1 - G(\epsilon_j))d\epsilon_j\right]. \end{aligned}$$

Note that the latter is increasing in ϵ_i,

$$\frac{d[u(\epsilon_k|\epsilon_i,\epsilon_j;\alpha) - u(\epsilon_j|\epsilon_i < 0;\alpha)]}{d\epsilon_i} = Q(\alpha)\big[1 - G(\epsilon_i) + G(\epsilon_i)^2\big] > 0,$$

and is negative at $\epsilon_i = 0$: $u(\epsilon_k|\epsilon_i,\epsilon_j;\alpha) - u(\epsilon_j|\epsilon_i = 0;\alpha) = -Q(\alpha)\int_0^{\delta}(1 - G(\epsilon_j))^2 d\epsilon_j < 0$. Therefore, the value of learning ϵ_j given $\epsilon_i \leq 0$ is

$$\max\{u(\epsilon_j|\epsilon_i \leq 0,\alpha) - c, u(\epsilon_k|\epsilon_i,\epsilon_j;\alpha) - 2c\} = u(\epsilon_j;\epsilon_i \leq 0,\alpha) - c,$$

so the student will learn ϵ_j if and only if $c < \overline{c}(\alpha)$.

Lastly, we analyze students' learning decisions at the beginning—that is, whether to learn ϵ_i and subsequently ϵ_j and ϵ_k. If the student does not at all, her expected payoff is

$V_0(\alpha) := Q(\alpha)q$. If she learns ϵ_i but not ϵ_j and ϵ_k, then her expected payoff is

$$U(\epsilon_i;\alpha) = Q(\alpha)\big\{\text{Prob}(\epsilon_i > 0)\mathbb{E}[q+\epsilon_i|\epsilon_i \geq 0] + \text{Prob}(\epsilon_i \leq 0)q\big\} = Q(\alpha)\left[q + \int_0^\delta (1-G(\epsilon_i))d\epsilon_i\right].$$

If she further learns ϵ_j but not ϵ_k, her expected payoff is

$$\begin{aligned} U(\epsilon_i,\epsilon_j;\alpha) &= Q(\alpha)\big\{\text{Prob}(\epsilon_i > 0)\mathbb{E}[u(\epsilon_j|\epsilon_i > 0;\alpha)] + \text{Prob}(\epsilon_i \leq 0)\mathbb{E}[u(\epsilon_j|\epsilon_i \leq 0;\alpha)]\big\} \\ &= Q(\alpha)\left[q + \int_0^\delta (1-G(\epsilon_i)^2)d\epsilon_i\right]. \end{aligned}$$

If the student learns all three ϵ's, her expected payoff is

$$U(\epsilon_i,\epsilon_j,\epsilon_k;\alpha) = \mathbb{E}[u(\epsilon_k|\epsilon_i,\epsilon_j;\alpha)] = Q(\alpha)\left[q + \int_{-\delta}^\delta (1-G(\epsilon_i)^2)G(\epsilon_i)d\epsilon_i\right].$$

Therefore, the value of learning is given by

$$V(\alpha) := \max\{U(\epsilon_i;\alpha) - c, U(\epsilon_i,\epsilon_j;\alpha) - 2c, U(\epsilon_i,\epsilon_j,\epsilon_k;\alpha) - 3c\},$$

and the student will choose to learn ϵ_i if and only if $V(\alpha) > V_0(\alpha)$.

We define the following terms to capture the incremental gains from learning:

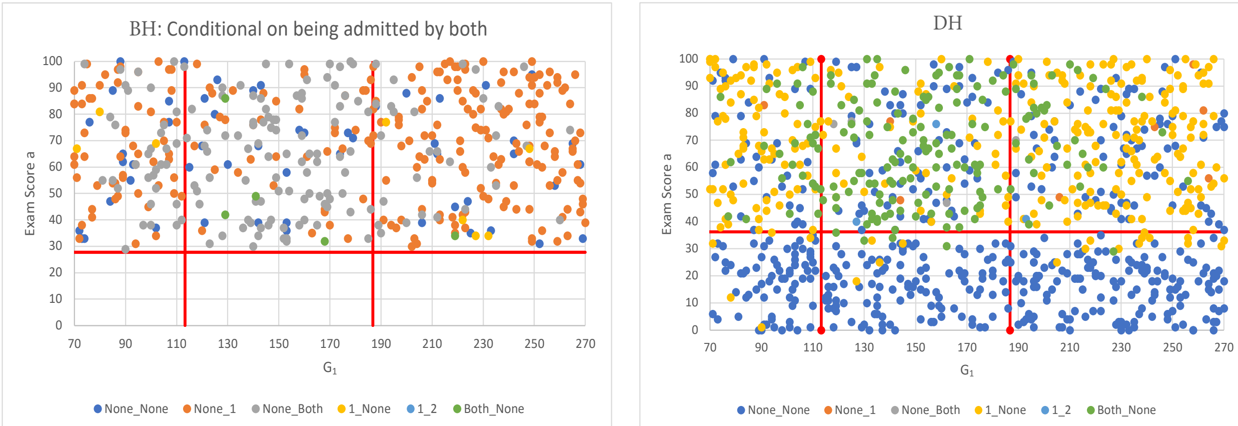
$$\begin{aligned} U(\epsilon_i;\alpha) - V_0(\alpha) &= Q(\alpha)\int_0^\delta (1-G(\epsilon))d\epsilon = \overline{c}(\alpha), \\ U(\epsilon_i,\epsilon_j;\alpha) - U(\epsilon_i;\alpha) &= Q(\alpha)\int_0^\delta (1-G(\epsilon))G(\epsilon)d\epsilon =: \hat{c}(\alpha), \\ U(\epsilon_i,\epsilon_j,\epsilon_k) - U(\epsilon_i,\epsilon_j;\alpha) &= Q(\alpha)\int_0^\delta (1-G(\epsilon))(2G(\epsilon)-1)d\epsilon =: \tilde{c}(\alpha). \end{aligned}$$

To compute $\tilde{c}(\alpha)$, observe that

$$\begin{aligned} U(\epsilon_i,\epsilon_j,\epsilon_k) - U(\epsilon_i,\epsilon_j;\alpha) &= Q(\alpha)\left[\int_{-\delta}^\delta (1-G(\epsilon)^2)G(\epsilon)d\epsilon - \int_0^\delta (1-G(\epsilon)^2)d\epsilon\right] \\ &= Q(\alpha)\left[\int_{-\delta}^0 (1-G(\epsilon)^2)G(\epsilon)d\epsilon - \int_0^\delta (1-G(\epsilon))(1-G(\epsilon)^2)d\epsilon\right] \\ &= Q(\alpha)\left[\int_0^\delta (1-(1-G(t))^2)(1-G(t))dt - \int_0^\delta (1-G(\epsilon))(1-G(\epsilon)^2)d\epsilon\right] \\ &= Q(\alpha)\left[\int_0^\delta (1-G(\epsilon))(2G(\epsilon)-1)d\epsilon\right] > 0, \end{aligned}$$

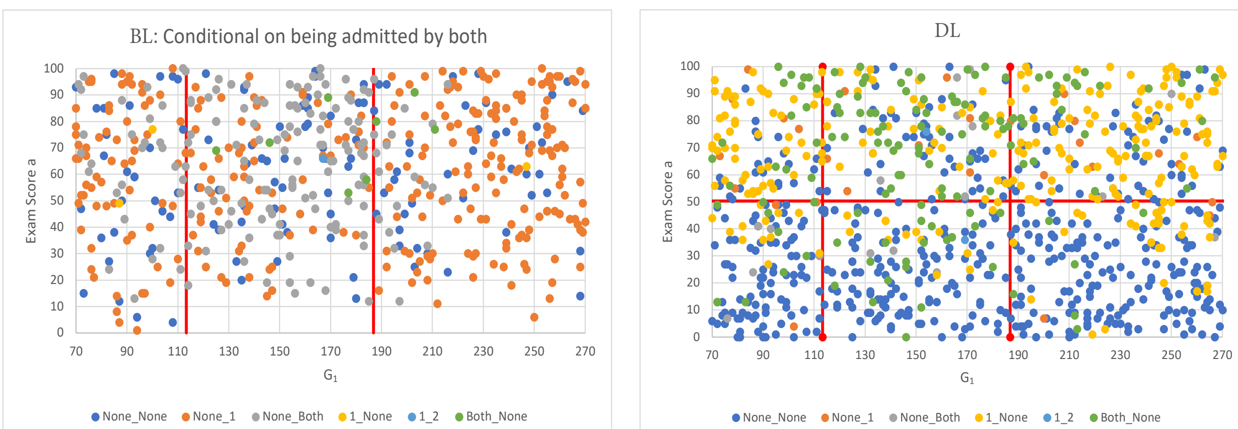
where the last inequality holds since $G(\epsilon) > \frac{1}{2}$ for any $\epsilon > 0$ by the symmetry of G. It is easy to see that $\tilde{c}(\alpha) < \hat{c}(\alpha) < \overline{c}(\alpha)$. Thus, the student will learn ϵ_i if and only if $c < \overline{c}(\alpha)$.

C Additional Figures and Tables



■ The label X_Y with $X, Y \in \{None, 1, 2, Both\}$ indicates that an individual learned the gain(s) from college(s) X in the pre-application stage and learned the gain(s) from college(s) Y in the post-admission stage.

Figure C1: High Perfectness Treatments – Outcome Comparison



■ The label X_Y with $X, Y \in \{None, 1, 2, Both\}$ indicates that an individual learned the gain(s) from college(s) X in the pre-application stage and learned the gain(s) from college(s) Y in the post-admission stage.

Figure C2: Low Perfectness Treatments – Outcome Comparison

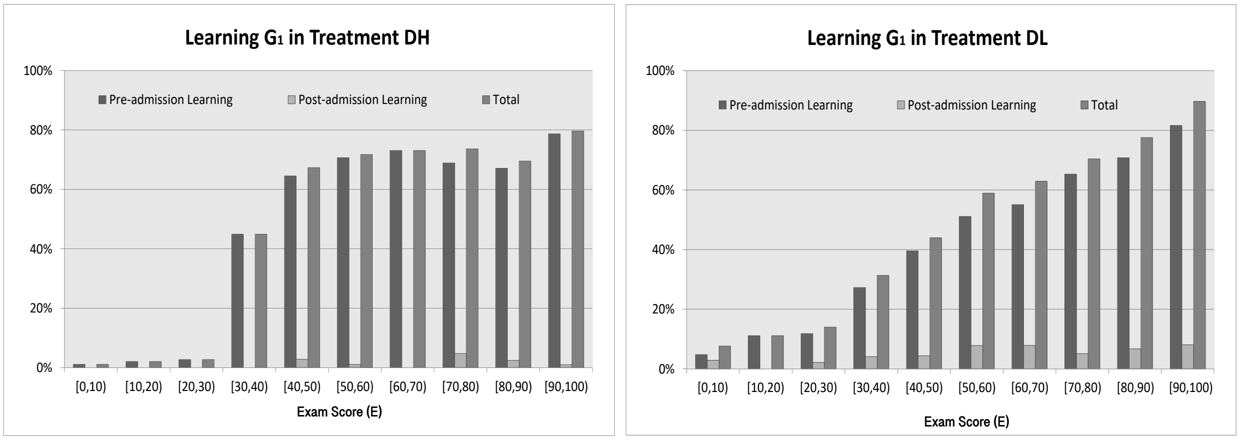


Figure C3: Learning G_1 in Treatments DH and DL - Histogram

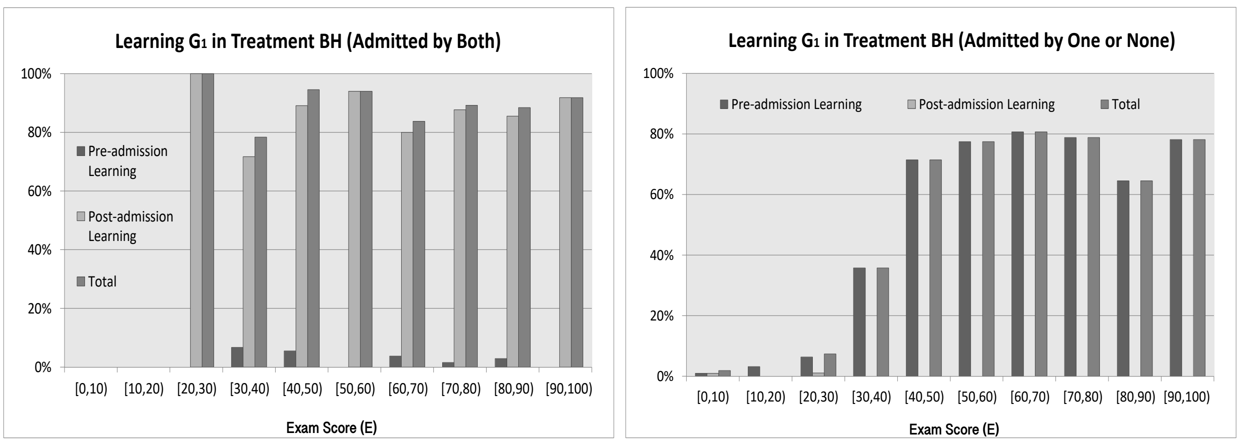


Figure C4: Learning G_1 in Treatment BH - Histogram

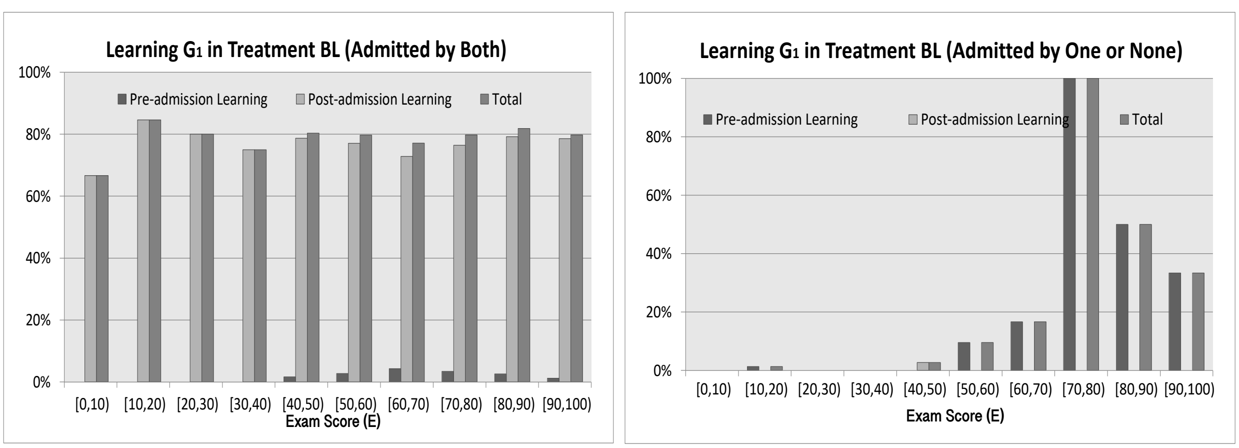
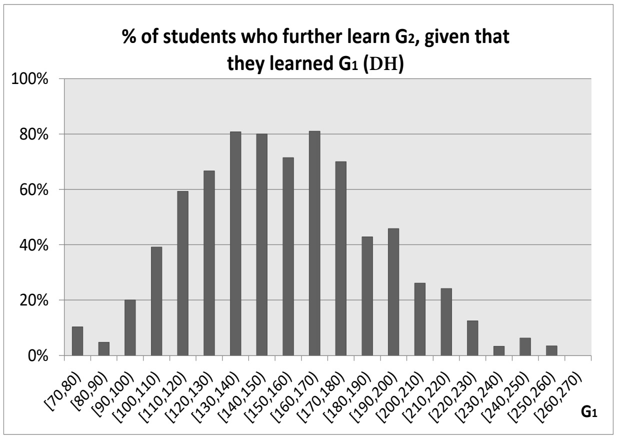


Figure C5: Learning G_1 in Treatment BL - Histogram



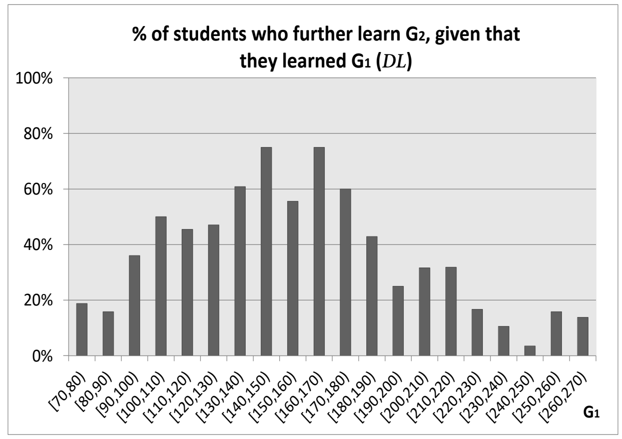
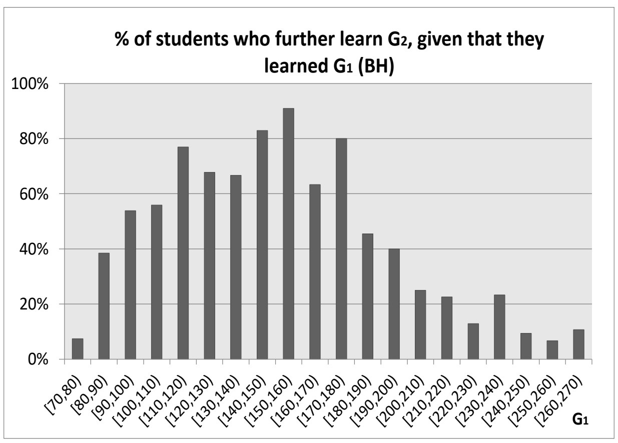


Figure C6: Learning G_2 in Treatments DH and DL - Histogram



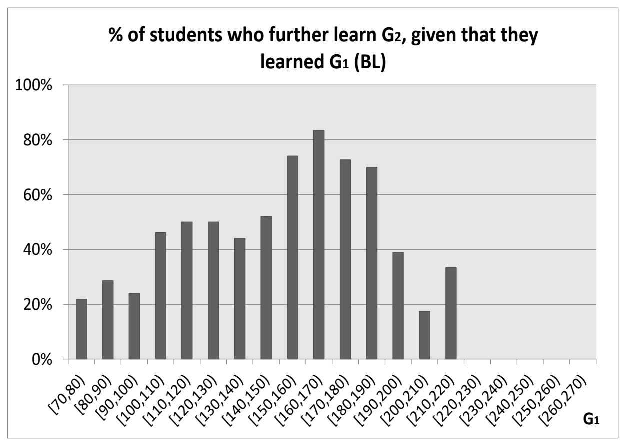
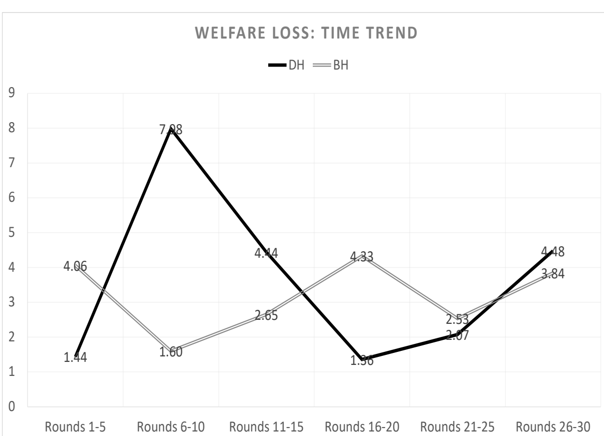


Figure C7: Learning G_2 in Treatments BH and BL - Histogram



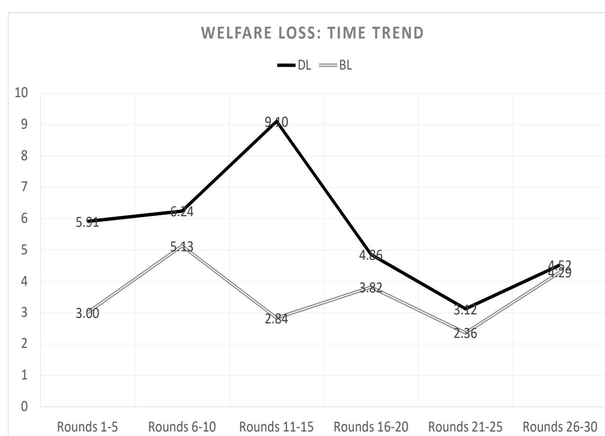


Figure C8: Welfare Loss - Time Trend

D Experimental Instructions

D.1 Treatment DH

Welcome to this experiment. Please read these instructions carefully. In the following one and a half hours or so, you will participate in 30 rounds of decision making. The payment you will receive from this experiment will depend on the decisions you make. The amount you earn will be paid **electronically via the HKUST Autopay System to the bank account you provide to the Student Information System (SIS)**. The auto-payment will be arranged by the Finance Office of HKUST, which takes about two weeks or more.

In this experiment, you are trying to enter a college. There are two colleges - College H and College K. The admission decision is based on your exam score and interview score, which will be further explained below.

Exam Score & Interview Score

At the beginning of each round, two scores will be generated for you according to the following procedure.

1. Exam Score (E): Your exam score is randomly drawn from $\{0, 1, 2, ..., 99, 100\}$.

   Each integer between 0 and 100 is equally likely to be drawn for your exam score. Then the exam score will be **announced to you**.

2. Interview Score (I): Your interview score is randomly drawn from $\{0, 1, 2, ..., 99, 100\}$.

   Each integer between 0 and 100 is equally likely to be drawn for your interview score. Your interview score will **not be revealed to you.**

   Note that the exam score and the interview score are independent with each other. That is, having higher or lower exam score E does not tell you anything about your interview score I.

3. Total Score (T): Your total score is calculated as follows.

$$\text{Total Score } (T) = 90\% \times E + 10\% \times I$$

Admission Procedure in Each College

After your exam score (E) is revealed to you, you (without knowing the total score) first need to decide if you want to send your application to College H, College K, or both colleges. There is no application fee at all.

Your application with your total score (T) is automatically sent to all college(s) you applied. Among the college(s) you applied, you will be admitted by a college if your **total score** (T) is higher than the following **admission cutoff**:

	College H	College K
Admission Cutoff (Total Score)	35	23

You will then receive admission(s) from none, one, or both of Colleges H and K. If you are admitted by only one college, you need to decide whether to pursue it or not. If you are admitted by both colleges, you need to decide whether to pursue none of them, College H, or College K.

Your Gain From College

Your gain from a college depends on how well the college suits you. More precisely,

- Your gain G_H (in tokens) from College H is randomly chosen between

$$\{\mathbf{70}, \mathbf{71}, \mathbf{72}, ..., \mathbf{269}, \mathbf{270}\}.$$

  Each integer between 70 and 270 is equally likely to be drawn for your G_H. The gain G_H becomes part of your earning if College H admits you and you decide to pursue it.

- Your gain G_K (in tokens) from College K is randomly chosen between

$$\{\mathbf{50}, \mathbf{51}, \mathbf{52}, ..., \mathbf{249}, \mathbf{250}\}.$$

  Each integer between 50 and 250 is equally likely to be drawn for your G_K. The gain G_K becomes part of your earning if College K admits you and you decide to pursue it.

Note that the gain G is college specific. That is, knowing G_H does not reveal anything about G_K, and vice versa.

Your Learning Decisions

G_H and G_K are **unknown** to you at the beginning of each round, but you will have opportunities to learn them. Learning incurs some costs to you.

Once each round begins, your decision screen always contains a panel that allows you to learn what the exact gain from College H (i.e., the value of G_H) is. The panel is randomly located either in the right half or in the left half of the screen. If you decide to learn it, you need to pay 10 tokens at the end of the round. Then you further decide whether to learn what the exact gain from College K (i.e., the value of G_K) is. If you decide to learn it, you need to pay additional 10 tokens at the end of the round.

Note that the options for you to learn G_H and G_K are **always** available in your decision screen and thus the following decisions are completely up to you:

1. **whether** to learn none/one/both of G_H and G_K and
2. **when** to learn them. You can learn none/one/both of them **before or after** the admission process begins or admission result is announced to you.

The learning cost you pay is constant at 10 tokens per college and does not depend on when you learn G_H and/or G_K.

Your Earnings

Your earning in each round will be

$$= \begin{cases} \text{Your Gain from Gollege} - \text{Cost of Learning You Paid}, & \text{if you pursue a college,} \\ \text{Default Gain (50 tokens)} - \text{Cost of Learning You Paid}, & \text{otherwise.} \end{cases}$$

For example,

1. Suppose that you paid 10 tokens and learned that $G_H = 150$, but decided to not learn G_K. You were admitted by College H and decided to pursue it. Your earning is 150 (Gain) - 10 (Cost of Learning) = 140.
2. Suppose that you paid 10 tokens and learned that $G_H = 150$. Then you further paid 10 tokens and learned that $G_K = 170$. It turned out that you were admitted by College K and decided to pursue it. Your earning is 170 (Gain) - 20 (Cost of Learning) = 150.
3. Suppose that you paid 10 tokens and learned that $G_H = 150$, but decided to not learn G_K. It turned out that you were admitted by College K and decided to pursue it. The realized gain was $G_K = 170$. Your earning is 170 (Gain) - 10 (Cost of Learning) = 160.
4. Suppose that you paid 10 tokens and learned that $G_H = 150$. Then you further paid 10 tokens and learned that $G_K = 170$. It turned out that you were not admitted by any college. Your earning is 50 (Default Gain) - 20 (Cost of Learning) = 30.
5. Suppose that you decided to not learn G_H nor G_K. It turned out that you were admitted by College K and decided to pursue it. The realized gain was $G_K = 180$. Your earning is 180 (Gain) - 0 (Cost of Learning) = 180.

Information Feedback

At the end of each round, the computer will provide you with some feedback, including 1) your exam score, 2) your interview score, 3) your total score, 4) which college(s) you are admitted, 5) which college you pursue, 6) your learning decisions, 7) your G_H and G_K (regardless of whether you paid to learn none, one, or both of them), and 8) your earning.

Your Payment

The computer randomly selects **1 round** out of the 30 rounds to calculate your cash payment. So it is in your best interest to take each round equally seriously. Your total payment in HKD will be the number of tokens you earned in the selected round (1 token = 1 HKD) plus a HKD 40 show-up fee.

A Practice Round

To ensure your understanding of the instructions, you will participate in a practice round. The practice round is part of the instructions and is not relevant to your cash payment. Its objective is to get you familiar with the computer interface and the flow of the decisions in each round. Once the practice round is over, the computer will tell you "The official rounds begin now!"

Completion of the Experiment

After the 30th round, the experiment will be over. You will be instructed to fill in the receipt for your payment. The amount you earn will be paid electronically via the HKUST Autopay System to the bank account you provide to the Student Information System (SIS). The auto-payment will be arranged by the Finance Office of HKUST.

[Round 0]

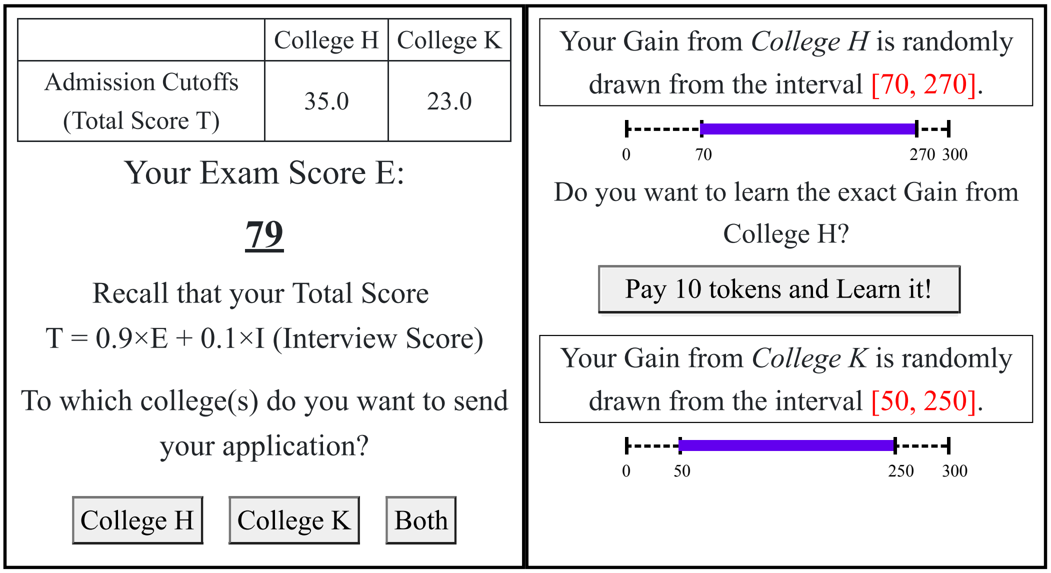


Figure D1: Screen Shot - Learning Panel on the Right

[Round 0]

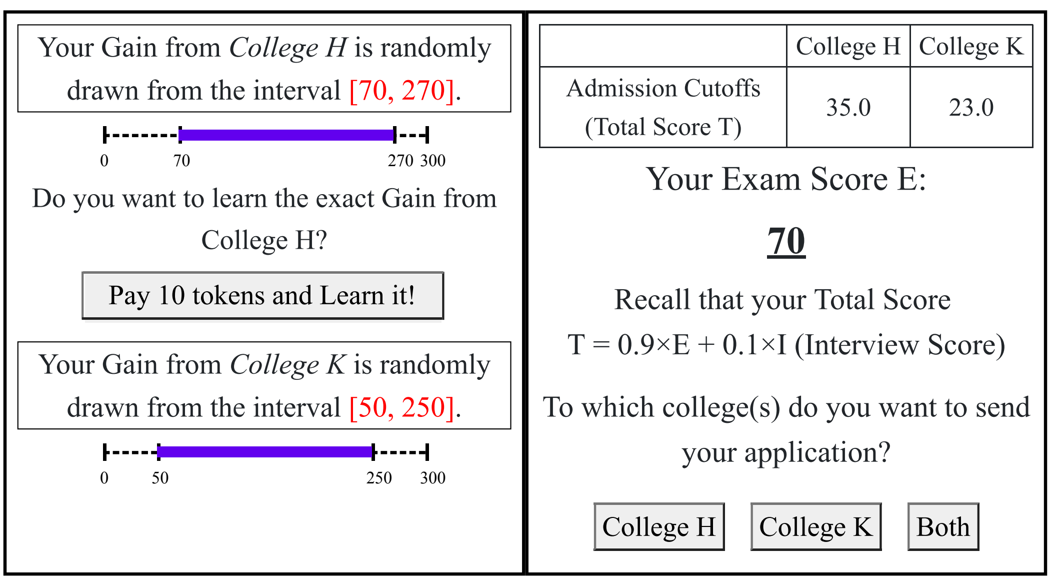


Figure D2: Screen Shot - Learning Panel on the Left

D.2 Treatment CH

Welcome to this experiment. Please read these instructions carefully. In the following one and a half hours or so, you will participate in 30 rounds of decision making. The payment you will receive from this experiment will depend on the decisions you make. The amount you earn will be paid **electronically via the HKUST Autopay System to the bank account you provide to the Student Information System (SIS)**. The auto-payment will be arranged by the Finance Office of HKUST, which takes about three weeks.

In this experiment, you are trying to enter a college. There are two colleges - College H and College K. They decide whether to admit you or not according to the following **centralized admission procedure**. First, you need to indicate your top choice between the two colleges to a central admission office. Second, the admission office makes admission decisions based on 1) the submitted rankings, and 2) your exam scores and interview scores, which will be further explained below.

Exam Score & Interview Score

At the beginning of each round, two scores will be generated for you according to the following procedure.

1. Exam Score (E): Your exam score is randomly drawn from $\{0, 1, 2, ..., 99, 100\}$.

   Each integer between 0 and 100 is equally likely to be drawn for your exam score. Then the exam score will be **announced to you**.

2. Interview Score (I): Your interview score is randomly drawn from $\{0, 1, 2, ..., 99, 100\}$.

   Each integer between 0 and 100 is equally likely to be drawn for your interview score. Your interview score will **not be revealed to you.**

   Note that the exam score and the interview score are independent with each other. That is, having higher or lower exam score E does not tell you anything about your interview score I.

3. Total Score (T): Your total score is calculated as follows.

$$\text{Total Score } (T) = 90\% \times E + 10\% \times I$$

Admission Procedure via Central Admission Office

After your exam score (E) is revealed to you, you (without knowing the total score) are asked to indicate your top choice between College H and College K as follows:

Please indicate your top choice:

College H *College K*

After you indicate your top choice, the admission procedure begins as follows:

1. The admission office sends your application to the college of your top choice.
2. The college accepts your application if your **total score** (T) is above the following **admission cutoff**, and reject otherwise:

	College H	College K
Admission Cutoff (Total Score)	36.3	23

3. If your application is accepted by the college of your top choice, the admission process is finalized.
4. Otherwise, the admission office sends your application to the college of your second choice.
5. The college decides whether to accept your application based on your total score T and the admission cutoff.
6. If your application is accepted by the college of your second choice, the admission process is finalized.
7. Otherwise, you are not admitted by any college and the process is finalized.

Note that the only thing you need to do is to indicate your top choice between College H and College K. All the steps described above take place in the admission system automatically, without any further inputs from you.

After the admission process is over, you will be informed whether you are admitted by College H, College K, or none of them. In case that you are admitted by a college, you need to decide whether to pursue the college or not.

Your Gain From College

Your gain from a college depends on how well the college suits you. More precisely,

- Your gain G_H (in tokens) from College H is randomly chosen between

$$\{\mathbf{70}, \mathbf{71}, \mathbf{72}, ..., \mathbf{269}, \mathbf{270}\}.$$

  Each integer between 70 and 270 is equally likely to be drawn for your G_H. The gain G_H becomes part of your earning if College H admits you and you decide to pursue it.

- Your gain G_K (in tokens) from College K is randomly chosen between

$$\{\mathbf{50}, \mathbf{51}, \mathbf{52}, ..., \mathbf{249}, \mathbf{250}\}.$$

  Each integer between 50 and 250 is equally likely to be drawn for your G_K. The gain G_K becomes part of your earning if College K admits you and you decide to pursue it.

Note that the gain G is college specific. That is, knowing G_H does not reveal anything about G_K, and vice versa.

Your Learning Decisions

G_H and G_K are **unknown** to you at the beginning of each round, but you will have opportunities to learn them. Learning incurs some costs to you.

Once each round begins, your decision screen always contains a panel that allows you to learn what the exact gain from College H (i.e., the value of G_H) is. The panel is randomly located either in the left half or right half of the screen. If you decide to learn it, you need to pay 10 tokens at the end of the round. Then you further decide whether to learn what the exact gain from College K (i.e., the value of G_K) is. If you decide to learn it, you need to pay additional 10 tokens at the end of the round.

Note that the options for you to learn G_H and G_K are **always** available in your decision screen and thus the following decisions are completely up to you:

1. **whether** to learn none/one/both of G_H and G_K and
2. **when** to learn them. You can learn none/one/both of them **before or after** the admission process begins or the admission result is announced to you.

The learning cost you pay is constant at 10 tokens per college and does not depend on when you learn G_H and/or G_K.

Your Earnings

Your earning in each round will be

$$= \begin{cases} \text{Your Gain from Gollege} - \text{Cost of Learning You Paid}, & \text{if you pursue a college,} \\ \text{Default Gain (50 tokens)} - \text{Cost of Learning You Paid}, & \text{otherwise.} \end{cases}$$

For example,

1. Suppose that you paid 10 tokens and learned that $G_H = 150$, but decided to not learn G_K. You were admitted by College H and decided to pursue it. Your earning is 150 (Gain) - 10 (Cost of Learning) = 140.

2. Suppose that you paid 10 tokens and learned that G_H = 150. Then you further paid 10 tokens and learned that G_K = 170. It turned out that you were admitted by College K and decided to pursue it. Your earning is 170 (Gain) - 20 (Cost of Learning) = 150.

3. Suppose that you paid 10 tokens and learned that G_H = 150, but decided to not learn G_K. It turned out that you were admitted by College K and decided to pursue it. The realized gain was G_K = 170. Your earning is 170 (Gain) - 10 (Cost of Learning) = 160.

4. Suppose that you paid 10 tokens and learned that G_H = 150. Then you further paid 10 tokens and learned that G_K = 170. It turned out that you were not admitted by any college. Your earning is 50 (Default Gain) - 20 (Cost of Learning) = 30.

5. Suppose that you decided to not learn G_H nor G_K. It turned out that you were admitted by College K and decided to pursue it. The realized gain was G_K = 180. Your earning is 180 (Gain) - 0 (Cost of Learning) = 180.

Information Feedback

At the end of each round, the computer will provide you with some feedback, including 1) your exam score, 2) your interview score, 3) your total score, 4) your top choice school, 5) which college you are admitted, 6) which college you pursue, 7) your learning decisions, 8) your G_H and G_K (regardless of whether you paid to learn none, one, or both of them), and 9) your earning.

Your Payment

The computer randomly selects **1 round** out of the 30 rounds to calculate your cash payment. So it is in your best interest to take each round equally seriously. Your total payment in HKD will be the number of tokens you earned in the selected round (1 token = 1 HKD) plus a HKD 40 show-up fee.

A Practice Round

To ensure your understanding of the instructions, you will participate in a practice round. The practice round is part of the instructions and is not relevant to your cash payment. Its objective is to get you familiar with the computer interface and the flow of the decisions in each round. Once the practice round is over, the computer will tell you “The official rounds begin now!”

Completion of the Experiment

After the 30th round, the experiment will be over. You will be instructed to fill in the receipt for your payment. The amount you earn will be paid electronically via the HKUST Autopay System to the bank account you provide to the Student Information System (SIS). The auto-payment will be arranged by the Finance Office of HKUST.

[Round 0]

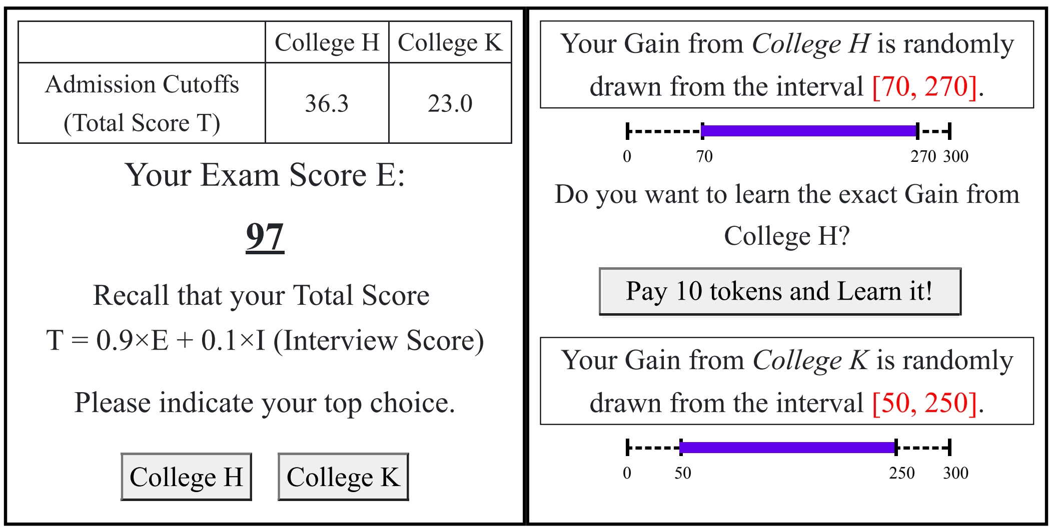


Figure D3: Screen Shot - Learning Panel on the Right

[Round 0]

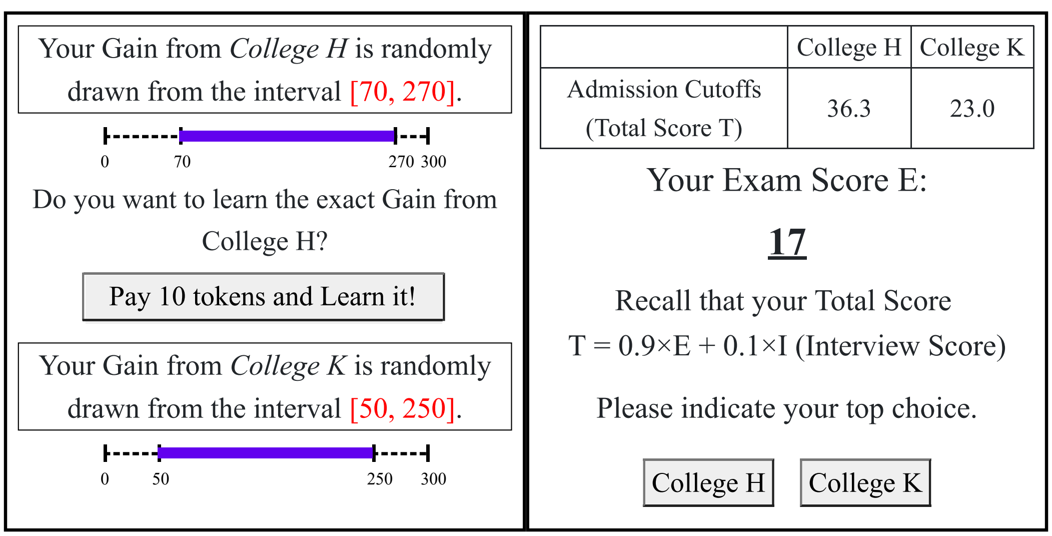


Figure D4: Screen Shot - Learning Panel on the Left