Costly information acquisition in centralized matching markets*

Rustamdjan Hakimov† Dorothea Kübler‡ Siqi Pan§

July 19, 2021

Abstract

Every year during school and college admissions, students and their parents devote considerable time and effort to acquiring costly information about their own preferences. In a market where students are ranked by universities based on exam scores, we explore ways to reduce wasteful information acquisition–that is, to help students avoid acquiring information about their out-of-reach schools or universities–using a market design approach. We find that, both theoretically and experimentally, a sequential serial dictatorship mechanism leads to less wasteful information acquisition and higher student welfare than a direct serial dictatorship mechanism. This is because the sequential mechanism informs students about which universities are willing to admit them, thereby directing their search. Additionally, our experiments show that the sequential mechanism has behavioral advantages because subjects deviate from the optimal search strategy less frequently under the sequential than under the direct mechanism. We also investigate the effects of providing historical cutoff scores under the direct mechanism. We find that the cutoff provision can increase student welfare, especially when the information costs are high, although the effect is weaker than that of a sequential mechanism.

Keywords: Matching market, serial dictatorship, information acquisition, lab experiment
JEL classification: C92, D47

---

*Our special thanks go to Nina Bonge, who helped us to program the experiments. We thank Inacio Bò, Jacob Leshno, Ignacio Rios, Renke Schmaker, Ran Shorrer, and the participants of the market design seminar in the University of Lausanne, the 17th Matching-in-Practice workshop, the ESA 2021 Global Around-the-Clock Virtual Conference, the 20th Annual SAET Conference, the Social, Behavioral and Experimental Economics (SBEE) seminar at the University of Michigan, and the seminars at Wirtschaftsuniversität Wien, Queensland University of Technology, and Monash University for their valuable comments. Financial support from the DFG through CRC TRR 190 (Kübler), the Leibniz SAW project MADEP (Kübler), the University of Melbourne (Pan), and the Swiss National Science Foundation project #100018_189152 (Hakimov) is gratefully acknowledged.

†University of Lausanne & WZB Berlin Social Science Center, Internef 536, Quartier de Chamberonne, CH-1015, Lausanne, Switzerland; email: rustamdjan.hakimov@unil.ch

‡WZB Berlin Social Science Center & Technical University Berlin, Reichpietschufer 50, 10785 Berlin, Germany; email: dorothea.kuebler@wzb.eu

§Department of Economics, The University of Melbourne, VIC 3010, Australia; email: siqi.pan@unimelb.edu.au

1 Introduction

Every year, many students and parents devote considerable time and effort to screening universities and study programs. These activities include searching university websites and brochures as well as talking to current students, alumni, and counsellors. University rankings and independent guidebooks are consulted to collect information. Finally, campus visits are scheduled, often involving costly and time-consuming trips. Parents and students gather information regarding academic quality and the programs offered by universities, in addition to costs and scholarships, the facilities and location, housing opportunities, etc. This information helps them form preferences about these universities and programs. A similar situation arises in school choice when parents and children have to decide which schools to put on their wish list of preferences.

If parents and prospective students search too little or in the wrong places, this can lead to an educational mismatch. It is often difficult to consider the right universities and to rank them properly. While search behavior has been studied in empirical work that takes the organization of the market as given, the question arises as to how admission procedures can be designed to facilitate the search process for students and direct them toward appropriate and realistic options.

Despite its importance, the need to collect information in order to form preferences in matching markets has not received much attention so far. The study of centralized and decentralized markets has mainly focused on the stability and efficiency of outcomes, and the incentive properties of the mechanisms. Regarding preferences, it is typically assumed that parents and students can rank universities at no cost. Such models are at odds with the observed activities of parents and students and the wealth of information on websites and in books and brochures. The assumption of costless preference formation has led to a strong emphasis on mechanisms that elicit the complete rank-order list from applicants. However, such mechanisms may turn out to be suboptimal when students have to first collect costly information about universities to be able to rank them.[1]

Our paper aims to study which matching mechanisms lead to higher student welfare with costly information acquisition to form preferences. As a first step, we use a simple model to derive optimal search strategies and demonstrate welfare differences between mechanisms given optimal search. As a second step, we empirically evaluate whether the mechanisms that perform better according to the theory are also superior in a laboratory experiment. We conduct this second step because in our setting as well as in real life, search strategy can be complicated and its complexity may vary in different mechanisms. The experimental method allows us to investigate, from a behavioral perspective, whether people can conduct optimal search with different mechanisms, and thus which mechanism results in the highest welfare empirically.

[1] It could be argued that the cost of information acquisition is low compared to the benefit of choosing a suitable study program. However, when the number of programs is high, gaining full knowledge of all details of all programs is impossible or very costly in terms of time and effort. In this case, students have to decide about which programs to acquire information, which is captured by the search costs in the model. We model search as a sequential process, such that the marginal benefit of an additional step of search may not exceed its cost. In our lab experiments, we investigate environments with low and high information costs.

To avoid wasteful information acquisition, it is important for parents and prospective students to search only among universities that are within their reach–that is, universities that would accept the students in the course of the procedure. We say that such universities belong to the students' budget set. When students are too pessimistic about their chances of being selected they may not acquire information about the universities in their budget set. The budget set depends on the preferences or priorities of universities, determined, for example, by the rank in admissions exams or the school GPA, as well as the preferences and choices of other students. We focus on situations where the priorities of universities are aligned and common knowledge, as is the case in single-exam or GPA-based university admissions where all universities rank students in the same way (e.g., in China, Turkey, Denmark, Sweden, Tunisia, and Germany). Additional applications include school choice based on a single average grade , for instance in Berlin (Basteck et al., 2021); centralized labor markets, for instance for doctors at public hospitals in France;[2] and other allocations based on priority and queues, for instance office allocations in new buildings (Baccara et al., 2012).

We model the formation of preferences as a costly process of information acquisition. Our search technology is motivated by centralized university admission systems that rely on ordinal rankings of the universities by the students, and allows students to learn their ordinal preferences of the universities.[3] In our model, although students may have different realized preferences over universities, their prior belief before search is that any ordinal ranking of the universities is equally likely. All universities have the same capacity.[4] The exact search technology is not crucial for our main theoretical results regarding student welfare but it must be simple enough to allow for closed-form solutions of the optimal search strategy for the experimental implementation. In the first step of the search, each student can pay a cost to compare any two universities. Then, for an additional cost, a student can choose another university to include in this ranking. Thus, to learn the ordinal ranking of m universities, a student has to pay the information cost $(m-1)$ times. Besides a uniform prior over all universities, we also consider a tiered prior structure. Specifically, all students prefer a university in a better tier to a university in a worse tier but their within-tier preference follows a uniform distribution. In this case, we assume that this search technology applies to each tier separately.[5] We assume that students can also apply to universities for which

[2]See https://www.anemf.org/wp-content/uploads/2019/07/procedure-des-choix.pdf, last accessed 4.7.2021.

[3]Ordinal preferences are enough to determine the optimal submission strategy, as we consider strategy-proof mechanisms. This would not be the case, for instance, in the Immediate Acceptance mechanism, where cardinal utilities might influence the equilibrium strategy.

[4]While the assumption of a uniform prior over all universities can be restrictive, it approximates the situation in some real-life markets. For instance, in countries where students apply to subject-specific tracks (as in most European countries, Russia, Brazil, etc.), the ranking of universities is often subject specific, and there is less vertical differentiation between universities across subjects. In the case of school choice, there is typically less information available about the quality of schools compared to the quality of university programs. Therefore, assuming a uniform prior for school rankings in the consideration set is reasonable. Regarding the assumption of equal size of universities, this can be interpreted as a situation of uncertainty where students are not informed about the number of seats before acquiring information about the universities.

[5]A tiered prior structure approximates the situation in many university admission markets. For instance, universities have an exogenously determined tier structure in China, see Chen and Kesten, 2017, where the ranking of

they have not acquired information, since there is typically no such requirement in university admissions.

In our environment where all universities rank students in the same way, we consider two mechanisms for implementing the serial dictatorship rule, namely one where students simultaneously submit their preference lists, called the Direct Serial Dictatorship (DirSD) mechanism, and one in which students sequentially select universities in the order of their priority, called the Sequential Serial Dictatorship (SeqSD) mechanism.[6]

The DirSD mechanism is praised in the matching literature (Abdulkadiroğlu and Sönmez, 1998), and variations of it are employed in Australia, China, Turkey, Greece, Denmark and Sweden for university admissions. In DirSD, all students simultaneously submit their rank-order lists to the clearinghouse. Then the serial dictatorship algorithm is run and the matching is determined. There is no opportunity to learn about the preferences and choices of other students. Students can only guess what their budget set and optimal search strategy are based on their expectations regarding others' choices. We consider DirSD as the baseline mechanism, and study two approches to improve the welfare of students relative to this baseline.

The first approach is a sequential mechanism, SeqSD, which is motivated by both theory and practice. In SeqSD, students take decisions sequentially in the order of their priority. The student with the highest priority chooses a university first, then the student with the second highest priority chooses a university among the universities that still have vacant seats, etc. Under this mechanism, students do not face any uncertainty about their budget set: when it is their turn, students observe their budget set, and can pick the most-preferred option. Importantly, students can acquire information about their preferences after they have learned their budget set in SeqSD. To our knowledge, this exact mechanism is not used for school or college admissions in practice, but some countries employ similar dynamic mechanisms in which students can observe their set of offers. For instance, France switched to a sequential university-proposing deferred-acceptance mechanism in 2018. In this mechanism, students receive offers from the universities over several weeks. Tunisia uses a three-step SD, in which the cohort is divided into three groups based on priority orders. Starting from the highest-ranked group, the three groups sequentially submit preference lists and are then assigned using the SD mechanism. After the assignment of each group, the remaining vacancies are published before the next group submit their preference lists. Germany and the Chinese province of Inner Mongolia also use dynamic mechanisms that reveal partial information about the budget set to the participants. SeqSD is also adopted to allocate faculty offices in renovated buildings in US professional schools, with lower ranked participants observing the choices of their peers with higher priority (Baccara et al., 2012). To our knowledge,

universities within tiers differs between students. Even when there is no official categorization of universities into tiers, the tiers could exist due to a general agreement regarding university quality or certain constraints such as universities in the home region being preferred to those in other regions.

[6] When all universities rank students in the same way, the outcome of the Serial Dictatorship rule is the same as that of the Top Trading Cycles rule and the Gale-Shapley Deferred Acceptance rule.

we are the first to investigate the effects of a sequential mechanism on student welfare when taking information acquisition into consideration.

The second approach is to provide historical cutoff scores under DirSD (Cutoff). When different cohorts have similar preferences over universities and similar distributions of exam scores, historical cutoffs contain useful information about the selectivity of the universities. Thus, by observing historical cutoffs, students receive information about their admission chances, that is, about their budget set. Cutoff is less time-consuming than SeqSD, especially in larger markets, because students search and choose offers one-by-one under SeqSD, but submit preferences simultaneously under Cutoff. However, given that SeqSD provides accurate information about the students' budget set, cutoffs can be less effective than SeqSD in reducing wasteful information acquisition if the distributions of preferences and exam scores fluctuate between years. For instance, Ajayi and Sidibe (2020) show that the correlation between the school cutoffs in 2007 and 2008 in Ghana was 0.84 for all schools, and as low as 0.37 for less selective schools, which may be driven by frequent changes in the mechanisms. Information about cutoffs is widely used in practice (see Immorlica et al. (2020) for an overview of college admission systems where cutoffs are published). However, to our knowledge there is no empirical evidence yet about the causal effects of historical cutoff provision on student welfare, and we aim to close this gap.

Based on a model of information acquisition and in line with the intuition described above, we show that student welfare under SeqSD is always higher than or equal to welfare under DirSD. The model also allows us to derive exact predictions concerning the optimal search behavior and submission strategies under potentially incomplete preference information. Regarding the provision of historical cutoff scores, student welfare is predicted to be higher than or equal to welfare under DirSD, as historical cutoffs contain information that can help students determine their budget set. However, welfare in treatment Cutoff is predicted to be lower than or equal to welfare in SeqSD, as the information contained in the cutoffs may not be accurate due to differences between cohorts, such as fluctuations in the distribution of scores in our experiments. These predictions regarding student welfare do not depend on a particular search technology. They also hold true for every student in the market regardless of her priority.

The theoretical results show that both the sequential mechanism and the cutoffs improve student welfare. However, the benefits of the two alternatives may not be fully captured by the theory. To derive the optimal search strategy in the direct mechanism, a student needs to form correct beliefs about the probability of each possible composition of her budget set and, based on these beliefs, to weigh the benefit of an additional search against its cost. This is a rather complicated problem to solve, and therefore inexperienced participants in real markets may deviate substantially from the optimal search. The sequential mechanism significantly mitigates the search complexity by precisely informing students about their budget set. Thus, it might have additional behavioral benefits over the direct mechanism, which the theory does not capture.[7] Cutoffs also

[7]Note that SeqSD is behaviorally simpler than DirSD for the students, even when we assume students have

provide useful information about the budget set to students, which might simplify search. However, the welfare gain from cutoffs relies on participants holding equilibrium beliefs about the search strategies of previous generations. This may also be challenging for students and thus an empirical test is crucial. Because search costs and student preferences are usually not observable in real-life, it is difficult to identify the optimality of search strategies from field data. Therefore, a laboratory experiment might be the best approach to address our research questions.

We design experiments to compare DirSD, SeqSD, and Cutoff in a centralized university admissions experiment where we vary the monetary cost of information acquisition and the prior of students about the quality of universities (a uniform prior or two tiers of universities). First, as predicted by the theory, we find that student welfare is highest under SeqSD, second-highest under Cutoff, and lowest under DirSD, with all differences being significant. The improvement of SeqSD relative to DirSD is driven by higher payoffs from the resulting matching and lower information acquisition costs. Cutoff and DirSD lead to similar average payoffs from the matching, but the information acquisition costs are significantly lower under Cutoff especially when information costs are high. With respect to optimal search strategies, we observe significant deviations in both directions (over- and under-search) in both DirSD and SeqSD. However, the deviations from the optimal search are significantly less frequent in SeqSD. Thus, SeqSD leads to higher welfare not only due to the provision of the budget set but also due to fewer strategic mistakes in the search strategies. As for Cutoff, we do not have point predictions, but we observe that participants avoid information acquisition for universities with cutoff scores higher than their score, especially when the search cost is high. Moreover, compared to DirSD and Cutoff, subjects under SeqSD make fewer strategic mistakes in submission decisions given the search. Thus, in addition to the theoretically predicted benefits of SeqSD, it has additional behavioral benefits because it is easier for participants to follow the optimal search and submission strategies under SeqSD than under DirSD. We believe that the behavioral benefits of SeqSD can be expected to be even more important with a more complex prior structure, which makes it more challenging for a student in DirSD to form beliefs about her own budget set when making search decisions. However, such belief formation is not necessary in the sequential mechanism where the budget set is known and therefore independent of the priors.

Related Literature

Our paper contributes to the recent literature on information acquisition in matching markets, which includes theoretical studies such as Bade (2015), Immorlica et al. (2020), Chen and He (2019), and Artemov (2019), and experimental studies like Chen and He (2018).

Bade (2015) studies the house allocation problem and shows that when information acquisition about one's own preferences is costly, serial dictatorship is the unique ex-ante Pareto optimal mechanism among all strategy-proof and nonbossy mechanisms. The paper allows for multiple levels of information acquisition through partitions of the state space but does not explicitly model

complete knowledge of their own preferences, because it is obviously strategy-proof (Li, 2017). We study whether the behavioral benefits of SeqSD extend to the search strategies.

the search process.

Immorlica et al. (2020) show that it is beneficial for students to know their budget set or set of feasible options before acquiring costly information about universities. This creates incentives to wait until the market has resolved before searching for information and accepting an offer. As a result, information deadlocks arise when there is a cycle of students in which each student's information acquisition decisions depend on the demand of others. The analysis of Immorlica et al. (2020) suggests that facilitating efficient price discovery by publishing cutoffs can improve student welfare. We find empirical support for this theoretical result. Additionally, we provide causal empirical evidence of the difference between direct and sequential mechanisms with regard to information acquisition.

Chen and He (2018, 2019) compare students' incentives to acquire information under the immediate acceptance mechanism and the student-proposing deferred acceptance mechanism. In both mechanisms, students have to submit rank-order lists upfront and do not receive information about their budget set. In their theoretical contribution, Chen and He (2019) show that only the immediate acceptance mechanism incentivizes students to learn their own cardinal and the other participants' preferences. In experiments, Chen and He (2018) find that overall, the willingness-to-pay for information is too high across treatments, lowering aggregate welfare. In contrast, we compare direct and sequential serial dictatorship mechanisms and study the effects of cutoff provision with respect to information acquisition and student welfare. Also, Chen and He (2018, 2019) model information acquisition as a binary choice: acquiring zero or full information. We develop a sequential search model in which agents can choose various stopping points. This captures search processes in many real-life scenarios, and it also provides us with rich data on search patterns.

Artemov (2019) finds that informational incentives provided by a random serial dictatorship mechanism fall short of the social optimum in most cases, and he proposes policies to improve information acquisition. Noda (2020) investigates the optimal disclosure policy regarding the choice set–that is, the set of objects available, under a random serial dictatorship mechanism. The paper concludes that the full-disclosure policy is typically Pareto inefficient due to the presence of a positive externality in information acquisition. Similar to Chen and He (2018, 2019), Artemov (2019) and Noda (2020) also simplify the search process to a binary choice of acquiring zero or full information. Harless and Manjunath (2018) consider a setup where information acquisition is costless but each agent can choose to learn his value for only one object. They show that the top trading cycles rule outperforms serial dictatorship in terms of fairness, though the allocation might not be Pareto efficient. Bó and Ko (2020) consider colleges' incentives to acquire information about the quality of applicants. They show that when screening costs are low, all schools acquire more information about applicants, but this does not improve the quality of the admitted pool for the lower-ranked colleges, as the best students are more likely to be assigned to better colleges.

Recent empirical work on school choice (Narita, 2018) and university admissions (Grenet et al., 2019) provides indirect evidence of students searching and learning about their preferences during the application process. Narita (2018) studies re-application behavior for high schools in New York City, and documents that a considerable proportion of students who have received their first choice decide to re-apply. Such changes in demand create a welfare loss if they cannot be accommodated by the market. Similarly, Grenet et al. (2019) document that university programs whose offers are received earlier are more likely to be ranked higher than programs whose offers arrive later. This can be explained with students' costly search regarding the programs.

Cutoff scores have been studied by Azevedo and Leshno (2016) in a two-sided matching framework with demand and supply. They show that cutoff scores can be interpreted as prices, such that at any vector of cutoffs that equates supply and demand, the demand function yields a stable matching. The cutoffs are also used in studies employing a regression-discontinuity design to estimate the causal impact of university or school attendance (see, for instance, Abdulkadiroğlu et al., 2014; Hastings et al., 2013; Zimmerman, 2019; Luflade, 2019). Ajayi et al., 2020 run a field experiment providing participants in school choice in Ghana with extensive information, including information on admission chances. The information provision changes application behavior, but it is hard to conclude which part of the information intervention drives the effect. To our knowledge, our paper is the first to empirically investigate the effect of providing historical cutoffs on search and market outcomes.

Our paper relates to the recent literature on dynamic mechanisms with known preferences. Li (2017) compares SeqSD and DirSD in the lab and finds that a significantly higher proportion of participants use the optimal submission strategies in SeqSD than in DirSD. Klijn et al. (2019) and Bó and Hakimov (2020) present similar results for the comparison of sequential and direct versions of the deferred acceptance mechanism, and Bó and Hakimov (2020) for the top trading cycles mechanism. Echenique et al. (2016) investigate the outcomes of the dynamic deferred acceptance mechanism in laboratory experiments in a two-sided setup. Moreover, several papers analyze dynamic mechanisms used in practice where offers, acceptances, and information can be spread out over time. Other related experiments on matching markets are surveyed in Hakimov and Kübler (2020) and Pan (2020). Bó and Hakimov (2019) and Gong and Liang (2016) analyze university admissions mechanisms used in Brazil and Inner Mongolia, respectively. Both mechanisms are sequential and include the provision of intermediate cutoff scores to students. Dur et al. (2017) analyze the submission mechanism used in the Wake County Public School System where parents can wait and observe how many others have applied to certain schools in order to gauge their chances of getting a seat. Luflade (2019) studies the university admissions in Tunisia, where the SD mechanism is implemented in three sequential stages, and documents that the sequential implementation has a positive effect on the students' welfare.

A set of papers (Das and Li, 2014; Kadam, 2015; Lee and Schwarz, 2017) analyzes interviews through which agents acquire costly information about their preferences in decentralized matching markets. Similar to the model of Weitzman (1979), these models assume that a school or firm must interview an applicant before making an offer. The search and matching literature also explores the role of costly search in decentralized matching markets. For example, Shimer and Smith (2000), Atakan (2006), and Eeckhout and Kircher (2010) focus on sequential and directed search while Chade and Smith (2006) and Shorrer (2019) investigate simultaneous search. From a different perspective, Rastegari et al. (2013) and Drummond and Boutilier (2014) consider eliciting information about agents' preferences using a minimal number of interviews, just enough to ensure that a stable match is found. They show that this problem is computationally intractable in general, but provide solutions for specific prior structures or approximate stability.

A large part of the theoretical work on matching markets assumes that applicants know their priority at schools or universities, based, for example, on grades or entrance exams. This assumption is relaxed when studying the consequences of not publicizing the results of entrance exams before students have to submit their rank-order lists (Lien et al., 2016, Pan, 2019). In our study, we provide full information about the priorities of students at universities but vary information about the preferences and behavior of others.

2 Theoretical analysis

In this section, we present a model to analyze the strategies and welfare of students under the Direct Serial Dictatorship (DirSD) and Sequential Serial Dictatorship (SeqSD) mechanisms. First, we modify the standard school choice problem Abdulkadiroğlu and Sönmez (2003) to allow students to acquire information about their own preferences before and during the matching process.[8] Next, we provide a detailed description of the DirSD and SeqSD mechanisms. We then discuss and compare these two mechanisms in terms of information acquisition, preference submission, and student welfare. Lastly, as an extension of our main theorem, we discuss the effect of providing additional information under DirSD and generate predictions for our Cutoff treatment.

2.1 A university admissions problem

Students want to be assigned a seat at one of the universities. Each student has strict preferences over all universities and each university has strict priorities over all students. There is a maximum capacity at each university, but the total number of seats exceeds the total number of students. We consider an environment in which every university's priority ordering over students is determined by exam rankings. Formally, the university admissions problem consists of:

1. A set of students $I = \{i_1, \ldots, i_n\}$, $n \geq 2$.

2. A set of universities $C = \{c_1, \ldots, c_m\}$, $m \geq 2$.

[8]The school and college admissions markets investigated in this paper are more similar to the school choice problem than to the college admissions problem (Gale and Shapley, 1962) in the matching literature because universities are not strategic and we focus on students when conducting the welfare analysis.

3. A capacity vector $q = (q_1, \dots, q_m)$.

4. A vector of students' ranks $r = (r_1, \dots, r_n)$ in an exam, where r_i denotes the rank of student i among all students (with 1 being the highest rank). The ranking determines their priority ordering at every university.

5. A list of strict student preferences $\succ_I = (\succ_{i_1}, \dots, \succ_{i_n})$. The preference relation $\succ_i$ of student i is a linear order over C, where $c \succ_i c'$ means that student i strictly prefers university c to university c'. Students prefer any university to remaining unmatched.[9]

6. For each student $i \in I$, a set of cardinal utilities $u_i = \{u_i^1, ..., u_i^m\}$ associated with her ordinal preferences: student i receives u_i^j when assigned to a university ranked jth in her preference relation $\succ_i$. For any $1 \leq j < j' \leq m$, $u_i^j > u_i^{j'}$.[10]

Let Ω be the set of all linear orders over C. The preference relation $\succ_i$ of student i is randomly and independently drawn from Ω following a prior probability distribution. The priors of all students are common knowledge to the entire market. Via costly information acquisition, student i can learn more about the realization of her own preferences $\succ_i$ but not the realization of other students' preferences $\succ_{i'}, i' \neq i$. The information acquired by each student is her private information. It is common knowledge that every student knows her own rank in the exam. We assume all market participants are risk neutral.

We define the *budget set* B_i as the set of all universities available to student i. That is, student i can be assigned to university $c \in B_i$ if she so desires, and cannot be assigned to any university in the complement set $C \setminus B_i$.[11] In Section 2.4, we will discuss how the budget set of a student is determined under each mechanism.

2.2 Mechanisms

Direct serial dictatorship mechanism (DirSD)

Every student simultaneously submits her rank-order list of universities. DirSD considers students in the order of their exam ranking.

[9] The assumptions that the total number of seats exceeds the total number of students, that students prefer any university to remaining unmatched, and that universities prefer any student to leaving a seat empty are made to simplify the exposition. They are not necessary for our discussion and can be relaxed easily.

[10] Our main theorem regarding student welfare allows for alternative ways of modeling cardinal utilities. This particular approach corresponds to our experimental setting. In the model, the cardinal utilities $u_i = \{u_i^1, ..., u_i^m\}$ are fixed and known in advance. Given that these utilities are associated with student i's ordinal preference $\succ_i$, the student only needs to acquire information about how she ranks the universities. For example, student i knows that the university of her first preference gives her a utility of u_i^1. She can discover which university is her first preference by acquiring information about her ranking of the universities $\succ_i$. Similar settings are used in other studies such as Coles and Shorrer (2014).

[11] We borrowed this terminology from Immorlica et al. (2020).

Step 1: The student who is ranked first in the exam (with the highest score) is assigned a seat at the first choice on her submitted list.

In general, Step κ ($\kappa \geq 2$) can be described as follows.

Step κ: The student ranked κth is assigned a seat at her best choice that still has vacant seats.

The procedure terminates when every student has been considered. Students can acquire information about their own preferences before submitting their rank-order lists.

Sequential serial dictatorship mechanism (SeqSD)

Under SeqSD, students sequentially select universities in the order of their exam ranking.

Step 1: The student who is ranked first in the exam (with the highest score) selects one university out of all universities. She is assigned a seat at this university.

In general, Step κ ($\kappa \geq 2$) can be described as follows.

Step κ: The student ranked κth selects one university out of all the universities that still have vacant seats. She is assigned a seat at this university.

The procedure terminates when every student has been considered. Students can acquire information about their own preferences before and after it is their turn to select universities.[12]

We use "preference submission" or simply "submission" to refer to the students' interaction with a mechanism–that is, submitting a rank-order list under DirSD and picking a university under SeqSD.

2.3 Information acquisition

We assume that each student i's realized preference relation $\succ_i$ is equally likely to be any linear order in Ω. In other words, all universities are equally likely to be of any rank in $\succ_i$. These uniform priors can be interpreted as the initial state in which no one has acquired any information about any university. In Appendix A.5, we introduce a tiered prior structure that allows for a common and a private factor in students' preferences and show that our main results with uniform priors can be generalized to the case of tiered priors.

Each student i, with information cost $k_i > 0$, can acquire information about her own preferences using the following search technology:

[12] We adopted a slightly different description of SeqSD in our experiments to facilitate understanding. We framed it as students sequentially receiving offers from universities that still have vacant seats and being asked to accept one offer.

Step 1: For a cost of k_i, the student can choose any two universities and learn which of these two universities is ranked higher in her own preference relation. Thus, for a cost of k_i she can learn the relative ordering of two universities.

Step 2: For an additional cost of k_i, the student can choose a third university and learn how it compares to the two universities previously investigated. Thus, for a total cost of $2k_i$ she can learn the relative ordering of three universities.

...

Step $(m-1)$**:** Finally, for an additional cost of k_i, the student can learn how the m-th university is compared to the $(m-1)$ universities investigated previously. Thus, for a total cost of $(m-1)\,k_i$ she can obtain full knowledge of her own preferences.

Students can choose to stop at any step in the above process. We use this search technology to model a student who investigates universities one by one.[13] After investigating each additional university, she knows how to compare it to all the universities she has previously investigated. Before the last step, she can only learn the relative ranking of the universities investigated, but not their exact ranks among all universities. In our setting, the investigation of only one university does not carry any information because the cardinal utility a student receives from a university is determined by the rank of that university in her preferences. Therefore, we start the process by having each student choose two universities to compare. One of these two universities can be considered as an option the student knows from the start, for instance a local university, and every investment in information informs her of a new university. This search technology captures important features of search in many real-life scenarios while being relatively simple and easy to understand for participants in the experiment.

Suppose student i stops searching at step α ($\alpha = 1, 2, \ldots, m-1$), and the set of universities she has chosen to search is given by C^S ($C^S \subseteq C$). This implies that for a cost of αk_i, the student learns the relative ranking of the $(\alpha+1)$ universities in C^S, denoted as $\succ_i^S$. She can then eliminate the possibility of all linear orders in Ω that are not consistent with this ranking, and redistribute the probability uniformly among the remaining rankings.[14]

With uniform priors, student i has the same expected utility for each university before searching, which is given by

$$V_i(0) = \frac{1}{m}\sum_{j=1}^{m} u_i^j.$$

[13] An obvious choice of a sequential search technology is the Pandora's box framework that is due to Weitzman (1979), in which a student directly observes her cardinal utilities when inspecting universities. However, this framework is not suitable for our research question because as shown in Doval (2018), the Pandora's box problem is not generally tractable without the assumption that students can only be assigned to universities they inspect. This assumption rarely holds in real-life college admissions, and centralized matching mechanisms like DirSD and SeqSD can assign students to a university independent of whether they have acquired information about it or not.

[14] Formally, we say a linear order $\omega \in \Omega$ is *consistent* with $\succ_i^S$ if $c \succ_i^S c'$ implies $c\omega c'$, $\forall c, c' \in C^S$.

After conducting α steps of search, her updated expected utility for the university that is relatively ranked γth in C^S according to $\succ_i^S$ ($\gamma = 1, \ldots, \alpha+1$) is given by

$$V_i^\gamma(\alpha) = \sum_{j=1}^{m} f^\gamma(j,\alpha)\, u_i^j,$$

in which

$$\begin{aligned} f^\gamma(j,\alpha) &= \frac{\binom{j-1}{\gamma-1}\binom{m-\alpha-1}{j-\gamma}(j-\gamma)! \times \binom{m-j}{\alpha+1-\gamma}(m-j-\alpha-1+\gamma)!}{\binom{m}{\alpha+1}(m-\alpha-1)!} \\ &= \frac{\binom{j-1}{\gamma-1}\binom{m-j}{\alpha-\gamma+1}}{\binom{m}{\alpha+1}} \end{aligned}$$

calculates the probability that the γth-ranked university in $\succ_i^S$ is ranked jth in $\succ_i$. Intuitively, because this university is ranked γth among the $(\alpha+1)$ searched universities, if it is ranked jth in the student's complete preference ordering $\succ_i$, we can identify $(\gamma-1)$ out of the j universities ranked above it and $(\alpha+1-\gamma)$ out of the $(m-j)$ universities ranked below it in $\succ_i$. The first term of the numerator $\binom{j-1}{\gamma-1}\binom{m-\alpha-1}{j-\gamma}(j-\gamma)!$ is the number of possible orderings of the universities ranked above it and the second term $\binom{m-j}{\alpha+1-\gamma}(m-j-\alpha-1+\gamma)!$ is the number of possible orderings of the universities ranked below it. The denominator $\binom{m}{\alpha+1}(m-\alpha-1)!$ is the permutation of all universities after knowing the relative ranking of $(\alpha+1)$ of them.[15] When $\alpha = m-1$, the student has full knowledge of her own preferences, and thus $V_i^\gamma(m-1) = u_i^\gamma$. When $\alpha < m-1$, the student's expected utility for those unsearched universities in $C \setminus C^S$ remains the same as the prior $V_i(0)$. We illustrate the belief updating process using the following example.

Example 1. Consider a market with three universities $C = \{c_1, c_2, c_3\}$. The first row of Table 1 lists all six possible preference orders over C (where we omit the subscript i to refer to student i). Without acquiring any information ($\alpha = 0$), a student holds the uniform prior belief. That is, she assigns the same probability $1/6$ to each of the six linear orders, which means that the expected

[15]Note that $f^\gamma(j,\alpha) = 0$ if $j < \gamma$ or $j > m-\alpha+\gamma-1$ because there have to be at least $(\gamma-1)$ universities ranked above the γth-ranked university in $\succ_i^S$ and at least $(\alpha+1-\gamma)$ universities ranked below it.

utility is the same for all three universities:

$$\mathrm{E}[c_1|\alpha=0]=\mathrm{E}[c_2|\alpha=0]=\mathrm{E}[c_3|\alpha=0]=V(0)=\frac{1}{3}\left(u^1+u^2+u^3\right).$$

1st preference (u^1)	c_1	c_1	c_2	c_2	c_3	c_3
2nd preference (u^2)	c_2	c_3	c_1	c_3	c_1	c_2
3rd preference (u^3)	c_3	c_2	c_3	c_1	c_2	c_1
$\Pr[\succ\|\alpha=0]$	1/6	1/6	1/6	1/6	1/6	1/6
$\Pr[\succ\|\alpha=1,c_3\succ c_2]$	0	1/3	0	0	1/3	1/3
$\Pr[\succ\|\alpha=2,c_1\succ c_3\succ c_2]$	0	1	0	0	0	0

Table 1: Belief Updating Example

Suppose the student chooses to conduct the first step of searching ($\alpha=1$). Suppose she picks universities c_2 and c_3 and learns that $c_3\succ c_2$. Now she is able to eliminate all orders over C that are inconsistent with $c_3\succ c_2$ and redistributes the probability uniformly among the remaining orders (see the third row of Table 1). Her updated expectation is

$$\begin{aligned}\mathrm{E}[c_1|\alpha=1,c_3\succ c_2]=&V(0)=\frac{1}{3}\left(u^1+u^2+u^3\right),\\ \mathrm{E}[c_2|\alpha=1,c_3\succ c_2]=&V^2(1)=\sum_{j=1}^{3}f^2(j,1)\,u^j=\frac{1}{3}u^2+\frac{2}{3}u^3,\\ \mathrm{E}[c_3|\alpha=1,c_3\succ c_2]=&V^1(1)=\sum_{j=1}^{3}f^1(j,2)\,u^j=\frac{2}{3}u^1+\frac{1}{3}u^2,\end{aligned}$$

with $\mathrm{E}[c_3|\alpha=1,c_3\succ c_2]>\mathrm{E}[c_1|\alpha=1,c_3\succ c_2]>\mathrm{E}[c_2|\alpha=1,c_3\succ c_2]$. Suppose the student continues searching ($\alpha=2$) and learns that $c_1\succ c_3\succ c_2$. She can then further eliminate the possibility of any order inconsistent with $c_1\succ c_3\succ c_2$ (see the fourth row of Table 1) and obtain full knowledge of her preferences:

$$\begin{aligned}\mathrm{E}[c_1|\alpha=2,c_1\succ c_3\succ c_2]=&V^1(2)=u^1,\\ \mathrm{E}[c_2|\alpha=2,c_1\succ c_3\succ c_2]=&V^3(2)=u^3,\\ \mathrm{E}[c_3|\alpha=2,c_1\succ c_3\succ c_2]=&V^2(2)=u^2.\end{aligned}$$

The example demonstrates an important feature of the updating process. In particular, after discovering $c_3\succ c_2$ at $\alpha=1$, the student realizes that c_2 cannot be her favorite university while c_3

cannot be her least favorite. As she still holds the prior belief about the rank of c_1, she now prefers c_3 to c_1 and c_1 to c_2 in expectation. The proposition below states that this is a general property: a student always prefers the higher-ranked searched universities to the unsearched universities and prefers the unsearched universities to the lower-ranked searched universities. This feature of the students' posterior beliefs allows us to characterize the students' strategies regarding preference submission and their expected utility functions under the two mechanisms.

Proposition 1. *For any $i \in I$, there exists a threshold rank $\hat{\gamma}_i(\alpha)$ at which $V_i^{\gamma}(\alpha) > V_i(0)$ for all $\gamma \leq \hat{\gamma}_i(\alpha)$ and $V_i^{\gamma}(\alpha) \leq V_i(0)$ otherwise. Thus, there is a step-specific threshold rank that splits all searched universities into two groups. The searched universities that are ranked higher than or equally to the threshold are preferred to all unsearched universities. All unsearched universities are weakly preferred to the searched universities that are ranked lower than the threshold.*

The complete proof of Proposition 1 can be found in Appendix A.1. In Example 1, the threshold rank after the first search step is $\hat{\gamma}_i(1) = 1$, meaning that the university ranked first among those searched has a higher expected utility than the unsearched university, while the university with a rank below one has a lower expected utility than the unsearched university. Thus, the higher-ranked searched university c_3 has a higher expected utility than the unsearched university c_1, and the lower-ranked searched university c_2 has a lower expected utility than the unsearched university c_1.

2.4 Preference submission and budget set

In the following proposition, we characterize the students' preference submission strategy under DirSD and SeqSD given partial or full knowledge of their own preferences.

Proposition 2. *(1) Truth-telling is an optimal submission strategy under DirSD. That is, under DirSD it is optimal for a student to rank universities according to the expected utilities (from high to low) in her rank-order list.*

(2) Truth-telling is an optimal submission strategy under SeqSD. That is, when it is her turn to choose, it is optimal for a student to select the university with the highest expected utility.

The proof of Proposition 2 is straightforward. In each step of DirSD, the student whose turn it is is assigned to the highest-ranked university in her submitted list from those that still have vacant seats. Therefore, a student is never better off by ranking a university with a lower expected utility above one with a higher expected utility. According to Proposition 2, she would rank the unsearched universities below the $\hat{\gamma}_i(\alpha)$th-ranked searched university, but above the $(\hat{\gamma}_i(\alpha) + 1)$th-ranked searched university. In SeqSD, a student whose turn it is chooses from all universities that are available to her, so it is optimal to simply choose the university with the highest expected utility. Proposition 2 relies on the strategy-proofness of DirSD and SeqSD in environments with complete information.

Recall that student i's budget set B_i is defined as the set of all universities available to her. Under DirSD, if $c \in B_i$, i can be assigned to c unless she is assigned to another university that is ranked higher than c in her submitted rank-order list; if $c \notin B_i$, i cannot be assigned to c no matter how she ranks c in her submitted list. Under SeqSD, when i selects universities, c would be available to her if $c \in B_i$, and unavailable if $c \notin B_i$.

Student i's budget set is determined by her exam rank and the submission strategies of those who are ranked above her. For instance, consider a market with three universities $C = \{c_1, c_2, c_3\}$, each of which has two seats. Under DirSD, the budget set of the student ranked third in the exam depends on the submitted rank-order lists of the two students ranked before her. If, for example, they both place university c_3 on the top of their lists, the budget set of the student ranked third contains only c_1 and c_2. Since all rank-order lists are submitted simultaneously under DirSD, a student decides what to search based on the ex ante probability distribution of her budget set $\left\{P_i\left(\tilde{B}\right)\right\}_{\tilde{B}\subseteq C}$, that is, $P_i\left(\tilde{B}\right) = \Pr\left[B_i = \tilde{B}\right]$, $\tilde{B} \subseteq C$. This requires students to form beliefs about the submission strategies of higher-ranked students. In contrast, under SeqSD a student selects the preferred university after the higher-ranked students have made their choices. She therefore observes the realization of her budget set before she makes her search and preference submission decisions.

To simplify the analysis, we make the following assumption about the market structure and students' strategies.

Assumption 1. *(1) All universities have the same capacity.*

(2) In each step of the search process, a student is equally likely to acquire information about any of the unsearched universities.

(3) If a student did not search all universities under DirSD, she is equally likely to choose any relative order over the unsearched universities in her submitted rank-order list. If a student did not search any universities that she is asked to select under SeqSD, she is equally likely to select any one of these universities.

Assumption 1 can be considered an anonymity assumption in that universities are not labeled and are thus selected at random when they have the same expected value. Together with the uniform prior structure and Proposition 2, it implies that a student i, who by assumption cannot observe what another student $i' \neq i$ has learned about her preferences $\succ_{i'}$, believes that i' is equally likely to submit any ranking in Ω under DirSD and is equally likely to select any university in C under SeqSD. In other words, for student i the submission strategy of i' always follows a uniform distribution, regardless of the search strategy of student i'. Thus, a student does not need to consider how much information other students acquire in equilibrium when forming beliefs about the submission strategies of others.

Therefore, the assumptions of same-capacity universities and uniform priors are mainly used to simplify the derivation of the ex ante probability distribution of a student's budget set, which is needed to find the optimal search strategy under DirSD. Without these assumptions, students may also need to form correct beliefs about the search strategies of others under DirSD. Even with these simplifying assumptions, decision-making under DirSD is challenging especially for the lower-ranked students, because they have to consider the submission strategies of other students.

In appendices A.2 and A.3, we present the strategies of information acquisition given the search technology. Under SeqSD, the marginal benefit of an additional search step among the available universities decreases, and we characterize the optimal stopping point of the search process. In contrast, under DirSD the marginal benefit of an additional search step may be non-monotonic. The optimal information acquisition strategy is not necessarily unique, but it is unique for the parameters that we chose in the experiment.

2.5 Welfare comparison

Now we introduce our main theorem, which compares student welfare between the two mechanisms SeqSD and DirSD. We focus on the welfare comparison at the ex-ante stage–that is, before students acquire any information about their preferences–and assume that all students acquire information optimally and adopt the truth-telling submission strategy under both mechanisms.

Theorem 1. *Every student is weakly better off under SeqSD than under DirSD ex ante if all students acquire information optimally and adopt the truth-telling submission strategy.*

While the complete proof of Theorem 1 can be found in Appendix A.4, the key intuition comes from the difference in the amount of information students have when they make their decisions. Under DirSD, all students simultaneously submit their rank-order lists. Thus, they do not have any opportunity to identify their budget sets by learning about the choices of others. In contrast, the dynamic nature of SeqSD can provide additional information about which universities other students have chosen and which are still available in the market, thus helping students to more accurately identify their budget sets. By focusing on search within the budget set, a student can reduce wasteful information acquisition and thus be weakly better off under SeqSD.

In some real-life markets with direct mechanisms, students are provided with cutoff scores that were necessary to be accepted by programs in previous years. Such information can be helpful to determine the budget set. However, this information is often not precise enough to pin down the exact budget set for every student in the market. Statistics such as the universities' historical cutoff scores or acceptance rates represent noisy information about the budget set because the distribution of student preferences, the distribution of exam scores, and the capacities of universities may change from year to year.

When students are provided with noisy information about their budget sets under DirSD, some students' posterior beliefs may still be non-deterministic. Formally, we say that the information is noisy if student i updates the probability distribution of her budget set to $\left\{\hat{P}_i\left(\tilde{B}\right)\right\}_{\tilde{B}\subseteq C}$, with

$P_i(B_i) \leq \hat{P}_i(B_i) \leq 1$ for all $i \in I$ and $\hat{P}_i(B_i) < 1$ for some i. Therefore, Theorem 1 still applies and the welfare advantage of SeqSD over DirSD holds.[16] Moreover, the provision of information cannot decrease welfare under DirSD because students in equilibrium interpret the information correctly and best respond to it. We summarize these insights in the following corollary.

Corollary 1. *1. Even when students are provided with noisy information about their budget sets (for instance historical cutoffs) under DirSD, every student is weakly better off under SeqSD than under DirSD ex ante if all students acquire information optimally and adopt the truth-telling submission strategy.*

2. Students are not worse off under DirSD ex ante when provided with noisy information about their budget sets (for instance historical cutoffs).[17]

Further comparison between DirSD with noisy information and DirSD without noisy information relies on the specific implementation. In Section 3.5, we show that in our experimental setting, some students can be strictly better off when provided with historical cutoffs.

Importantly, the proofs of Theorem 1 and Corollary 1 do not rely on a particular search technology. Also, they hold true for any ex ante distribution over a student's budget set, which means that they apply to every student regardless of the exam rank. We show in Appendix A.5 that theoretical results generalize to a tiered prior structure that allows for a common component and a private component in students' preferences. Specifically, all students prefer a university in a better tier to a university in a worse tier but may have different preferences over the universities within each tier (their within-tier preference follows a uniform distribution). We also relax Assumption 1 to allow universities in different tiers to have different capacities. In this setting, each student only needs to consider universities in one tier as long as all students adopt the truth-telling submission strategy. Therefore, we can consider each tier of universities, together with the students who consider that tier, as an independent market, and this market is identical to the market with uniform priors.

3 Experimental Design

We conducted an experiment to test the predictions of our information acquisition model, and to compare the welfare of students under the three centralized matching procedures. The three procedures (treatments) are studied in four different environments that are characterized by the one and two-tier preferences of students' preferences and two different costs of information acquisition.

[16]The proof of this statement only requires replacing $P_i\left(\tilde{B}\right)$ with $\hat{P}_i\left(\tilde{B}\right)$ in Appendix A.4.

[17]The ex-ante stage in Corollary 1 is before students acquire any information about their preferences but after they are provided with noisy information about their budget sets.

3.1 Setup

In the experiments, 12 students competed for 12 seats at six universities. Each university had two seats. All universities ranked students based on the exam scores. The score of each student was randomly and independently drawn from a uniform distribution between 1 and 100. Students were played by experimental subjects while the universities were not strategic and their actions were simulated by the computer. Students knew their score and the rank of their score among the other students in their group.[18]

Participants received 40 AUD for the assignment to their most-preferred university, 34 AUD for the second most-preferred university, 28 AUD for the third most-preferred university, etc., and 10 AUD for the least-preferred university. At the beginning of each round, participants did not know their preferences over universities but were told that each ranking was equally likely. They had the opportunity to acquire costly information about their own ordinal preferences. The exact timing, technology, and costs of this search process varied between environments and treatments.

Each session consisted of 24 participants who were split into two groups of 12 for the entire session. We used fixed matching groups to increase the number of independent observations. Each round represented a new university admissions process for the students. In total, there were eight rounds in the experiment. At the end of the experiment, one round was drawn randomly to determine the subjects' payoff.

3.2 Treatments

We conducted three treatments between subjects by varying the centralized allocation procedure and the information provided:

In **Treatment DirSD**, the direct serial dictatorship mechanism is adopted. Participants can learn their preferences at a cost before the procedure starts, i.e., before they submit the rank-order lists to the system.

In **Treatment SeqSD**, the sequential serial dictatorship mechanism is adopted. Students can search before or after observing the set of available universities.[19]

In **Treatment Cutoff**, we provide historical cutoff scores under the direct serial dictatorship mechanism. All participants observe the cutoffs of all universities of the previous cohorts before the procedure starts. To generate the cutoffs, we used the results of the DirSD sessions. More precisely, we provided the average cutoff scores from all previous DirSD markets in the same environment. Students with the same rank in the previous sessions had the same preferences over universities as in the current session, which was explained to the participants in the current session. Note that this design choice ensures the informativeness of the cutoff scores and their relevance for

[18] Although students under DirSD and SeqSD only need to know their ranks to make decisions, we also provided them with scores because this enables us to conduct the treatment with cutoffs scores.

[19] Participants can also start searching before observing the set of available universities, and continue with the search after observing it.

the optimal search strategies (see section 3.5 for details). After learning the cutoffs, subjects can acquire information about their own preferences and then submit their rank-order list, just as in DirSD.

3.3 Environments

There were four environments in the experiments. The environments varied in two dimensions, namely the prior about the quality of the universities and the cost of information acquisition.

Dimension 1: Tiered versus non-tiered preferences. The first dimension of our environments was whether the preferences of students were tiered. That is, we varied the prior belief of students about the position of universities in their preferences, and thus the degree of correlation of preferences between students. We considered two preference structures:

Two tiers. The six universities A1, A2, A3, B1, B2, and B3 belonged to two different tiers: Universities A1, A2, and A3 belonged to tier A, and Universities B1, B2, and B3 belonged to tier B. Every student preferred a university in tier A to a university in tier B, which was common knowledge. Students could have different preferences within each tier. For each tier, the within-tier ordinal preferences of each student were independently and randomly drawn from the set of all possible orderings of the three universities in that tier. Each ordering was equally likely.

For each tier of universities, the search process was as follows:

1. For a cost of \$X, a student could pick any two universities belonging to the same tier and learn which of these two universities was ranked higher in her preference ordering. Thus, for a cost of \$X a student could learn the relative ordering of two universities within a tier.

2. For an additional cost of \$X, a student could learn how the third university from the same tier compared to the two universities that she had chosen previously. Thus, for a cost of \$2X a student could learn her full ranking of universities within a tier.

The same process applied to both tiers of universities. Thus, for a total cost of \$4X a student was able to obtain full knowledge of her preferences.

One tier. The six universities, namely A, B, C, D, E, and F, belonged to one tier. The ordinal preferences of each student were independently and randomly drawn from the set of all possible orderings of the six universities. Each ordering was equally likely to be drawn.

The search process was as follows:

1. For a cost of \$X, a student could pick any two universities and learn which of these two universities was ranked higher in her preference ordering. Thus, for a cost of \$X a student could learn the relative ordering of two universities.

2. Next, for an additional cost of \$X, a student could learn how a third university compares to the two that she had chosen previously. Thus, for a cost of \$2X a student could learn the relative ordering of three universities.

3. ...

4. Finally, for an additional cost of \$X, a student could learn the preference ordering of all six universities. Thus, for a total cost of \$5X a student was able to obtain full knowledge of her preferences.

We used the two-tier environment for two reasons. First, the strategies are more straightforward than in the case of one tier, as each tier essentially represented a smaller separate market, and our theoretical results apply to each tier separately.[20] In equilibrium, the six higher-ranked students are assigned to tier-A universities while the six lower-ranked students are assigned to tier-B universities. Thus, each student in equilibrium only has incentives to search within one tier. Moreover, tiered priors are more realistic in university admissions than uniform priors, since in practice students often agree on how to group universities in terms of quality, but may have different preferences within each group. We used one-tier environments because they generate higher variation in the optimal search strategies, both between students of different ranks and between treatments.

Dimension 2: Cost of information acquisition. The second dimension of our environments was the cost of information acquisition (the value of X). We considered two cost levels:

Low cost (X=\$0.5) and High cost (X=\$2.3).

When varying the cost, the predictions regarding the centralized admission procedures vary greatly. The exact parameters were chosen such that the optimal search strategies ranged from full search to no search, depending on the rank of a student. Thus, we allowed for deviations in both directions, namely under-search and over-search. Furthermore, the low-cost environment that induces full search by all students (except those with the lowest score in each tier) guarantees the informativeness of the cutoff scores under optimal search.

Table 2 presents the summary of the design by sessions, including the treatments and the order of the environments. The first four rounds were always the rounds with two tiers, while rounds five to eight were the rounds with one tier. We chose to fix this order, since the rounds with two tiers are simpler in terms of finding the optimal search strategies for subjects, and since we do not intend to directly compare the two environments. Within the two-tier and one-tier environments, there were two rounds with high cost and two rounds with low cost. To control for order effects regarding the costs, we used two different orders: either two consecutive high-cost rounds preceded the two low-cost rounds or vice versa.[21]

[20] See Appendix A.5 for details.

[21] Note that we do not randomize the order of tiers. As we are not interested in comparative statics between environments, we always start with the simpler environments with two tiers.

Sessions	Round	1	2	3	4	5	6	7	8
1, 2, 9	Treatment	DirSD							
	Tier	2				1			
	Cost	High	High	Low	Low	High	High	Low	Low
4, 7, 10	Treatment	DirSD							
	Tier	2				1			
	Cost	Low	Low	High	High	Low	Low	High	High
3, 5, 12	Treatment	SeqSD							
	Tier	2				1			
	Cost	High	High	Low	Low	High	High	Low	Low
6, 8, 11	Treatment	SeqSD							
	Tier	2				1			
	Cost	Low	Low	High	High	Low	Low	High	High
16, 17, 18	Treatment	Cutoff							
	Tier	2				1			
	Cost	High	High	Low	Low	High	High	Low	Low
13, 14, 15	Treatment	Cutoff							
	Tier	2				1			
	Cost	Low	Low	High	High	Low	Low	High	High

Table 2: Summary of sessions by treatments and environments

3.4 Experimental procedures

Across all treatments, we assigned the same randomly generated preferences to students with the same rank in the corresponding rounds and environments. For instance, the preferences of a student with rank 1 were the same in round 1 of sessions 1, 2, 11, 3, 5, 12, 16, 17, 18, and in round 3 of all other sessions. This is because we implemented two different orderings of the costs (see Table 2). For the cutoffs from the previous sessions of the DirSD treatment to be informative for the subjects in the Cutoff treatment, a correlation between the preferences of the two cohorts is necessary. This correlation was created by students with the same rank who had identical preferences in all treatments.

We used the same randomly generated scores in DirSD and SeqSD but re-generated the scores for the Cutoff treatment. This was explained to the participants in the Cutoff treatment. This design ensures that the cutoffs were informative about the competitiveness of universities, but did not provide perfect information due to the fluctuation in the distribution of scores. It also prevented the cutoffs from directly informing the students about the preferences of the previous cohort.[22]

[22]Without regenerating scores for the Cutoff treatment, a subject in a Cutoff session could have directly observed the allocation of her "copy" and thus infer the realization of her own preferences. For example, if a subject with score 85 learns that the cutoff score of University A was 85, she would know that her copy was allocated to University A. This would have affected her decisions regarding information acquisition and preference submission. This is not the type of information cutoffs carry in real markets, but is a result of a small and discrete market implemented in the experiment.

The experiment was conducted in the Experimental Economics Laboratory of the University of Melbourne (E^2MU) and was programmed using z-Tree. Upon entering the lab, subjects were provided with experimental instructions for the treatment in the two-tier environment. Before the start of the one-tier environment, an additional set of instructions was distributed. In total, we conducted 18 sessions with 24 subjects each. Thus, we had 432 participants with average earnings of 28 AUD. The sessions lasted around 80 minutes.

3.5 Theoretical predictions

The main goal of the experiment is to compare the three treatments across the different environments. Our design is not aimed at comparisons between environments under the same matching procedure. Instead, we aim to compare different matching procedures in a variety of market environments. Our main interest is the welfare of students.

Prediction 1 (Welfare): *In terms of student welfare, the following relationships hold for all students in all environments: DirSD≤Cutoff≤SeqSD.*

The comparison between DirSD and SeqSD follows from Theorem 1 and its extension to tiered priors in Appendix A.5. Because the cutoff scores represent noisy information regarding the students' budget sets under DirSD, the comparison between SeqSD and Cutoff follows from Corollary 1 as well as its extension to tiered priors in Appendix A.5.

According to the theory, the provision of cutoffs cannot decrease welfare because subjects in equilibrium hold correct beliefs about the informativeness of the cutoffs and respond optimally to them. Given the way the cutoffs are generated, they can make some students better off because the rank of a cutoff among all six cutoffs carries information about the chances of being accepted by a certain university if the previous generation has searched enough. For example, consider the one-tier low-cost environment. In equilibrium under DirSD, all students except the rank 12 student acquire full information about their preferences and submit truthful rank-order lists. The highest cutoff score among all universities cannot be the score of the students in ranks 8 to 12. Thus, we know that the university with the highest cutoff cannot be in the budget set of the students in ranks 8 to 12, the school with the second-highest cutoff cannot be in the budget sets of the students in ranks 9 to 12, and so on.

In this way, cutoffs can help some students to narrow down the options that are potentially available to them. Therefore, we predict that cutoffs can improve upon the welfare under DirSD in some settings.

Next we turn to the search behavior of students.

Prediction 2 (Search): *Participants acquire information about their own preferences following the predictions of the model. In Cutoff, students are less likely to search the universities with cutoffs higher than their score compared to the universities with cutoffs lower than their score.*

The optimal search strategies under DirSD and SeqSD are provided in Figures 4 and 5 in Appendix B.1. The predictions regarding search depend not only on the treatment, but also on each student's rank. Because the rank essentially determines the budget set of the student, the benefit of searching varies greatly between ranks. For instance, it is an optimal strategy to never search for rank 12 student, as she always receives the last available seat. Similarly, the rank 6 student never searches in the two-tier environment, since she prefers to be matched to the last available seat in tier A. Note that the search incentives depend on the size and probability distribution of the budget set. This explains why the incentive to search does not necessarily decrease for lower-ranked students, and why students do not always search weakly less under SeqSD than under DirSD, given their rank.

In contrast to DirSD and SeqSD, it is challenging to derive point predictions for the optimal search strategies in the Cutoff treatment. The main reason is that given the public information on cutoffs, students update their beliefs about their budget sets, and their prior beliefs before information acquisition are no longer uniform.[23] Recall from Section 2.4 that uniform priors are important in keeping our derivations of the optimal search strategies theoretically tractable. Without uniform priors, each student needs to consider what information other students choose to acquire in equilibrium when forming beliefs about her own budget set. We therefore choose to focus on the empirical exploration of the effects of cutoffs in this paper. Specifically, we investigate a simple strategy in the Cutoff treatment: whether subjects are less likely to search the universities with cutoffs higher than their scores. This helps us understand whether subjects use the cutoff information to narrow down the options in their budget sets and respond to it in their search strategies.

Prediction 3 (Submission): *In DirSD and Cutoff, students submit their preferences in the order of decreasing expected utility of universities, given the updated beliefs after the search. In SeqSD, a student who has searched selects the highest-ranked university among the searched ones.*

The prediction directly follows from Proposition 2.

4 Experimental Results

We start with the analysis of market outcomes and then move to the analysis of individual strategies. We can pool the sessions with different environment orderings (information cost), since the order does not significantly affect the main variables of interest (see Table B.2 in Appendix B.2). All results reported are significant at the 5% level if not stated otherwise. For all tests, we use

[23] For example, suppose that based on the cutoff information, other students can infer that the fourth ranked student knows that her budget set is more likely to include Universities A and B than the other universities. Then they would know this student is more likely to search Universities A and B and as a result, is more likely to rank A or B as her top choice than other universities. This means that the submission strategy of this student no longer follows a uniform distribution, which affects the beliefs of other students about their own budget sets.

the p-values of the coefficient of the treatment dummy in regressions on the variable of interest. Standard errors are clustered at the level of matching groups, and the sample is restricted to the treatments that are of interest for the test. We use the sign ">" between treatments to express "significantly higher," and the sign "=" to express "no significant difference."[24]

4.1 Market outcomes

4.1.1. Welfare

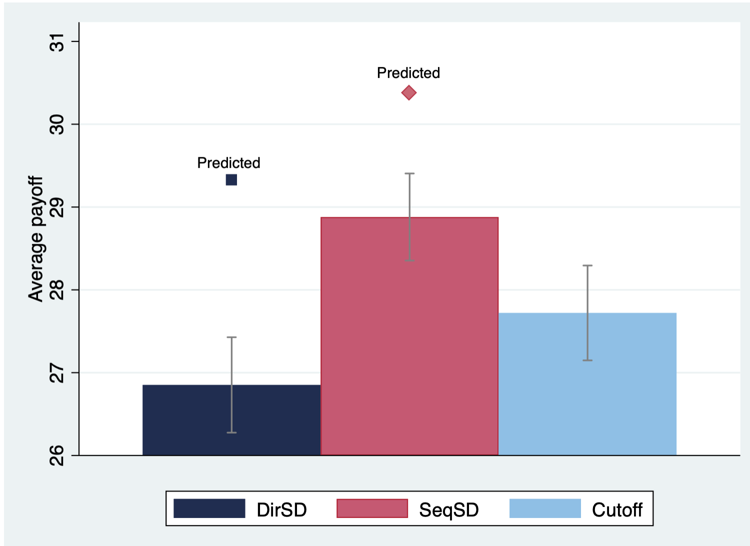
Figure 1 shows the average payoffs of participants by treatments, aggregated over all environments. The average payoff is highest in SeqSD, with the difference to DirSD and Cutoff being significant. The markers in Figure 1 indicate the theoretical predictions for the average payoffs in DirSD and SeqSD, showing that the higher welfare of participants under SeqSD compared to DirSD is in line with the theoretical predictions. However, in both SeqSD and DirSD welfare is lower than predicted: in DirSD, the average payoffs are 2.2 AUD lower than predicted, while in SeqSD the difference is 1.2 AUD. This can be due to either suboptimal search strategies or suboptimal submission strategies. We will investigate this in more detail in the following sections. In the Cutoff treatment, as predicted by the theory, we observe that average payoffs of students are higher than in DirSD ($p = 0.05$) and lower than in SeqSD ($p < 0.01$).[25]

To understand in which environments the Cutoff and SeqSD procedures have the greatest advantage over DirSD, Table 3 presents the average payoffs of participants by treatments for each tier and cost combination. First, SeqSD has significantly higher average payoffs than DirSD in all environments (see column (4) for the p-values). This confirms our theoretical prediction. Regarding the policy of providing historical cutoffs, we observe that cutoffs significantly improve the welfare of students relative to DirSD in two out of four environments, namely in the environments with a high cost of information acquisition. In these environments, the increase in welfare under Cutoff is similar to the increase under SeqSD relative to DirSD. SeqSD still generates higher welfare, but the difference is not significant. In the environments with low costs, average payoffs of participants are significantly higher in SeqSD than in Cutoff.

Next we consider the two components of welfare separately. Figure 2 presents the average payoffs of participants from the university assignments and the average search costs by treatments. It emerges that the welfare benefits of SeqSD relative to DirSD can be attributed to both sources, namely more efficient matching outcomes and lower costs of information acquisition. Both treatment differences are predicted by the theory and turn out to be significant in the experiment

[24]Due to a mistake in our code of the z-tree program, in one market of treatment SeqSD (the second round with two tiers and high cost), rank 12 students were asked to choose a university before rank 11 students. In order to keep the comparison of welfare balanced between procedures, we deleted the observations for the students ranked 11th and 12th in this round of all treatments. The inclusion of these data or the exclusion of these subjects from analyses does not change the qualitative results.

[25]Since we do not calculate a point prediction for the Cutoff treatment, there is no marker for the predicted average payoff under Cutoff in Figure 1.



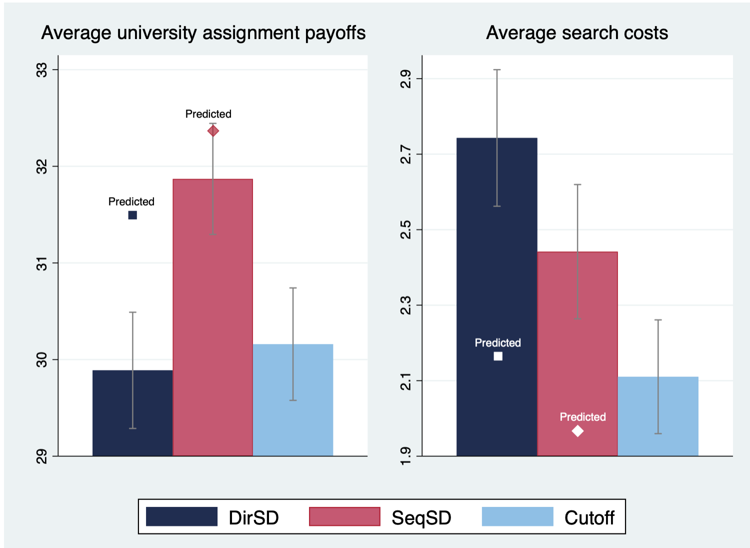
Notes: Vertical gray bars represent the 95% confidence intervals. The square marker indicates the theoretical prediction for DirSD. The diamond marker indicates the theoretical prediction for SeqSD. The y-scale is in AUD.

Figure 1: Average payoffs by treatments

	Treatment			p-value for test of equality		
	DirSD (1)	SeqSD (2)	Cutoff (3)	DirSD=SeqSD (4)	DirSD=Cutoff (5)	SeqSD=Cutoff (6)
Two tiers & low cost	26.7	27.6	26.9	0.01	0.53	0.02
Two tiers & high cost	25.0	26.2	26.1	0.00	0.01	0.61
One tier & low cost	32.6	34.9	32.4	0.01	0.79	0.00
One tier & high cost	24.1	27.9	26.7	0.00	0.00	0.07
All	26.9	29.4	27.7	0.00	0.05	0.00

Notes: For the tests in columns 4-6, we use the p-values for the coefficient of the treatment dummy in the OLS regression of payoffs on this dummy with standard errors clustered at the level of matching groups and with a sample restricted to the treatments that are of interest for the test.

Table 3: Average payoffs of subjects by treatments and environments



Notes: Vertical gray bars represent the 95% confidence intervals. Square markers indicate the theoretical predictions for DirSD. Diamond markers indicate the theoretical predictions for SeqSD. The y-scales are in AUD.

Figure 2: Average payoffs from the assignment and search costs by treatments

($p < 0.01$). At the same time, in both DirSD and SeqSD, participants receive lower than predicted payoffs from the university assignment, despite higher than predicted search costs on average. The left panel shows that the difference between the predicted and realized payoffs from the assignment is higher in DirSD than in SeqSD in the sense that in SeqSD, the predicted payoff lies in the 95% confidence interval of the realized payoffs, which is not the case in DirSD.

For the Cutoff treatment, the improvement in welfare relative to DirSD is mainly due to lower search costs. There is no significant difference in the assignment payoffs between DirSD and Cutoff ($p = 0.58$) but search costs are significantly lower under Cutoff than under DirSD ($p < 0.01$). The search costs in Cutoff are also significantly lower than in SeqSD ($p = 0.01$).

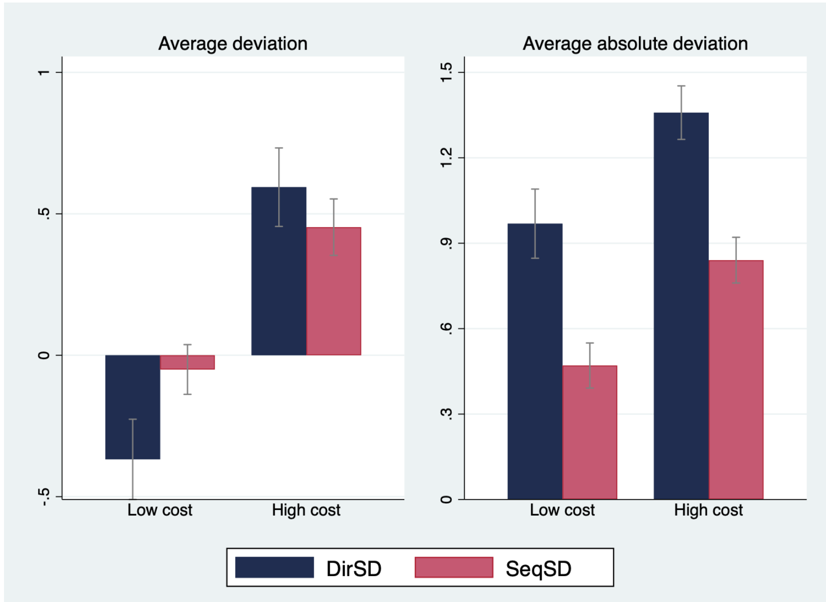
We summarize these findings in the following result.

Result 1 (Welfare):
(i) For the average payoff of participants the following relationships hold: SeqSD>Cutoff>DirSD .
(ii) For the average payoff of participants from the university assignments the following relationships hold: SeqSD>DirSD, SeqSD>Cutoff, DirSD=Cutoffs. For search costs the following relationships hold: DirSD>SeqSD>Cutoff.

4.2 Individual behavior

4.2.1 Search strategies

In this section we study the subjects' search strategies across the different treatments and environments.



Notes: Vertical gray bars represent the 95% confidence intervals. The left panel presents average deviations with under-search as a negative value and over-search as a positive value. The right panel presents average absolute deviations which are calculated as the absolute number of differences between the optimal number of searches and the observed number of searches.

Figure 3: Deviations from the optimal number of searches by treatments

First we present the main results on the search optimality in DirSD and SeqSD. The detailed analysis of search strategies by ranks and environments is presented in Appendix B.1. In the low-cost environments in DirSD, all subjects, except those ranked last in each tier, are predicted to obtain full knowledge about their preferences, but they under-search on average. Unlike in DirSD, in the low-cost environments in SeqSD, the deviations from the predicted number of searches are small. In the high-cost environments, none of the subjects are predicted to obtain full knowledge about their preferences, but we find that on average they search too much in both DirSD and SeqSD.

The left panel of Figure 3 presents the average deviation from the predicted number of searches for low- and high-costs environments in DirSD and SeqSD. In low-cost environments, students under-search on average, with significantly more under-search in DirSD than in SeqSD. In high-cost environments, we observe over-search on average. Note, however, that under-search is the only possible deviation for all but the last ranked students in each tier in the low-cost environments, and thus it has to be interpreted with caution.[26] Our results concerning over-search in high-cost environments are in line with previous experimental findings on information acquisition (see Chen and He, 2018 for school choice, Bhattacharya et al., 2017 for voting, and Gretschko and Rajko, 2015 for auctions).

[26]The findingof over-search in low-cost and under-search in high-cost environments is in line with recent findings of Descamps et al. (2021). However, under-search is not due to the equilibrium being a corner solution in their experiment.

Combining and averaging positive and negative deviations can substantially mask the actual deviations from the theory. Therefore we also consider absolute deviations. The right panel of Figure 3 presents the average absolute deviation from the predicted number of searches for low- and high-cost environments in DirSD and SeqSD. The average absolute deviation is significantly lower in SeqSD than in DirSD independent of the costs. The difference is significant for the pooled sample ($p < 0.01$) and for each environment separately ($p < 0.01$). Thus, SeqSD does not only lead to lower search costs in theory, but it also induces behavior in the lab which is more in line with the predictions than DirSD. One possible explanation for this result is that the optimal search is more straightforward for participants under SeqSD than under DirSD.

Next, we turn to the search behavior in the Cutoff treatment. On average, when the cost is low, the search under Cutoff is not significantly different from DirSD in the two-tier environment ($p = 0.32$), but is significantly lower than under DirSD in the one-tier environment ($p < 0.01$). When the cost is high, the search under Cutoff is significantly lower than under DirSD ($p < 0.01$ for both one- and two-tier environments), and under SeqSD for the one-tier environment ($p < 0.01$), but not for the two-tier environment ($p = 0.16$). Thus, the participants rely on cutoffs more in high-cost environments than in low-cost environments. As we do not form point predictions for the optimal search under Cutoff, we use regressions to analyze the empirical patterns of the subjects' reaction to cutoffs. Specifically, we investigate the simple strategy identified in Prediction 2 that subjects are less likely to search the universities with cutoffs higher than their own score compared to universities with cutoffs below their own score. Table 4 presents the marginal effects of the probit model for information acquisition about a university, depending on the cutoff of this university.

Model (1) of Table 4 presents the results for all environments of the Cutoff treatment. The coefficient of "Higher cutoff, dummy" is negative and statistically significant. Thus, on average, participants are less likely to search among the universities with cutoff scores higher than their score. This suggests that participants believe that these universities are less likely to be within their budget set. We consider the each environment separately, and find that the effect is strongest in the environments with high information costs (see models (3) and (5)). In contrast, in low-cost environments, the effect is either not significant or only marginally significant (see models (2) and (4)). Model (6) considers the absolute difference between a cutoff and a student's score for all environments. Again, the higher the cutoff is relative to the student's score, the less likely the student is to acquire information about this university. However, students are also less likely to acquire information about universities with cutoffs below their scores. The magnitude of the effect is smaller, but still significant. Models (7) and (8) study two- and one-tier environments separately. The effect of lower cutoffs remains significant in the two-tier environments. This can be explained by students in ranks 1 to 6 not searching tier-B universities. In the one-tier environments, the lower cutoff scores do not decrease the probability of search significantly, which is rational given the independence of the preferences.

	Cutoff all	Two tiers & low cost	Two tiers & high cost	One tier & low cost	One tier & high cost	Cutoff all	Two tiers	One tier
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Cost of search	-0.11					-0.11	-0.12	-0.13
	(0.01)					(0.01)	(0.01)	(0.01)
Dummy for two tiers	-0.16					-0.14		
	(0.01)					(0.01)		
Higher cutoff, dummy	-0.08	-0.02	-0.11	-0.05	-0.13			
	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)			
Higher cutoff, difference						-0.009	-0.013	-0.004
						(0.001)	(0.001)	(0.001)
Lower cutoff, difference						-0.004	-0.009	-0.0005
						(0.0003)	(0.0004)	(0.0007)
Observations	6768	1728	1584	1728	1728	6768	3456	3456

Note: Marginal effects of probit regressions regarding information acquisition about a university in Cutoff. "Higher cutoff, dummy" is a dummy that is equal to one if the cutoff score of the university minus the score of the student is greater than zero. "Higher cutoff, difference" is equal to the cutoff score of the university minus the score of the student if the difference is positive and zero otherwise. "Lower cutoff, difference" is equal to the score of the student minus the cutoff score of the university if the difference is positive and zero otherwise. Standard errors are clustered at the level of matching groups.

Table 4: Probability of information acquisition about a university depending on the cutoff

We summarize our findings regarding individual search strategies in the following result.

Result 2 (Search strategies):

- *In low-cost environments, the average number of searches in SeqSD is not statistically different from the predicted optimal strategy, while there is significant under-search in DirSD. In high-cost environments, there is over-search in DirSD and SeqSD, with larger deviations from the optimal strategy in DirSD.*

- *The average absolute deviation from the optimal search strategy is lower in SeqSD than in DirSD.*

- *Students are less likely to search universities with cutoffs higher than their score compared to universities with cutoffs below their score, especially in the high-cost environments.*

4.2.2 Submission strategies

In this section, we analyze the subjects' strategies for ranking and choosing universities. When a participant has learned her preferences completely, the optimal submission strategy is to list all universities in the order of her true preferences in DirSD and Cutoff, and to select the most preferred university from the available ones according to her true preferences in SeqSD. Note that a student's submitted list in DirSD and Cutoff is relevant only up to her guaranteed university. For example, a rank 4 participant is guaranteed a seat at the university of her second preference, since each university has two seats. Similarly, a rank 7 participant is guaranteed a seat at the university of her fourth preference. If a student does not have full knowledge of her preferences, Proposition 2 presents the optimal submission strategies for the case of one tier and Proposition 5 in Appendix A.5 presents the optimal strategies for the case of two tiers. Depending on the treatment, the optimal submission strategy leads to the following behavior:

- In DirSD, under the optimal submission strategy, universities are ranked in decreasing order of expected values. By Proposition 1, this implies listing the higher-ranked searched universities above all unsearched ones, followed by the lower-ranked searched universities.[27] The unsearched universities can be ordered in any way. If a student does not search any university, any list is optimal (respecting tiers). When counting optimal submission strategies, we only consider a student's submitted list up to her guaranteed university. For instance, for a student with rank 1, only the top choice is considered when evaluating the optimality of her strategy.

- In SeqSD, the optimal submission strategy implies choosing the highest-ranked university among those searched from the set of available universities. If a student does not search any available university, any choice is optimal (respecting tiers).

- In Cutoff, the optimal submission strategy is the same as in DirSD. Note, however, that the cutoffs might lead to multiple optimal strategies. For instance, if a student believes, based on the cutoff, that a university is out of her budget set, she is indifferent with respect to how to rank this university. Thus, some unsearched universities can be placed anywhere in the submitted rank-order list. As a benchmark, we compare the strategies in Cutoff to the optimal strategies in DirSD, thereby potentially underestimating the proportion of optimal strategies in Cutoff. The reason is that we neglect the fact that universities not in the budget set can be placed anywhere in the rank-order list.

Table 5 presents the proportion of optimal submission strategies conditional on the subjects' search behavior. The highest rate of optimal submission strategies is observed in SeqSD with differences being significant in all environments relative to DirSD and Cutoff. In SeqSD, we observe almost universally optimal submission strategies. This might be because it is rather simple to derive the optimal submission strategy under SeqSD. It does not require the ability to compare the expected utilities of all the searched and the unsearched universities. In DirSD and Cutoff, the overall proportions of optimal submission strategies are 80.4% and 75.7%, respectively, with the

[27] In our experimental setting, if the number of searched universities is even, the higher-ranked half of the searched universities should be listed above the unsearched universities, followed by the lower-ranked half of the searched universities. If the number of searched universities is odd, the optimal submission strategy is the same, but the middle-ranked searched university is treated like an unsearched university.

difference being significant for the pooled sample but not in any of the environments separately. In DirSD and Cutoff, the deviation from the optimal submission strategies might be driven either by the participants' attempt to manipulate the rank-order lists submitted to the mechanism or by the complexity of comparing the expected utilities of the searched to the unsearched universities. We find that when subjects have full knowledge of their preferences, the rate of manipulations of the submitted lists is only 8.2% in DirSD and 14.7% in Cutoff. With partial knowledge, the rates of reports where the relative positions of the searched universities are misreported are 5.7% in DirSD and 7.6% in Cutoff, which are just 29% and 31% of all suboptimal strategies in DirSD and Cutoff respectively. The remaining deviations from optimal strategies (71% in DirSD and 69% in Cutoff) are driven by incorrect rankings of unsearched universities relative to searched ones. Thus, the complexity of comparing the expected utilities of searched universities to unsearched ones contributes to the higher-than-predicted difference in welfare between SeqSD and DirSD.[28]

Result 3 (Submission strategies): *In all environments, the proportion of optimal submission strategies is significantly higher in SeqSD than in DirSD and Cutoff.*

		Treatment			p-value for test of equality		
		DirSD	SeqSD	Cutoff	DirSD=SeqSD	DirSD=Cutoff	SeqSD=Cutoff
		(1)	(2)	(3)	(4)	(5)	(6)
Optimal strategies given search	Two tiers & low cost	78.8%	97.9%	*74.0%*	0.00	0.22	0.00
	Two tiers & high cost	79.2%	98.5%	*72.3%*	0.00	0.08	0.00
	One tier & low cost	85.4%	99.3%	*78.1%*	0.00	0.07	0.00
	One tier & high cost	78.1%	98.6%	*78.1%*	0.00	1.00	0.00
	All	80.4%	98.6%	*75.7%*	0.00	0.03	0.00

Notes: For the tests in columns 4-6, we use the p-values for the coefficient of the treatment dummy in the probit regression of the optimal submission strategy. Standard errors are clustered at the level of matching groups, and the sample is restricted to the treatments that are of interest for the test. The proportions of optimal submission strategies in Cutoff are italicized as the prediction ignores that multiple optimal strategies can exist.

Table 5: Proportions of optimal submission strategies by treatments and environments

Summing up the section on individual strategies, we observe that the welfare benefits of SeqSD relative to DirSD are driven both by smaller deviations from the optimal search strategies and by a higher proportion of optimal submission strategies in SeqSD. As for the Cutoff treatment, its improvement in welfare over DirSD is driven solely by saving on search costs. Thus, in markets

[28]The rate of truthful reporting given that the students have full knowledge of their preferences is higher than in Li (2017) and Bó and Hakimov (2020), where subjects misreport their rank-order lists in 33% and 28% of cases, respectively, in DirSD. This difference may be driven by the selected sample of those who acquire full information or because our participants concentrate on optimal search strategies in our experiment, thinking less about the submission strategies, which improves truthfulness. As for SeqSD, the rate of truthful behavior is 98% in Bó and Hakimov (2020), which is similar to ours, and higher than the 83% in the first round of Li (2017) (the average for all rounds of the experiments is not reported).

where search costs are high relative to differences in the payoffs from assignments, the provision of historical cutoffs can improve welfare.

5 Discussion and conclusions

In this paper, we explore how students search university programs, how wasteful information acquisition can be reduced, and how student welfare can be improved in a market where students are ranked by universities based on exam scores. Theoretically, a sequential serial dictatorship mechanism leads to less wasteful information acquisition and higher student welfare than a direct serial dictatorship mechanism. However, we find that the theory underestimates these benefits, as in the experiments participants make superior decisions both when searching and when choosing among universities under sequential serial dictatorship compared to under direct serial dictatorship. We also find that the provision of cutoffs can increase student welfare, especially when information costs are high, although the effects of cutoffs are not as strong as the effects of using a sequential mechanism. With cutoffs, we observe that participants follow a simple strategy and avoid searching universities with cutoffs higher than their scores, especially when the costs are high. This simple strategy results in higher welfare with cutoff provision than under direct serial dictatorship without cutoff provision in high-cost environments.

Admittedly, the symmetry of universities and the uniform and tiered prior structures adopted in this paper do not describe all school and college admissions markets in practice. We have chosen this simple environment in order to keep our theoretical analysis tractable and to derive predictions for the optimal search strategies in the direct mechanism. With a more complex market structure, a student in the direct mechanism needs to consider the search strategies of higher-ranked students when forming beliefs about her own budget set. However, this is not necessary in the sequential mechanism where the budget set is known with any priors. Therefore, a more complex market structure should only strengthen our results favoring the sequential mechanism.

Our results support switching to a sequential mechanism to improve student welfare in markets where information acquisition about preferences is costly. The practical applicability of sequential mechanisms may be limited, especially if students take a considerable amount of time to acquire information, which can make the matching process substantially longer. This is an important concern. However, sequential mechanisms have become more widespread recently, enabled by the digitization of assignment procedures and the possibility to coordinate repeatedly and in real time through online platforms. In the dynamic mechanism used in France, students have five days to decide among their offers during the first weeks of the procedure, three days in the following weeks, and one day towards the end of the procedure. With these deadlines for decisions, students have enough time to study in detail the content of each program and other relevant features, such as the costs of living. The procedure has been run successfully for four years, and the main admissions are made within 50 days.[29] While in practice it is unlikely that students will move one by one, sequential decisions by groups of students, depending on their scores, are realistic, and the welfare benefits relative to the direct mechanism persist, according to our theoretical analysis.[30] Sequential decisions based on groups preserve the benefits of sequential mechanisms from the perspective of information acquisition, as lower-ranked groups of students do not have to acquire information about universities that are filled by higher-ranked students.[31] Finally, if it is not possible to implement a sequential mechanism, the policy of providing historical cutoffs also improves welfare. While this policy has already been implemented in some countries, our study provides empirical support for this practice, especially in markets when information acquisition costs are high.

References

Abdulkadiroğlu, A., J. Angrist, and P. Pathak (2014). The elite illusion: Achievement effects at boston and new york exam schools. *Econometrica 82*(1), 137–196.

Abdulkadiroğlu, A. and T. Sönmez (1998). Random serial dictatorship and the core from random endowments in house allocation problems. *Econometrica 66*(3), 689–701.

Abdulkadiroğlu, A. and T. Sönmez (2003, June). School choice: A mechanism design approach. *The American Economic Review 93*(3), 729–747.

Ajayi, K. and M. Sidibe (2020). School choice under imperfect information. Technical report, mimeo.

Ajayi, K. F., W. H. Friedman, and A. M. Lucas (2020). When information is not enough: Evidence from a centralized school choice system. Technical report, National Bureau of Economic Research.

Artemov, G. (2019). Assignment mechanisms: common preferences and information acquisition. Technical report, Working Paper.

Atakan, A. E. (2006). Assortative matching with explicit search costs. *Econometrica 74*(3), 667–680.

[29]Indeed, in the first year, the procedure took too long to converge, which led to a change in the procedures with a decrease from seven to five days during the first weeks of the procedure.

[30]Tunisia uses a sequential version of SD, with top, middle, and bottom grades deciding sequentially, after the choices of the previous group are finalized. Also, the Chinese province of Inner Mongolia uses a dynamic mechanism where the choices of students are finalized sequentially by groups, depending on students' scores. Students with higher scores leave the procedure before students with lower scores. See Gong and Liang (2016) for details.

[31]A group-by-group sequential SD can help students narrow down their budget sets by eliminating schools that are filled by higher-ranked groups. But the information provided by the mechanism is not precise enough to pin down the exact budget sets because students in the same group move simultaneously. Thus, the information is noisy and Corollary 1 in our paper applies.

Azevedo, E. M. and J. D. Leshno (2016). A supply and demand framework for two-sided matching markets. *Journal of Political Economy 124* (5), 1235–1268.

Baccara, M., A. İmrohoroğlu, A. J. Wilson, and L. Yariv (2012). A field study on matching with network externalities. *American Economic Review 102* (5), 1773–1804.

Bade, S. (2015). Serial dictatorship: The unique optimal allocation rule when information is endogenous. *Theoretical Economics 10(2)*, 385–410.

Basteck, C., B. Klaus, and D. Kübler (2021). How lotteries in school choice help to level the playing field. *Games and Economic Behavior*.

Bhattacharya, S., J. Duffy, and S. Kim (2017). Voting with endogenous information acquisition: Experimental evidence. *Games and Economic Behavior 102*, 316–338.

Bó, I. and R. Hakimov (2019). The iterative deferred acceptance mechanism. *Available at SSRN 2881880*.

Bó, I. and R. Hakimov (2020). Iterative versus standard deferred acceptance: Experimental evidence. *The Economic Journal 130* (626), 356–392.

Bó, I. and R. Hakimov (2020). Pick-an-object mechanisms. *Available at SSRN*.

Bó, I. and C. Y. Ko (2020). Competitive screening and information transmission. *Journal of Public Economic Theory*.

Chade, H. and L. Smith (2006). Simultaneous search. *Econometrica 74* (5), 1293–1307.

Chen, Y. and Y. He (2018). Information acquisition and provision in school choice: an experimental study. Technical report, mimeo.

Chen, Y. and Y. He (2019). Information acquisition and provision in school choice: a theoretical investigation. Technical report, mimeo.

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

Coles, P. and R. Shorrer (2014). Optimal truncation in matching markets. *Games and Economic Behavior 87*, 591–615.

Das, S. and Z. Li (2014). The role of common and private signals in two-sided matching with interviews. In *International Conference on Web and Internet Economics*, pp. 492–497. Springer.

Descamps, A., S. Massoni, and L. Page (2021). Learning to hesitate. *Experimental Economics*, 1–25.

Doval, L. (2018). Whether or not to open pandora's box. *Journal of Economic Theory 175*, 127–158.

Drummond, J. and C. Boutilier (2014). Preference elicitation and interview minimization in stable matchings. In *Twenty-Eighth AAAI Conference on Artificial Intelligence*.

Dur, U., R. G. Hammond, and O. Kesten (2017). Sequential school choice: Theory and evidence from the field and lab. Technical report, Working Paper, North Carolina State University.

Echenique, F., A. J. Wilson, and L. Yariv (2016). Clearinghouses for two-sided matching: An experimental study. *Quantitative Economics 7*(2), 449–482.

Eeckhout, J. and P. Kircher (2010). Sorting and decentralized price competition. *Econometrica 78*(2), 539–574.

Gale, D. and L. S. Shapley (1962, January). College admissions and the stability of marriage. *The American Mathematical Monthly 69*(1), 9–15.

Gong, B. and Y. Liang (2016). A dynamic college admission mechanism in inner mongolia: Theory and experiment. Technical report, Working paper, 2016. 2, 10.

Grenet, J., Y. He, and D. Kübler (2019). Decentralizing centralized matching markets: Implications from early offers in university admissions. Technical report, mimeo.

Gretschko, V. and A. Rajko (2015). Excess information acquisition in auctions. *Experimental Economics 18*(3), 335–355.

Hakimov, R. and D. Kübler (2020). Experiments on centralized school choice and college admissions: a survey. *Experimental Economics*, 1–55.

Harless, P. and V. Manjunath (2018). Learning matters: Reappraising object allocation rules when agents strategically investigate. *International Economic Review 59*(2), 557–592.

Hastings, J. S., C. A. Neilson, and S. D. Zimmerman (2013). Are some degrees worth more than others? evidence from college admission cutoffs in chile. Technical report, National Bureau of Economic Research.

Immorlica, N. S., J. D. Leshno, I. Y. Lo, and B. J. Lucier (2020). Information acquisition in matching markets: The role of price discovery. Mimeo, CRC TRR 190.

Kadam, S. V. (2015). Interviewing in matching markets. Technical report, Working Paper.

Klijn, F., J. Pais, and M. Vorsatz (2019). Static versus dynamic deferred acceptance in school choice: Theory and experiment. *Games and Economic Behavior 113*, 147–163.

Lee, R. S. and M. Schwarz (2017). Interviewing in two-sided matching markets. *The RAND Journal of Economics 48*(3), 835–855.

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

Lien, J. W., J. Zheng, and X. Zhong (2016). Preference submission timing in school choice matching: testing fairness and efficiency in the laboratory. *Experimental Economics 19*(1), 116–150.

Luflade, M. (2019). The value of information in centralized school choice systems (job market paper).

Narita, Y. (2018). Match or mismatch? learning and inertia in school choice. Unpublished.

Noda, S. (2020). Strategic experimentation with random serial dictatorship. *Available at SSRN 2956548*.

Pan, S. (2019). The instability of matching with overconfident agents. *Games and Economic Behavior 113*, 396–415.

Pan, S. (2020). Experiments in market design. In *Handbook of Experimental Game Theory*. Edward Elgar Publishing.

Rastegari, B., A. Condon, N. Immorlica, and K. Leyton-Brown (2013). Two-sided matching with partial information. In *Proceedings of the fourteenth ACM conference on Electronic commerce*, pp. 733–750.

Shimer, R. and L. Smith (2000). Assortative matching and search. *Econometrica 68*(2), 343–369.

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

Weitzman, M. L. (1979). Optimal search for the best alternative. *Econometrica: Journal of the Econometric Society*, 641–654.

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

Appendix

A Proofs and additional theoretical analysis

In the subsequent analysis, we omit the subscript i when referring to any student.

A.1 Proof of Proposition 1

The proof of Proposition 1, as well as several subsequent results, will use the following lemma. Suppose that $\lambda^1(x)$ and $\lambda^2(x)$ are two probability mass functions (PMFs) of distributions over the same discrete domain Ψ, and that and $\Lambda^1(x)$ and $\Lambda^2(x)$ are their corresponding cumulative distribution functions (CDFs). Let $\eta(x)$ be the difference between these two PMFs, that is, $\eta(x) \equiv \lambda^1(x) - \lambda^2(x)$.

Lemma 1. *If there exists a threshold $\hat{x} \in \Psi$ such that $\eta(x) \leq 0$ for $x \leq \hat{x}$ and $\eta(x) > 0$ otherwise, then Λ^1 first-order stochastically dominates Λ^2, that is, $\Lambda^1(x) \leq \Lambda^2(x), \forall x$.*

Proof. Denote the smallest and largest values in Ψ as $\underline{x}$ and $\bar{x}$, respectively. Denote x^+ as the smallest element in Ψ that is greater than x (for $x < \bar{x}$). Given the definition of $\hat{x}$, we know that when $x \leq \hat{x}$,

$$\Lambda^1(x) - \Lambda^2(x) = \sum_{x'=\underline{x}}^{x} \eta(x') \leq 0.$$

When $x > \hat{x}$,

$$\begin{aligned}
\Lambda^1(x) - \Lambda^2(x) &= \sum_{x'=\underline{x}}^{x} \eta(x') \\
&= \sum_{x'=\underline{x}}^{\hat{x}} \eta(x') + \sum_{x'=\hat{x}^+}^{x} \eta(x') \\
&\leq \sum_{x'=\underline{x}}^{\hat{x}} \eta(x') + \sum_{x'=\hat{x}^+}^{x} \eta(x') + \sum_{x'=x^+}^{\bar{x}} \eta(x') \\
&= \sum_{x'=\underline{x}}^{\bar{x}} \eta(x') \\
&= 0.
\end{aligned}$$

The last step is due to the definition of $\eta(x)$. Therefore, we have $\Lambda^1(x) \leq \Lambda^2(x), \forall x$. The inequality holds strictly for some x as long as the two distributions are not identical. Hence, Λ^1 first-order stochastically dominates Λ^2. $\square$

Now we start to prove Proposition 1.

Proof. We write the expected utility for those unsearched universities in $C \setminus C^S$ as

$$V(0) = \sum_{j=1}^{m} f^0(j,\alpha)\, u^j$$

and the updated expected utility for the university relatively ranked γth in C^S $(\gamma = 1, \ldots, \alpha+1)$ as

$$V^\gamma(\alpha) = \sum_{j=1}^{m} f^\gamma(j,\alpha)\, u^j,$$

in which

$$f^0(j,\alpha) = \frac{1}{m}$$

and

$$f^\gamma(j,\alpha) = \frac{\begin{pmatrix} j-1 \\ \gamma-1 \end{pmatrix}\begin{pmatrix} m-j \\ \alpha-\gamma+1 \end{pmatrix}}{\begin{pmatrix} m \\ \alpha+1 \end{pmatrix}}$$

are the PMFs of the distributions over the set of cardinal utilities $\{u^1, ..., u^m\}$; let $F^0(j,\alpha)$ and $F^\gamma(j,\alpha)$ be the corresponding CDFs.

(1) We first show that $V^1(\alpha) > V(0)$ for any $\alpha = 1, 2, \ldots, m-1$. Let $g^{1,0}(j,\alpha)$ be the difference between the two PMFs $f^1(j,\alpha)$ and $f^0(j,\alpha)$, that is,

$$g^{1,0}(j,\alpha) \equiv f^1(j,\alpha) - f^0(j,\alpha) = \frac{\begin{pmatrix} m-j \\ \alpha \end{pmatrix}}{\begin{pmatrix} m \\ \alpha+1 \end{pmatrix}} - \frac{1}{m}.$$

We can see from the above definition that (i) given α and m, $g^{1,0}(j,\alpha)$ is decreasing in j;[32] (ii) $g^{1,0}(m,\alpha) = -\frac{1}{m} < 0$; and (iii) $g^{1,0}(1,\alpha) = \frac{\alpha+1}{m} > 0$. Therefore, there exists an integer $\hat{j}$ such that $g^{1,0}(j,\alpha) \leq 0$ when $\hat{j} \leq j \leq m$, and $g^{1,0}(j,\alpha) > 0$ when $1 \leq j \leq \hat{j}' - 1$. Because $u^1 > u^2 > \ldots > u^m$, $u^{\hat{j}}$ is equivalent to the threshold $\hat{x}$ in Lemma 1. According to Lemma 1, $F^1(j,\alpha)$ first-order stochastically dominates $F^0(j,\alpha)$, that is, $V^1(\alpha) > V(0)$ for any $\alpha = 1, 2, \ldots, m-1$.

Next, we show $V^{\alpha+1}(\alpha) < V(0)$ for any $\alpha = 1, 2, \ldots, m-1$. Let $g^{0,\alpha+1}(j,\alpha)$ be the difference

[32] $f^1(j,\alpha)$ equals zero when $j > m - \alpha$ and is strictly decreasing in j when $j \leq m - \alpha$.

between the two PMFs $f^0(j,\alpha)$ and $f^{\alpha+1}(j,\alpha)$, that is,

$$g^{0,\alpha+1}(j,\alpha) \equiv f^0(j,\alpha) - f^{\alpha+1}(j,\alpha) = \frac{1}{m} - \frac{\begin{pmatrix} j-1 \\ \alpha \end{pmatrix}}{\begin{pmatrix} m \\ \alpha+1 \end{pmatrix}}.$$

We know from the above definition that (i) given α and m, $g^{0,\alpha+1}(j,\alpha)$ is decreasing in j;[33] (ii) $g^{0,\alpha+1}(m,\alpha) = -\frac{\alpha}{m} < 0$; and (iii) $g^{0,\alpha+1}(1,\alpha) = \frac{1}{m} > 0$. Therefore, there exists an integer $\hat{j}'$ such that $g^{0,\alpha+1}(j,\alpha) \leq 0$ when $\hat{j}' \leq j \leq m$, and $g^{0,\alpha+1}(j,\alpha) > 0$ when $1 \leq j \leq \hat{j}'-1$. Again, according to Lemma 1, $F^0(j,\alpha)$ first-order stochastically dominates $F^{\alpha+1}(j,\alpha)$, that is, $V^{\alpha+1}(\alpha) < V(0)$ for any $\alpha = 1,2,\ldots,m-1$.

(2) We first show that $V^{\gamma}(\alpha) > V^{\gamma+1}(\alpha)$ for any $\gamma = 1,2,\ldots,\alpha+1$ and $\alpha = 1,2,\ldots,m-1$. Let $g^{\gamma,\gamma+1}(j,\alpha)$ be the difference between the two PMFs $f^{\gamma}(j,\alpha)$ and $f^{\gamma+1}(j,\alpha)$, that is,

$$\begin{aligned} g^{\gamma,\gamma+1}(j,\alpha) &\equiv f^{\gamma}(j,\alpha) - f^{\gamma+1}(j,\alpha) \\ &= \frac{\begin{pmatrix} j-1 \\ \gamma-1 \end{pmatrix}\begin{pmatrix} m-j \\ \alpha-\gamma+1 \end{pmatrix}}{\begin{pmatrix} m \\ \alpha+1 \end{pmatrix}} - \frac{\begin{pmatrix} j-1 \\ \gamma \end{pmatrix}\begin{pmatrix} m-j \\ \alpha-\gamma \end{pmatrix}}{\begin{pmatrix} m \\ \alpha+1 \end{pmatrix}} \end{aligned}$$

Because $f^{\gamma}(j,\alpha) = f^{\gamma+1}(j,\alpha) = 0$ when $j > m-\alpha+\gamma$ or $j < \gamma$, we re-define $f^{\gamma}(j,\alpha)$, $f^{\gamma+1}(j,\alpha)$, and $g^{\gamma,\gamma+1}(j,\alpha)$ to be the PMFs over the set $\{u^{\gamma},...,u^{m-\alpha+\gamma}\}$. For $\gamma < j < m-\alpha+\gamma$, $f^{\gamma}(j,\alpha) > 0$, $f^{\gamma+1}(j,\alpha) > 0$, and

$$g^{\gamma,\gamma+1}(j,\alpha) \propto (m+1)\gamma - (\alpha+1)j.$$

Because $g^{\gamma,\gamma+1}(\gamma,\alpha) = f^{\gamma}(\gamma,\alpha) - 0 > 0$ and $g^{\gamma,\gamma+1}(m-\alpha+\gamma,\alpha) = 0 - f^{\gamma+1}(m-\alpha+\gamma,\alpha) < 0$, we know $g^{\gamma,\gamma+1}(\gamma,\alpha) \leq 0$ if $\frac{(m+1)\gamma}{\alpha+1} \leq j \leq m-\alpha+\gamma$ and $g^{\gamma,\gamma+1}(\gamma,\alpha) > 0$ if $\gamma \leq j < \frac{(m+1)\gamma}{\alpha+1}$. According to Lemma 1, $F^{\gamma}(j,\alpha)$ first-order stochastically dominates $F^{\gamma+1}(j,\alpha)$, that is, $V^{\gamma}(\alpha) > V^{\gamma+1}(\alpha)$ for any $\alpha = 1,2,\ldots,m-1$.[34]

Now we have shown that given any $\alpha = 1,2,\ldots,m-1$, $V^1(\alpha) > V(0)$, $V^{\alpha+1}(\alpha) < V(0)$, and $V^{\gamma}(\alpha) > V^{\gamma+1}(\alpha), \forall\gamma = 1,2,\ldots,\alpha+1$. Therefore, by the mean value theorem, there exists a threshold $\hat{\gamma}(\alpha)$ at which (i) $V^{\gamma}(\alpha) > V(0)$ for all $\gamma \leq \hat{\gamma}(\alpha)$, and (ii) $V^{\gamma}(\alpha) \leq V(0)$ otherwise. □

[33] $f^{\alpha+1}(j,\alpha)$ equals zero when $j < \alpha+1$ and is strictly increasing in j when $j \geq \alpha+1$.

[34] Since $\frac{(m+1)\gamma}{\alpha+1}$ is not necessarily an integer, the threshold in Lemma 1 can be considered as $u^{\left\lceil\frac{(m+1)\gamma}{\alpha+1}\right\rceil}$, where $\left\lceil\frac{(m+1)\gamma}{\alpha+1}\right\rceil$ is the ceiling of $\frac{(m+1)\gamma}{\alpha+1}$, i.e., the smallest integer greater than $\frac{(m+1)\gamma}{\alpha+1}$.

A.2 Information acquisition under DirSD

In this section, we discuss the role of information and students' information acquisition strategy under DirSD.

Proposition 3. *Under DirSD, the marginal benefit of additional information is non-negative and can be non-monotonic.*

The proof of Proposition 3 will use the following two lemmas.

Lemma 2. *For any $\alpha = 2, \ldots, m-1$ and $\gamma = 1, 2, \ldots, \alpha+1$, $V^{\gamma}(\alpha) > V^{\gamma}(\alpha-1)$.*

Proof. Suppose a student has completed $(\alpha-1)$ steps of searching and is considering the benefit of step α, $\alpha = 2, \ldots, m-1$. When this additional step of search is conducted, the change in expected value is given by

$$\begin{aligned} V^{\gamma}(\alpha) - V^{\gamma}(\alpha-1) &= \sum_{j=1}^{m} f^{\gamma}(j,\alpha)\, u^{j} - \sum_{j=1}^{m} f^{\gamma}(j,\alpha-1)\, u^{j} \\ &= \sum_{j=\gamma}^{m-\alpha+\gamma-1} \frac{\begin{pmatrix} j-1 \\ \gamma-1 \end{pmatrix}\begin{pmatrix} m-j \\ \alpha-\gamma+1 \end{pmatrix}}{\begin{pmatrix} m \\ \alpha+1 \end{pmatrix}} u^{j} - \sum_{j=\gamma}^{m-\alpha+\gamma} \frac{\begin{pmatrix} j-1 \\ \gamma-1 \end{pmatrix}\begin{pmatrix} m-j \\ \alpha-\gamma \end{pmatrix}}{\begin{pmatrix} m \\ \alpha \end{pmatrix}} u^{j}. \end{aligned}$$

Let $h(j)$ be the difference between the two PMFs $f^{\gamma}(j,\alpha)$ and $f^{\gamma}(j,\alpha-1)$, that is,

$$\begin{aligned} h(j) &\equiv f^{\gamma}(j,\alpha) - f^{\gamma}(j,\alpha-1) \\ &= \frac{\begin{pmatrix} j-1 \\ \gamma-1 \end{pmatrix}\begin{pmatrix} m-j \\ \alpha-\gamma+1 \end{pmatrix}}{\begin{pmatrix} m \\ \alpha+1 \end{pmatrix}} - \frac{\begin{pmatrix} j-1 \\ \gamma-1 \end{pmatrix}\begin{pmatrix} m-j \\ \alpha-\gamma \end{pmatrix}}{\begin{pmatrix} m \\ \alpha \end{pmatrix}}. \end{aligned}$$

When $j = m-\alpha+\gamma$, $h(m-\alpha+\gamma) = 0 - \dfrac{\begin{pmatrix} m-\alpha+\gamma-1 \\ \gamma-1 \end{pmatrix}}{\begin{pmatrix} m \\ \alpha \end{pmatrix}} < 0$. When $j \leq m-\alpha+\gamma-1$,

$$h(j) \propto (m+1)\gamma - (\alpha+1)j.$$

We can see that $h(j) \leq 0$ if $\frac{(m+1)\gamma}{\alpha+1} \leq j \leq m-\alpha+\gamma$ and $h(j) > 0$ if $\gamma \leq j < \frac{(m+1)\gamma}{\alpha+1}$. According to Lemma 1, $F^{\gamma}(j,\alpha)$ first-order stochastically dominates distribution $F^{\gamma}(j,\alpha-1)$. Hence, $V^{\gamma}(\alpha) > V^{\gamma}(\alpha-1)$ for any $\alpha = 2, \ldots, m-1$ and $\gamma = 1, 2, \ldots, \alpha+1$. □

Lemma 3. *For any $\alpha = 2, \ldots, m-1$ and $\gamma = 1, 2, \ldots, \alpha$, $V^{\gamma}(\alpha - 1) > V^{\gamma+1}(\alpha)$.*

Proof. The proof of this lemma is similar to the proof of Lemma 3. Given $\alpha = 2, \ldots, m-1$ and $\gamma = 1, 2, \ldots, \alpha$,

$$\begin{aligned} V^{\gamma}(\alpha-1) - V^{\gamma+1}(\alpha) &= \sum_{j=1}^{m} f^{\gamma}(j, \alpha-1) u^{j} - \sum_{j=1}^{m} f^{\gamma+1}(j,\alpha) u^{j} \\ &= \sum_{j=\gamma}^{m-\alpha+\gamma} \frac{\begin{pmatrix} j-1 \\ \gamma-1 \end{pmatrix} \begin{pmatrix} m-j \\ \alpha-\gamma \end{pmatrix}}{\begin{pmatrix} m \\ \alpha \end{pmatrix}} u^{j} - \sum_{j=\gamma+1}^{m-\alpha+\gamma} \frac{\begin{pmatrix} j-1 \\ \gamma \end{pmatrix} \begin{pmatrix} m-j \\ \alpha-\gamma \end{pmatrix}}{\begin{pmatrix} m \\ \alpha+1 \end{pmatrix}} u^{j}. \end{aligned}$$

Let $h'(j)$ be the difference between the two PMFs $f^{\gamma}(j, \alpha-1)$ and $f^{\gamma+1}(j,\alpha)$, that is,

$$\begin{aligned} h'(j) &\equiv f^{\gamma}(j,\alpha-1) - f^{\gamma+1}(j,\alpha) \\ &= \frac{\begin{pmatrix} j-1 \\ \gamma-1 \end{pmatrix} \begin{pmatrix} m-j \\ \alpha-\gamma \end{pmatrix}}{\begin{pmatrix} m \\ \alpha \end{pmatrix}} - \frac{\begin{pmatrix} j-1 \\ \gamma \end{pmatrix} \begin{pmatrix} m-j \\ \alpha-\gamma \end{pmatrix}}{\begin{pmatrix} m \\ \alpha+1 \end{pmatrix}}. \end{aligned}$$

When $j = \gamma$, $h'(\gamma) = \frac{\begin{pmatrix} m-\gamma \\ \alpha-\gamma \end{pmatrix}}{\begin{pmatrix} m \\ \alpha \end{pmatrix}} - 0 > 0$. When $j \geq \gamma + 1$,

$$h'(j) \propto (m+1)\gamma - (\alpha+1) j.$$

Therefore, $h'(j) \leq 0$ if $\frac{(m+1)\gamma}{\alpha+1} \leq j \leq m - \alpha + \gamma$ and $h'(j) > 0$ if $\gamma \leq j < \frac{(m+1)\gamma}{\alpha+1}$. According to Lemma 1, $F^{\gamma}(j, \alpha-1)$ first-order stochastically dominates $F^{\gamma+1}(j,\alpha)$. Thus, $V^{\gamma}(\alpha-1) > V^{\gamma+1}(\alpha)$ for any $\alpha = 2, \ldots, m-1$ and $\gamma = 1, 2, \ldots, \alpha$. □

Now we move to prove Proposition 3: under DirSD, the marginal benefit of additional information (1) is non-negative, and (2) can be non-monotonic.

Proof. (1) Suppose the student submits a list $\hat{\succ}_i$ under DirSD. Let Q_i^{θ} be the probability that she is accepted by the θth ranked university in $\hat{\succ}_i$. Recall that in each step of DirSD, the student whose turn it is is assigned to the highest-ranked university in her submitted list from those that still have vacant seats. Thus, for any probability distribution over one's budget set, DirSD ensures that $Q^{\theta} \geq Q^{\theta'}$ if $\theta < \theta'$, that is, a student is more likely to be admitted by a university if it is higher ranked in her submitted list.

A student who stops searching at step α and chooses the optimal strategy of truth-telling under DirSD, according to Propositions 1 and 2, would rank the unsearched universities below the $\hat{\gamma}(\alpha)$th-ranked searched university, but above the $(\hat{\gamma}(\alpha)+1)$th-ranked searched university, and would rank the searched universities according to the discovered relative preferences. Hence, her expected utility is given by

$$u^{DirSD}(\alpha) = \sum_{\theta=1}^{\hat{\gamma}(\alpha)} Q^\theta V^\theta(\alpha) + \sum_{\theta=\hat{\gamma}(\alpha)+1}^{\hat{\gamma}(\alpha)+m-\alpha-1} Q^\theta V(0) + \sum_{\theta=\hat{\gamma}(\alpha)+m-\alpha}^{m} Q^\theta V^{\theta-m+\alpha+1}(\alpha) - \alpha k,$$

in which αk is the total cost of information. For any $\alpha = 2, \ldots, m-1$, the benefit of conducting an additional search step under DirSD is given by $A(\alpha) - A(\alpha-1)$, where

$$A(\alpha) \equiv \sum_{\theta=1}^{\hat{\gamma}(\alpha)} Q^\theta V^\theta(\alpha) + \sum_{\theta=\hat{\gamma}(\alpha)+1}^{\hat{\gamma}(\alpha)+m-\alpha-1} Q^\theta V(0) + \sum_{\theta=\hat{\gamma}(\alpha)+m-\alpha}^{m} Q^\theta V^{\theta-m+\alpha+1}(\alpha),$$

and thus

$$A(\alpha-1) = \sum_{\theta=1}^{\hat{\gamma}(\alpha-1)} Q^\theta V^\theta(\alpha-1) + \sum_{\theta=\hat{\gamma}(\alpha-1)+1}^{\hat{\gamma}(\alpha-1)+m-\alpha} Q^\theta V(0) + \sum_{\theta=\hat{\gamma}(\alpha-1)+m-\alpha+1}^{m} Q^\theta V^{\theta-m+\alpha}(\alpha-1).$$

Recall that $\hat{\gamma}(\alpha)$ is the threshold at which $V^\gamma(\alpha) > V(0)$ for all $\gamma \leq \hat{\gamma}(\alpha)$ and $V^\gamma(\alpha) \leq V(0)$ otherwise. From Lemma 2, we know that $V^\gamma(\alpha) > V^\gamma(\alpha-1)$, $\forall\gamma$, therefore we have $\hat{\gamma}(\alpha-1) \leq \hat{\gamma}(\alpha)$.

First, when $\hat{\gamma}(\alpha-1) = \hat{\gamma}(\alpha)$,

$$\begin{aligned}
&A(\alpha) - A(\alpha-1) \\
=&\sum_{\theta=1}^{\hat{\gamma}(\alpha)} Q^\theta \Bigg[\underbrace{V^\theta(\alpha) - V^\theta(\alpha-1)}_{>0}\Bigg] + \sum_{\theta=\hat{\gamma}(\alpha)+1}^{\hat{\gamma}(\alpha)+m-\alpha-1} Q^\theta \Bigg[\underbrace{V(0) - V(0)}_{=0}\Bigg] \\
&+ Q^{\hat{\gamma}(\alpha)+m-\alpha} \Bigg[\underbrace{V^{\hat{\gamma}(\alpha)+1}(\alpha) - V(0)}_{\leq 0}\Bigg] + \sum_{\theta=\hat{\gamma}(\alpha)+m-\alpha+1}^{m} Q^\theta \Bigg[\underbrace{V^{\theta-m+\alpha+1}(\alpha) - V^{\theta-m+\alpha}(\alpha-1)}_{<0}\Bigg].
\end{aligned}$$

In the above equation, $\left[V^\theta(\alpha) - V^\theta(\alpha-1)\right]$ is positive according to Lemma 2, the second term is zero, $\left[V^{\hat{\gamma}(\alpha)+1}(\alpha) - V(0)\right]$ is non-positive according to the definition of $\hat{\gamma}(\alpha)$, and $[V^{\theta-m+\alpha+1}(\alpha) - V^{\theta-m+\alpha}(\alpha-1)]$ is negative according to Lemma 3. Since for any α the total expected value of all universities is a constant equal to $\sum_{j=1}^{m} u^j$, we have

$$\begin{aligned}
&\sum_{\theta=1}^{\hat{\gamma}(\alpha)} \left[V^{\theta}(\alpha) - V^{\theta}(\alpha-1)\right] \\
&= -\left\{ \left[V^{\hat{\gamma}(\alpha)+1}(\alpha) - V(0)\right] + \sum_{\theta=\hat{\gamma}(\alpha)+m-\alpha+1}^{m} \left[V^{\theta-m+\alpha+1}(\alpha) - V^{\theta-m+\alpha}(\alpha-1)\right] \right\}.
\end{aligned}$$

Because Q^{θ} weakly increases as θ decreases, the positive term outweighs the negative, that is,

$$\begin{aligned}
&\sum_{\theta=1}^{\hat{\gamma}(\alpha)} Q^{\theta}\left[V^{\theta}(\alpha) - V^{\theta}(\alpha-1)\right] \\
&\geq -\left\{ Q^{\hat{\gamma}(\alpha)+m-\alpha}\left[V^{\hat{\gamma}(\alpha)+1}(\alpha) - V(0)\right] + \sum_{\theta=\hat{\gamma}(\alpha)+m-\alpha+1}^{m} Q^{\theta}\left[V^{\theta-m+\alpha+1}(\alpha) - V^{\theta-m+\alpha}(\alpha-1)\right] \right\}.
\end{aligned}$$

Therefore, we can conclude that $A(\alpha) \geq A(\alpha-1)$ for any $\alpha = 2, \ldots, m-1$.

Next, when $\hat{\gamma}(\alpha-1) < \hat{\gamma}(\alpha)$,

$$\begin{aligned}
&A(\alpha) - A(\alpha-1) \\
&= \sum_{\theta=1}^{\hat{\gamma}(\alpha-1)} Q^{\theta}\left[\underbrace{V^{\theta}(\alpha) - V^{\theta}(\alpha-1)}_{>0}\right] + \sum_{\theta=\hat{\gamma}(\alpha-1)+1}^{\hat{\gamma}(\alpha)} Q^{\theta}\left[\underbrace{V^{\theta}(\alpha) - V(0)}_{>0}\right] + \sum_{\theta=\hat{\gamma}(\alpha)+1}^{\hat{\gamma}(\alpha-1)+m-\alpha} Q^{\theta}\left[\underbrace{V(0) - V(0)}_{=0}\right] \\
&\quad + \sum_{\theta=\hat{\gamma}(\alpha-1)+m-\alpha+1}^{\hat{\gamma}(\alpha)+m-\alpha-1} Q^{\theta}\left[\underbrace{V(0) - V^{\theta-m+\alpha}(\alpha-1)}_{\geq 0}\right] + \sum_{\theta=\hat{\gamma}(\alpha)+m-\alpha}^{m} Q^{\theta}\left[\underbrace{V^{\theta-m+\alpha+1}(\alpha) - V^{\theta-m+\alpha}(\alpha-1)}_{<0}\right].
\end{aligned}$$

In the above equation, $\left[V^{\theta}(\alpha) - V^{\theta}(\alpha-1)\right]$ is positive according to Lemma 2, $\left[V^{\theta}(\alpha) - V(0)\right]$ is positive according to the definition of $\hat{\gamma}(\alpha)$, the third term is zero, $\left[V(0) - V^{\theta-m+\alpha}(\alpha-1)\right]$ is non-negative according to the definition of $\hat{\gamma}(\alpha-1)$, and $\left[V^{\theta-m+\alpha+1}(\alpha) - V^{\theta-m+\alpha}(\alpha-1)\right]$ is negative according to Lemma 3. Similar to the previous case, since the positive difference equals the absolute value of the negative difference but has more weight, we can again conclude that $A(\alpha) \geq A(\alpha-1)$ for any $\alpha = 2, \ldots, m-1$.

Lastly, when $\alpha = 1$, we know from Proposition 1 that $V^1(1) > V(0)$ and $V^2(1) < V(0)$. Thus,

$$\begin{aligned}
A(1) &= Q^1 V^1(1) + \sum_{\theta=2}^{m-1} Q^{\theta} V(0) + Q^m V^2(1) \\
&\geq Q^1 V(0) + \sum_{\theta=2}^{m-1} Q^{\theta} V(0) + Q^m V(0) \\
&= V(0) \equiv A(0)
\end{aligned}$$

Again, the inequality is due to the fact that $V^1(1) - V(0) = -[V^2(1) - V(0)]$ and $Q^1 \geq Q^m$.

Therefore, we can conclude that $A(\alpha) \geq A(\alpha')$ for any $\alpha > \alpha'$. That is, the marginal benefit of additional information is non-negative under DirSD.

(2) Under DirSD, the benefit of information is rescaled by the probabilities Q^θ's. Therefore, depending on the ex ante probability distribution of a student's budget set, the marginal benefit of additional information is not necessarily decreasing.

With Assumption 1 and uniform priors, each student knows that the rank-order list submitted by any other student is equally likely to be any ranking in Ω. Thus, from the perspective of student i, she always has an equal chance at every university, and this chance is given by $\left\{P_i\left(\tilde{B}\right)\right\}_{\tilde{B}\subseteq C}$. This makes the "name" of a university irrelevant to the student. Let $\{p_i(\beta)\}_{\beta=1,2,\ldots,m}$ be the probability distribution of the number of universities in student i's budget set, that is, $p_i(\beta) = \Pr[|B_i| = \beta]$, $\beta = 1, 2, \ldots, m$.[35] We thus have $P_i\left(\tilde{B}\right) = P_i\left(\tilde{B}'\right) = p_i(\beta)/\begin{pmatrix} m \\ \beta \end{pmatrix}$ for all $\left|\tilde{B}\right| = \left|\tilde{B}'\right| = \beta$. That is, only the number of universities in a student's budget set but not its specific composition matters for her decisions under DirSD. For instance, consider a market with three universities $C = \{c_1, c_2, c_3\}$, each of which has two seats. The budget set of the student ranked third in the exam depends on the submitted rank-order lists of the two students ranked above her. If, for example, they both place university c_3 on the top of their lists, which occurs with probability $\frac{1}{3} \times \frac{1}{3} = \frac{1}{9}$, then the budget set of the student ranked third contains only c_1 and c_2. The same probability $\frac{1}{9}$ should be assigned to all possible two-university compositions of her budget set: $\{c_1, c_2\}$, $\{c_1, c_3\}$, and $\{c_2, c_3\}$.

Suppose a student submits a list $\hat{\succ}$ under DirSD. Then given $\{p(\beta)\}_{\beta=1,2,\ldots,m}$, the probability that she is accepted by the θth ranked university in $\hat{\succ}$ is given by

$$Q^\theta = \sum_{\beta=1}^{m-\theta+1} \frac{\begin{pmatrix} m-\theta \\ \beta-1 \end{pmatrix}}{\begin{pmatrix} m \\ \beta \end{pmatrix}} p(\beta).$$

If a student is assigned to her θth choice, her budget set B has to include her θth choice and exclude the $(\theta - 1)$ universities listed above it. With probability $p(\beta)$, B includes β universities. One of them has to be her θth choice and the remaining $(\beta - 1)$ ones cannot be her top θ choices, which means $\begin{pmatrix} m-\theta \\ \beta-1 \end{pmatrix}$ out of the $\begin{pmatrix} m \\ \beta \end{pmatrix}$ possible compositions can occur. Thus, Q^θ sums up, for all possible values of β, the probability that the student is accepted by her θth choice. We can see that for any probability distribution over one's budget set, DirSD ensures that $Q^\theta \geq Q^{\theta'}$

[35] A student has at least one university in her budget set because we assume the total number of seats exceeds the total number of students. This assumption simplifies our analysis, but is not crucial.

if $\theta < \theta'$, that is, a student is more likely to be admitted by a university if it is higher ranked in her submitted list. This, again, proves the optimality of the truth-telling strategy stated in Proposition 2.

Now consider our experimental market with one tier. There are six universities, and each has two seats. The cardinal utilities of every student are determined by the experimental payments $\{u^1, u^2, u^3, u^4, u^5, u^6\} = \{40, 34, 28, 22, 16, 10\}$.

For the student ranked first in the exam, $\{p(1), p(2), p(3), p(4), p(5), p(6)\} = \{0, 0, 0, 0, 0, 1\}$ and the marginal benefit of each additional step of searching is $A(1) - A(0) = 7$, $A(2) - A(1) = 3.5$, $A(3) - A(2) = 2.1$, $A(4) - A(3) = 1.4$, $A(5) - A(4) = 1$, which is decreasing.

However, for the student ranked tenth in the exam, $\{p(1), p(2), p(3), p(4), p(5), p(6)\} \approx \{0, 0.57, 0.43, 0, 0, 0\}$ and the marginal benefit of each additional step of searching is approximately $A(1) - A(0) \approx 2.84$, $A(2) - A(1) \approx 1.42$, $A(3) - A(2) \approx 1.87$, $A(4) - A(3) \approx 1.25$, $A(5) - A(4) \approx 1.14$, which is clearly non-monotonic.

This implies that under DirSD, the optimal information acquisition strategy is not necessarily unique in the general setting. However, we ensure the uniqueness for every student in each treatment of our experimental design. □

A.3 Information acquisition under SeqSD

In this section, we discuss the role of information and students' information acquisition strategy under SeqSD.

Proposition 4. *Under SeqSD,*

(1) the marginal benefit of an additional step of searching among available universities is non-negative and decreasing;

(2) the optimal stopping point α^{SeqSD} in a student's search process is characterized as (i) $\alpha^{SeqSD} = 0$ if $V^1(1) - V(0) < k$; (ii) $\alpha^{SeqSD} = 1$ if $V^1(1) - V(0) > k$ and $V^1(2) - V^1(1) \leq k$; and (iii) α^{SeqSD} solves $\left[V^1\left(\alpha^{SeqSD}\right) - V^1\left(\alpha^{SeqSD} - 1\right)\right] > k$ and $\left[V^1\left(\alpha^{SeqSD} + 1\right) - V^1\left(\alpha^{SeqSD}\right)\right] \leq k$ otherwise.

Proof. (1) First, we show that the marginal benefit of an additional step of searching is non-negative.

Under SeqSD, each student, when being considered, is asked to select from all universities that still have vacant seats, that is, from all universities in her budget set B. Obviously, a student would not search outside her budget set because the information about unavailable universities cannot affect her selection. When searching within B, a student who stops at step α and chooses the optimal strategy of truth-telling under SeqSD, according to Propositions 1, would choose the university with the highest expected utility. Hence, her expected utility at this point is given by $V^1(\alpha) - \alpha k$, in which αk is the total cost of information.

Suppose that when a student is considered by SeqSD, there is only one university left available, that is, her budget set includes only one university ($|B| = 1$). Thus, she obviously has no incentive to invest in any information and the marginal benefit of additional information is constantly zero.

Suppose a student is asked by SeqSD to choose from multiple universities ($|B| > 1$). According to Proposition 2, it is an optimal strategy for her to choose the university with the highest expected utility. Then the marginal benefit of conducting the first step of searching is $V^1(1) - V(0)$ and the marginal benefit of conducting an additional subsequent search step is given by $V^1(\alpha) - V^1(\alpha - 1)$, $\alpha = \{2, \ldots, |B| - 1\}$. According to Lemma 2, we know that $V^1(\alpha) > V^1(\alpha - 1)$ for any $\alpha = 2, \ldots, m-1$, and we have already shown that $V^1(1) > V(0)$ in Proposition 1. Therefore, $V^1(\alpha) > V^1(\alpha')$ for any $\alpha > \alpha'$ when $|B| > 1$.

Combining the cases of $|B| = 1$ and $|B| > 1$, we can conclude the marginal benefit of additional information is non-negative under SeqSD.

Next, we consider the change in marginal benefit during a student's search process under SeqSD.

We only consider a student with $|B| > 2$ because one with $|B| \leq 2$ would not conduct multiple steps of search. The difference in marginal benefits between an increase from $(\alpha - 1)$ to α and an increase from α to $(\alpha + 1)$, $\alpha = 2, \ldots, |B| - 2$ is given by

$$
\begin{aligned}
& \left[V^1(\alpha) - V^1(\alpha - 1)\right] - \left[V^1(\alpha + 1) - V^1(\alpha)\right] \\
=& 2V^1(\alpha) - V^1(\alpha + 1) - V^1(\alpha - 1) \\
=& 2\sum_{j=1}^{m-\alpha} f^1(j, \alpha) u^j - \sum_{j=1}^{m-\alpha-1} f^1(j, \alpha + 1) u^j - \sum_{j=1}^{m-\alpha+1} f^1(j, \alpha - 1) u^j.
\end{aligned}
$$

Define $\chi(j)$ as the corresponding difference in PMFs:

$$
\begin{aligned}
\chi(j) &\equiv 2f^1(j, \alpha) - f^1(j, \alpha + 1) - f^1(j, \alpha - 1). \\
&= 2\frac{\begin{pmatrix} m - j \\ \alpha \end{pmatrix}}{\begin{pmatrix} m \\ \alpha + 1 \end{pmatrix}} - \frac{\begin{pmatrix} m - j \\ \alpha + 1 \end{pmatrix}}{\begin{pmatrix} m \\ \alpha + 2 \end{pmatrix}} - \frac{\begin{pmatrix} m - j \\ \alpha - 1 \end{pmatrix}}{\begin{pmatrix} m \\ \alpha \end{pmatrix}}.
\end{aligned}
$$

We can calculate that $\chi(m - \alpha + 1) = -\frac{1}{\begin{pmatrix} m \\ \alpha \end{pmatrix}} < 0$, $\chi(m - \alpha) = 2\frac{1}{\begin{pmatrix} m \\ \alpha + 1 \end{pmatrix}} - \frac{\alpha}{\begin{pmatrix} m \\ \alpha \end{pmatrix}}$, and thus

$$\begin{aligned}\chi(m-\alpha+1)+\chi(m-\alpha) &= \frac{2}{\begin{pmatrix} m \\ \alpha+1 \end{pmatrix}} - \frac{\alpha+1}{\begin{pmatrix} m \\ \alpha \end{pmatrix}} \\ &\propto \alpha+2-m \\ &\leq 0.\end{aligned}$$

Therefore, the difference in CDFs is non-positive when $j \geq m-\alpha$. For $1 \leq j \leq m-\alpha-1$,

$$\chi(j) \propto (j-1)\left[2(m+1)-(\alpha+2)j\right].$$

We can see that $\chi(j)<0$ if $\frac{2(m+1)}{\alpha+2} < j \leq m-\alpha-1$ and $\chi(j) \geq 0$ if $1 \leq j \leq \frac{2(m+1)}{\alpha+2}$. According to Lemma 1, the difference in CDFs are non-positive at any j, which indicates first-order stochastic dominance. Hence, we conclude that $\left[V^1(\alpha)-V^1(\alpha-1)\right] > \left[V^1(\alpha+1)-V^1(\alpha)\right]$ for any $\alpha = 2,\ldots,|B|-2$.

When $\alpha=1$, the difference in marginal benefits between an increase from $(\alpha-1)$ to α and an increase from α to $(\alpha+1)$ is given by

$$\begin{aligned}&\left[V^1(1)-V(0)\right]-\left[V^1(2)-V^1(1)\right] \\ =&2\sum_{j=1}^{m-1} f^1(j,1)\,u^j - \sum_{j=1}^{m-2} f^1(j,2)\,u^j - \frac{1}{m}\sum_{j=1}^{m} u^j\end{aligned}$$

Define $\chi^1(j)$ as the corresponding difference in PMFs:

$$\begin{aligned}\chi^1(j) &\equiv 2f^1(j,1)-f^1(j,2)-\frac{1}{m} \\ &= 2\frac{m-j}{\begin{pmatrix} m \\ 2 \end{pmatrix}} - \frac{\begin{pmatrix} m-j \\ 2 \end{pmatrix}}{\begin{pmatrix} m \\ 3 \end{pmatrix}} - \frac{1}{m}\end{aligned}$$

When $j \leq m-2$,

$$\chi^1(j) = \frac{(j-1)\left(2(m+1)-3j\right)}{m(m-1)(m-2)}.$$

Therefore, $\chi^1(j)<0$ when $\frac{2(m+1)}{3} < j \leq m-2$, and $\chi^1(j) \geq 0$ when $1 \leq j \leq \frac{2(m+1)}{3}$. We can also calculate that $\chi^1(m) = -\frac{1}{m} < 0$ and $\chi^1(m-1) = \frac{4}{m(m-1)} - \frac{1}{m}$. When $m>4$, $\chi^1(m-1) \leq 0$ and thus $\chi^1(j) \leq 0$ for $\frac{2(m+1)}{3} < j \leq m$ and $\chi^1(j) \geq 0$ for $1 \leq j \leq \frac{2(m+1)}{3}$. According to Lemma 1, the difference in CDFs is non-positive at any j, which indicates first-order stochastic dominance and thus $[V^1(1) - V(0)] > [V^1(2) - V^1(1)]$. When $m = 3$, we know $V(0) = \frac{1}{3}(u^1 + u^2 + u^3)$, $V^1(1) = \frac{2}{3}u^1 + \frac{1}{3}u^2$, $V^1(2) = u^1$, and thus $[V^1(1) - V(0)] > [V^1(2) - V^1(1)]$ as $u^1 > u^2 > u^3$. When $m = 4$, we know $V(0) = \frac{1}{4}(u^1 + u^2 + u^3 + u^4)$, $V^1(1) = \frac{1}{2}u^1 + \frac{1}{3}u^2 + \frac{1}{6}u^3$, $V^1(2) = \frac{3}{4}u^1 + \frac{1}{4}u^2$, and thus $[V^1(1) - V(0)] > [V^1(2) - V^1(1)]$ as $u^1 > u^2 > u^3 > u^4$. Therefore, $[V^1(1) - V(0)] > [V^1(2) - V^1(1)]$ holds for any $m > 2$.

To sum up, we can conclude that the marginal benefit of information acquisition within one's budget set decreases under SeqSD.

(2) Since the marginal benefit of an additional step of searching within B decreases and the marginal cost is constantly k, it is optimal for a student to adopt another step of searching as long as the marginal benefit exceeds the marginal cost, and stop searching otherwise. Specifically, the optimal stopping point α^{SeqSD} in the search process is characterized as

(i) $\alpha^{SeqSD} = 0$ if $V^1(1) - V(0) < k$;

(ii) $\alpha^{SeqSD} = 1$ if $V^1(1) - V(0) > k$ and $V^1(2) - V^1(1) \leq k$; and

(iii) α^{SeqSD} solves $\left[V^1\left(\alpha^{SeqSD}\right) - V^1\left(\alpha^{SeqSD} - 1\right)\right] > k$ and $\left[V^1\left(\alpha^{SeqSD} + 1\right) - V^1\left(\alpha^{SeqSD}\right)\right] \leq k$ otherwise.

Due to the discreteness of the problem, under some parameters a student may be indifferent between two optimal stopping points if the marginal benefit of the last step of searching equals k; here we assume the student chooses the smaller one. Otherwise the optimal stopping point is unique. $\square$

A.4 Proof of Theorem 1

Proof. This proof does not rely on a particular search technology. We define student i's search decision as a choice s_i from the set S_i and denote its cost as $d_i(s_i)$. For the search technology specified in Section 2.3, student i's search decision s_i represents her stopping point $\alpha_i \in \{0, 1, \ldots, m-1\}$ and its cost is $d_i(s_i) = \alpha_i k_i$.

According to Proposition 2, a student who adopts the optimal strategy of truth-telling ranks universities according the expected utilities from high to low under DirSD and chooses the university with the highest expected utility under SeqSD. In both cases, the student bases her submission strategy on her updated beliefs about her preferences after search and is accepted by the university with the highest expected utility in her budget set. The expected utility of this university, denoted as $EU\left(s, \tilde{B}\right)$, is thus determined by the student's search decision and her budget set. With Assumption 1 and uniform priors, the ex ante probability distribution of a student's budget set $\left\{P\left(\tilde{B}\right)\right\}_{\tilde{B}\subseteq C}$ does not depend on the search strategies of others and is the same under DirSD and SeqSD.

Under DirSD, all students simultaneously submit their rank-order lists. A student takes her search decision s based on the ex ante probability distribution of her budget set $\left\{P\left(\tilde{B}\right)\right\}_{\tilde{B}\subseteq C}$ and needs to pay the information cost $d(s)$. Thus, the optimization problem under DirSD is given by

$$U^{DirSD} = \max_{s \in S} \left[\left(\sum_{\tilde{B} \subseteq C} P\left(\tilde{B}\right) EU\left(s, \tilde{B}\right) \right) - d(s) \right].$$

Under SeqSD, a student selects the preferred university after the higher-ranked students have made their choices. She therefore observes the realization of her budget set before she makes her search decision. Thus, the optimization problem under SeqSD is given by

$$U^{SeqSD} = \sum_{\tilde{B} \subseteq C} P\left(\tilde{B}\right) \max_{s \in S} \left[EU\left(s, \tilde{B}\right) - d(s) \right].$$

Therefore, we have

$$\begin{aligned}
U^{DirSD} &= \max_{s \in S} \left[\left(\sum_{\tilde{B} \subseteq C} P\left(\tilde{B}\right) EU\left(s, \tilde{B}\right) \right) - d(s) \right] \\
&= \max_{s \in S} \left[\sum_{\tilde{B} \subseteq C} P\left(\tilde{B}\right) \left(EU\left(s, \tilde{B}\right) - d(s) \right) \right] \\
&\leq \sum_{\tilde{B} \subseteq C} P\left(\tilde{B}\right) \max_{s \in S} \left[EU\left(s, \tilde{B}\right) - d(s) \right] \\
&= U^{SeqSD}.
\end{aligned}$$

Hence, we conclude that a student with any probability distribution for her budget set is weakly better off under SeqSD than under DirSD. □

A.5 Tiered priors

In this appendix, we consider the following prior structure. Universities belong to different "tiers," ranked from better to worse. All students have the same between-tier preference: they all prefer any university in a better tier to any university in a worse tier. However, students may have different within-tier preferences: each student's preference over universities in the same tier follows a uniform distribution, that is, it is equally likely to be any linear order over these universities. Formally, let $\{T_t\}_{t=1,2,\ldots,\tau}$ be a partition of the set of universities C. For any $c \in T_t$, $c' \in T_{t'}$, and $i \in I$, we have $c \succ_i c'$ if $t < t'$. That is, all students prefer any university in T_1 to any university in T_2, prefer any university in T_2 to any university in T_3, and so on. This between-tier preference is common knowledge to the entire market. Via costly information acquisition, a student can learn more about the realization of her own within-tier preferences, but not the realization of other students' within-tier preferences. That is, the information acquired by each student is her private information. Unlike the uniform priors introduced in Section 2.3, this prior structure allows for both a common and a private factor in students' preferences. For example, in many real-life university admission markets, there is usually a common consensus or a clear definition as to which universities belong to the top tier, to the second tier, and so on. But students' tastes over universities in the same tier may vary depending on the location, family culture, personal taste, etc.

Each student can acquire further information about her own within-tier preferences for zero, one, or multiple tiers. The search in each tier follows the same technology as described in Section 2.3. A student starts by choosing any two universities from this tier and learns their relative ordering. In each of the subsequent steps, she chooses one more university from the same tier to learn the relative ordering of all the universities she has chosen. Students can stop at any step in the process. With a cost of $(|T_t|-1)\,k$, the student can fully discover her preferences over universities in T_t. With a total cost of $\sum_{t=1}^{\tau}(|T_t|-1)\,k$, the student can obtain full knowledge of her own preferences.

A.5.1 Preference submission

The optimality of truth-telling strategies under DirSD or SeqSD does not depend on the prior structure. Therefore, Proposition 2 still holds with tiered priors. With aligned between-tier preferences, we can further characterize the truth-telling strategies under the two mechanisms.

Proposition 5. *In a market with tiered priors, a student who adopts the truth-telling strategies, regardless of her knowledge about her within-tier preferences, would always:*

(i) rank any university in a better tier above any university in a worse tier in her submitted rank-order list under DirSD, and

(ii) choose a university in the best tier among those available to her under SeqSD.

We use q_{T_t} to denote the total capacity of all universities in T_t, that is, $q_{T_t}=\sum_j q_j$ for all j such that $c_j \in T_t$. From the strategies characterized above, we know that in equilibria with truth-telling strategies under both mechanisms, students with the exam rank $r \le q_{T_1}$ are admitted to universities in T_1, students with $q_{T_1} \le r \le q_{T_1}+q_{T_2}$ are admitted to universities in T_2, and so on. In general, students with $\sum_{\tilde{t}=1}^{t-1} q_{T_{\tilde{t}}} \le r \le \sum_{\tilde{t}=1}^{t} q_{T_{\tilde{t}}}$ are admitted to universities in T_t.

A.5.2 Information acquisition and welfare comparison

Proposition 5 implies that we can also categorize students by tiers: we say those with $\sum_{\tilde{t}=1}^{t-1} q_{T_{\tilde{t}}} \le r \le \sum_{\tilde{t}=1}^{t} q_{T_{\tilde{t}}}$ are "tier-t students" since they would be admitted to a tier-t university under DirSD and SeqSD as long as all students adopt truth-telling strategies. For a tier-t student, universities in a better tier $T_{t'}(t'<t)$ are definitely not in her budget set. On the other hand, although universities in a worse tier $T_{t'}(t'>t)$ are certainly available to the student, she can always secure a seat at a tier-t university by adopting the truth-telling strategy. This means a tier-t student only needs to consider universities in T_t when choosing strategies; universities in other tiers are essentially irrelevant for her decision-making. Therefore, for any given $t = 1, 2, \ldots, \tau$, we can consider all the tier-t universities and tier-t students as an independent market, and this market is identical to the market with uniform priors. We extend Assumption 1 as follows.

Assumption 2. *(1) All universities in the same tier have the same capacity; universities in different tiers can have different capacities.*

(2) In each step of the search process, a tier-t student is equally likely to choose any one of the unsearched tier-t universities to investigate.

(3) If a tier-t student did not search all universities under DirSD, she is equally likely to choose any relative order over the unsearched tier-t universities in her submitted rank-order list. If a tier-t student did not search any tier-t universities that she is asked to select under SeqSD, she is equally likely to select any one of these tier-t universities.

We summarize the conclusion regarding information acquisition in the following proposition.

Proposition 6. *In a market with tiered priors,*

(i) a tier-t student only searches among tier-t universities if she chooses to acquire information; and

(ii) her search strategy among tier-t universities is the same as that in a uniform-prior market with only tier-t universities.

Because with tiered priors each tier can be treated as a separate market, we can apply Theorem 1 and Corollary 1 and conclude the following.

Theorem 2. *In the case of tiered priors, every student is weakly better off under SeqSD than under DirSD ex ante if all students acquire information optimally and adopt the truth-telling submission strategy.*

Corollary 2. *In the case of tiered priors,*

1. *even when students are provided with noisy information about their budget sets (for instance historical cutoffs) under DirSD, every student is weakly better off under SeqSD than under DirSD ex ante if all students acquire information optimally and adopt the truth-telling submission strategy;*

2. *students cannot be worse off under DirSD ex ante when provided with noisy information about their budget sets (for instance historical cutoffs).*

Therefore, the advantage of SeqSD in student welfare persists in environments with tiered priors.

B Additional experimental results

B.1 Details on individual search strategies

We consider each environment separately, since optimal search strategies differ greatly depending on search costs and whether the preferences are tiered.[36] Figure 4 presents the average cost of information acquisition by treatments when the cost is low at \$0.5.

The left panel of Figure 4 presents optimal and actual search strategies for the two-tier environment. In DirSD under low information costs, the optimal search strategy for all subjects (except for subjects ranked sixth and 12th) is to invest \$1 to obtain full certainty about their own preferences in the respective tier. Note that subjects with score ranks 1-5 should only consider the universities in tier A while subjects with score ranks 7-11 should only consider universities in tier B. On average, we observe that subjects search too little, except for rank 7 subjects who, on average, over-search by investing in information about universities in both tiers. The excessive search by rank 7 subjects may be driven by optimism that some of the subjects ranked 1 to 6 will be assigned to a tier B university, due to suboptimal preference submission.[37] In Cutoff, the behavior is similar to DirSD (p-value for the test of difference is 0.32). Thus, the cutoff provision does not have a significant effect on search strategies in the two-tier markets with low costs. On the one hand, the cutoffs are informative due to the full uncertainty resolution in the equilibrium of DirSD. On the other hand, the benefit of relying on cutoff information is relatively small, as the total cost of optimal information acquisition is just \$1. Thus, subjects might not risk saving \$1 by relying on cutoff information. As for subjects with ranks 6 and 12, they should not invest in information at all, as they both get the only free seat of the corresponding tier in equilibrium. However, we observe a high degree of over-search by these subjects.[38] As for SeqSD, the actual search behavior of subjects is, on average, remarkably in line with the theoretical predictions. The actual search costs are significantly lower than in DirSD and Cutoff (the p-value for the test of difference is $<$0.01 for both comparisons). Thus, the optimal search strategy in SeqSD is more straightforward for subjects than in DirSD. This is not surprising, as the optimal strategy consists of full investment in resolving uncertainty about one's available universities, and the only deviation could be under-search or searching before the allocation procedure started–that is, before one learns which universities are available to her.

The right panel of Figure 4 presents the predictions and actual search strategies for the one-tier environment with low costs. In DirSD, the optimal search strategy for all subjects (except rank

[36]As explained in Section 3.5, we do not derive point predictions for optimal search strategies in the Cutoff treatment.

[37]In total, a rank 7 participant had the potential choice between universities in the top tier due to suboptimal strategies of higher-ranked participants in only 1 out of 48 rounds of DirSD.

[38]Note that in our experimental setup all students had to submit the full rank-order list of universities in DirSD and Cutoff, or had to choose one university in SeqSD, thus making it impossible to remain unassigned. Therefore, not searching is an optimal strategy for rank 12 students.

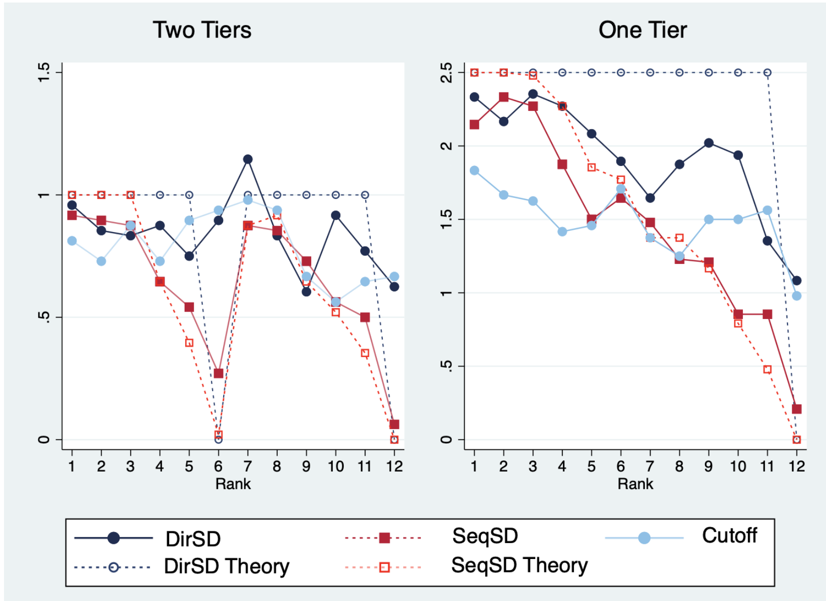


Figure 4: Average costs of information acquisition with low costs by treatments

12 subjects) is to invest \$2.5 to obtain full certainty about their preferences. We observe that rank 1 to 11 subjects search too little, which is even more pronounced for subjects with ranks 6 to 11 than for subjects with ranks 1 to 5. Note that the relative benefit of search decreases with the rank, thus it can be partially driven by the risk aversion of subjects. Another possibility is that subjects perceive the preferences as correlated, and thus overestimate the chances that the most-preferred universities will be assigned to the higher-ranked subjects. In Cutoff, the actual search costs are significantly lower than in DirSD for ranks 1 to 10 (p-value for the test of difference is <0.01 for these ranks, and for all ranks). Again, cutoffs are informative due to full uncertainty resolution in the equilibrium of DirSD. In the two-tier environment, the potential benefit of relying on cutoffs for subjects is only \$1. In the one-tier environment, the optimal information cost is \$2.5 and the potential benefit of cutoffs from the perspective of saving search costs is higher. As for rank 12 subjects, they should not invest in information at all, but they invest on average \$1.08 in DirSD and \$0.98 in Cutoff. This violates the optimal strategy of not searching. As for SeqSD, the actual search behavior of subjects is remarkably in line with the theoretical predictions on average. The most substantial deviation is under-search of the subjects ranked 1 to 5. Again, the optimal strategy in SeqSD consists of obtaining full certainty about the ranking of all available universities, and the only deviation could be under-search or search before the allocation procedure started. When the optimal strategy requires an investment of \$2.5, and thus five steps of search, subjects often stop after four steps of search, thus underestimating the probability of the last university being preferred to the other five universities. This under-search in SeqSD is similar to the under-search of rank 1 to 3 subjects in DirSD. Overall, the actual search costs in SeqSD are significantly lower than in DirSD, but not significantly different from Cutoff (p-value for the test of difference

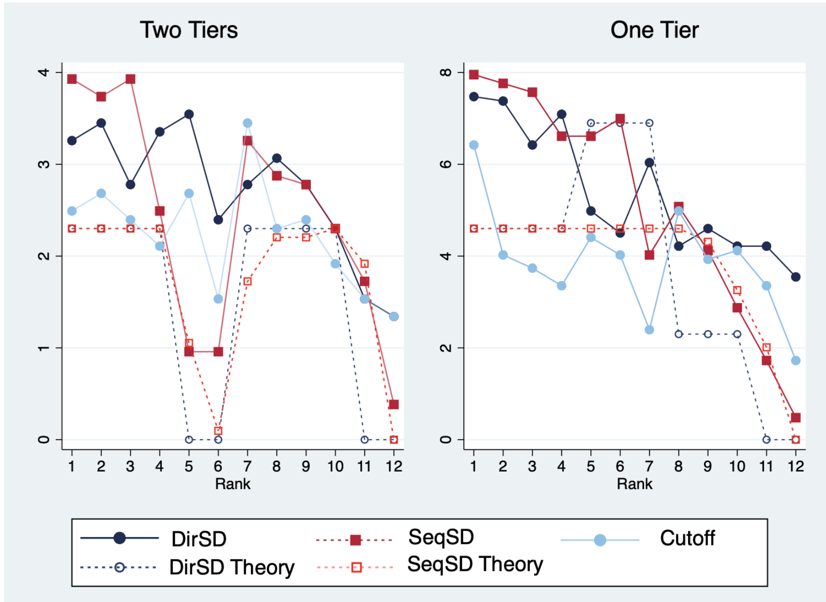


Figure 5: Average costs of information acquisition with high costs by treatments

is <0.01 and equal to 0.79 respectively).

Figure 3 presents the average cost of information acquisition by treatments when the cost is high at \$2.3. The left panel of Figure 5 presents predictions and actual search strategies for the two-tier environment. First, in DirSD the optimal search strategy for rank 1 to 4 and 7 to 10 subjects is to invest \$2.3 in resolving uncertainty about the relative ranking of any two universities in the respective tier. Thus, in the high-cost treatments subjects never obtain full certainty about the university rankings. Students ranked 5, 6, 11, and 12 should not invest in search at all. Unlike in treatments with low costs, we observe significant over-search in DirSD for all ranks. This finding is in line with previous experimental findings on information acquisition (see Chen and He, 2018 for school choice, Bhattacharya et al., 2017 for voting, and Gretschko and Rajko, 2015 for auctions). In Cutoff, the actual search costs are lower than in DirSD (p-value<0.01) with the highest difference for the lower-ranked students. Unlike the two-tier low-cost environment when the potential benefit of relying on cutoffs saves subjects only up to \$1, in the two-tier high-cost environment the optimal information cost is \$2.3. Thus, the potential benefit of cutoffs for saving information costs is much higher. Subjects rely on the cutoffs following the higher potential saving of information costs. In SeqSD with high costs, unlike in SeqSD with low costs where the actual search behavior of subjects is mostly in line with theoretical predictions, we observe a high degree of over-search for students ranked 1 to 3 and 7 to 9. The over-search for ranks 1 to 3 is even higher than in DirSD. As for ranks 5, 6, 11, and 12, the behavior is more in line with the theory than in the other treatments. Overall, in the two-tier high-cost environment, there is no significant difference in the average actual search costs between SeqSD and DirSD (p=0.12), and between SeqSD and Cutoff (p=0.16).

Finally, the right panel of Figure 5 presents predicted and actual search strategies for the one-tier high-cost environment. In DirSD, the optimal search strategy for rank 1 to 4 subjects is to invest \$4.6 in resolving the uncertainty about the ranking of any three universities. Similar to the two-tier environment with high costs, students ranked 1 to 4 over-search relative to the optimal strategy. Ranks 5 to 7 have an optimal strategy of investing \$6.9 to resolve uncertainty about the ranking of four out of six universities. Note that this is the only case where the lower-ranked subjects search more in theory than the higher-ranked subjects. This pattern, however, finds no support in the data, as the students ranked 5 to 7 search less than students ranked 1 to 4. As for ranks 8 to 12, they all invest on average around \$4 in information acquisition, despite an optimum of \$2.3 for ranks 9 and 10 and an optimum of \$0 for ranks 11 and 12. In Cutoff, the actual search costs are lower than in DirSD ($p<0.01$), with larger differences for the higher-ranked students. Just as in the two-tier environment with high costs, subjects rely on the cutoffs leading to lower information costs than in DirSD. Yet again, they ignore the fact that in the high-cost environments, the cutoffs are less informative about the preferences of the previous cohort than in the low-cost environments, as many submissions of the previous cohort are made without resolving preference uncertainty. Note, however, that in both high-cost environments, in DirSD subjects over-invest in information relative to the optimal strategy. Thus, the cutoffs are more informative than in equilibrium. As for SeqSD, we observe a high degree of over-search for students ranked 1 to 6. As for ranks 7 to 12, the behavior is more in line with the theory than in the other treatments. Overall, in the one-tier high-cost environment, the average actual search costs in SeqSD are significantly higher than in Cutoff ($p<0.01$), and not significantly different from DirSD ($p=0.52$).

B.2 Order effects

	Total payoff	Total payoff	Number of searches	Number of searches	Optimal strategy	Optimal strategy
Order	-0.00	-0.00	0.06	0.06	0.00	0.00
	(0.41)	(0.29)	(0.10)	(0.07)	(0.06)	(0.02)
SeqSD		2.05		-0.38		0.26
		(0.36)		(0.07)		(0.02)
Cutoff		0.90		-0.47		-0.14
		(0.43)		(0.10)		(0.02)
Tiers		-3.17		-1.19		0.09
		(0.49)		(0.12)		(0.03)
Cost of search		-2.34		-0.44		-0.03
		(0.12)		(0.03)		(0.01)
Period		0.07		0.04		0.01
		(0.10)		(0.03)		(0.01)
Observations	3384	3384	3384	3384	3384	3384
R	0.00	0.09	0.00	0.22	0.00	0.16

Note: Results of OLS regressions with clustering of standard errors on the level of matching groups. Order is a dummy variable equal to 0 when Low cost preceded High cost, and equal to 1 when High cost preceded Low cost. SeqSD is a dummy for treatment SeqSD, Cutoff is a dummy for treatment Cutoff. Tier is equal to 1 in One-tier environments and equal to 2 in Two-tier environments.

Table 6: Order effects