Priority Uncertainty and Cutoff Signals in Decentralized College Admissions*

Hiroto Sato† Ryo Shirakawa‡

October 11, 2022

Abstract

Priority structures are uncertain in real-life college admissions markets. This study investigates how information structures on priority affect resulting allocations. To do this, we focus on a class of real-life information structures that are represented as *cutoff signals*, which privately tell whether a priority exceeds a cutoff, in a simple decentralized college admissions market. The main result shows that the cutoff signals implement all implementable state-independent distributions of students across schools. In addition, we find that the cutoff signals have rich properties: All allocations implementable by cutoff signals and straightforward strategies are ex-ante fair. Moreover, these allocations are implementable as unique equilibrium outcomes. The applications include both multiple-sender persuasion and information design. Thus, the results indicate that equilibrium allocations are sensitive to priority uncertainty and could be undesirable without appropriate policy intervention. We also characterize implementable outcomes in celebrated centralized admissions systems.

*We thank Fuhito Kojima for providing helpful comments. We also acknowledge the financial support from the Japan Society for the Promotion of Science KAKENHI: Grant Number 22J10043 (Sato) and 22J20710 (Shirakawa).

†Graduate School of Economics, the University of Tokyo (email address: hirotosato@g.ecc.u-tokyo.ac.jp).

‡Graduate School of Economics, the University of Tokyo (email address: r.shirakawa0723@gmail.com).

Contents

1 Introduction

College admissions systems are enormous, and crucial markets whose outcomes affect the educational experiences and future job placements of the participants. Allocations are determined depending on students' application profiles to schools and the schools' priorities over the students. According to Chen and Kesten (2017), around 10 million high school students participate in the Chinese college admissions system each year. Many studies have explored college admissions markets because of their importance and size.

One notable but overlooked characteristic of college admissions markets is that students often face inevitable *priority uncertainty* in deciding which schools to apply to. This is mainly because students are often required to submit their applications either before taking entrance exams or without knowing their scores on these exams, which would allow them to determine their priorities. In Japan, for instance, students must submit college applications before decentralized university-specific exams are held. The college admissions systems in South Korea and in some cities in China also have this requirement. Thus, the existence of priority uncertainty may be a non-negligible element of college admissions systems because it could significantly affect students' applications and the resulting allocations.

Furthermore, college admissions markets have certain information structures on priorities. Among others, well-observed informative signals consist of cutoffs, which tell students whether their priorities are higher than a given threshold. These signals are provided by private sector organizations, schools, or governments. For example, private tutoring schools in Japan conduct mock exams/aggregate scores at the central exam, and then students receive five-grade evaluations. Most US universities impose a minimum threshold for grades and SAT scores for applications. Additionally, some cites in China inform individual students if their score on the national exam, the *Gaokao*, meets the cutoff for applications to top schools. We view these instances as information provision and model the pith of these information structures as *cutoff signals*.

Our primary goal is to answer the following question: *In college admissions market, what allocations can arise depending on information structures?* We model a simple decentralized college admissions market with priority uncertainty based on Abdulkadiroğlu, Che, and Yasuda (2011), in which a continuum of students have common ordinal preferences over finite schools. To reflect the real-life heterogeneity of students' exam performance, we further extend the base model so that the students are divided into finite groups we call *tiers*, where students with higher tiers have strictly higher priorities at any given state of the world. By contrast, students in the same group are uncertain about priority orderings and form beliefs over priorities given an information structure. Based on their beliefs, students submit single applications in our decentralized market, which determines allocation.

This decentralized market formulation follows Hafalir et al. (2018). They observe that the number of possible applications is highly limited in real-life decentralized systems because decentralized college-specific exams are often held on the same day. Further, even some centralized markets share this feature: Avery, Lee, and Roth (2014) report that before the 1994 reform, the South Korean college admissions market held the national exam *Sunun* on two days, and only one application was allowed per exam. Some cities in China, which has a centralized system, also allow for only a few applications. Hence, the decentralized market defined in our model can also be considered a real-life centralized market with a limited number of possible applications.

Focusing on state-independent distributions of students across schools, Theorem 1 shows that a *cutoff signal*, which indicates to students if they meet a given *cutoff* for each school, can implement a distribution whenever some information structure can. In other words, cutoff signals achieve all implementable state-independent distributions. Moreover, the cutoff signal can implement the outcome with *straightforward strategies*, wherein students apply to the most preferred school that exceeds the cutoff. Proposition 2 further proves that cutoff signals induce an allocation that is *ex-ante fair* and the closest to being *ex-post fair* among the allocations sharing the same distribution. These results indicate that cutoff signals inherit implementability, simplicity, and fairness, which might explain why in practice, information provision is usually conducted using cutoff signals.

The multiplicity of equilibrium allocations is relevant for policy-makers who send signals to achieve an intended allocation: It remains a possibility that another unintended equilibrium allocation is realized by an *adversarial equilibrium selection*. We therefore examine which information structures robustly implement which outcomes against adversarial equilibrium selections. Proposition 4 shows that, through *public* cutoff signals, policy-makers can implement *all* implementable state-independent distributions as unique equilibrium outcomes. Uniqueness of equilibrium implies the robustness of implementation against adversarial equilibrium selection, and thus Proposition 4 indicates that there is no trade-off between the set of feasible outcomes and robust implementation.

Our model has two applications: multiple-sender Bayesian persuasion and information design. One plausible interpretation of cutoff signals is that each school independently commits to their own cutoff, which then induces a cutoff signal. We first explore the property of equilibrium allocations in this multiple-sender Bayesian persuasion problem based on Gentzkow and Kamenica (2016, 2017). Assuming that each school selects a cutoff to maximize the number of applications, Proposition 5 indicates that equilibrium allocations could be far from being ex-post fair. In other words, competitions among schools for their own benefit may result in undesirable allocations, though their desirability depends on the context.

Another interpretation of cutoff signals is that a policy-maker sets cutoffs as a policy intervention. In particular, we consider a benevolent planner who seeks to maximize the aggregated welfare of those who have ex-ante priority below a given threshold. The objective reduces to the usual utilitarianism when the threshold is sufficiently high. If the low-priority students consist of those from minority groups, this objective function may also be seen as an affirmative action policy. Proposition 6 shows an optimal cutoff profile for these objectives and finds that lower thresholds lead to less ex-post fair matching outcomes. Thus, our result suggests that caution is needed for trade-offs between the welfare of low-priority students and the ex-post fairness of allocations.

Although our main analysis focuses on decentralized college admissions systems, we also characterize the implementable outcomes in celebrated centralized markets. As research on market design has grown, centralized matching mechanisms have been introduced in practice globally. The *deferred acceptance* mechanism of Gale and Shapley (1962) and the *top trading cycles* mechanism of Shapley and Scarf (1974) are celebrated examples of centralized direct mechanisms. Perhaps the most crucial common characteristic of these two mechanisms is that they are *strategy-proof*, that is, the honest revelation of true preferences is a weakly dominant strategy for students. Proposition 7 characterizes the implementable allocations under centralized markets with strategy-proof mechanisms, independent of information structures. In other words, information structures play no role in equilibrium allocations under strategy-proof mechanisms.

While many centralized school choice systems have started to employ strategy-proof mechanisms, some are still using manipulable mechanisms. One popular manipulable mechanism is the *Immediate Acceptance* mechanism, also referred to as the *Boston* mechanism. Although many studies since Abdulkadiroğlu and Sönmez (2003) have pointed out various shortcomings of the Boston mechanism, it remains one of the most prevalent practical mechanisms. Kojima and Ünver (2014) note that Minneapolis, Lee County in Florida, Denver, and Cambridge, Massachusetts employ or have employed the Boston mechanism. Thus, we also characterize the set of implementable outcomes under the Boston mechanism. Specifically, given a sufficient number of students relative to the number of schools, which is often the case, Proposition 8 finds that the set of implementable state-independent distributions under the Boston mechanism coincides with that under the decentralized system. In this sense, centralized college admissions systems adopting the Boston mechanism may not significantly differ from decentralized systems.

1.1 Related literature

This study is the first to explore the information structures of schools' priority structures and thus contributes to the literature on both matching markets and Bayesian persuasion/information design.

Most studies on college admissions and school choice, including the seminal works of Gale and Shapley (1962) and Abdulkadiroğlu and Sönmez (2003), assume that agents have complete information on their preferences and priorities. Recent studies have also examined college admissions models with preference uncertainty. Under the presence of uncertainty on the intensities of common ordinal preferences, Abdulkadiroğlu, Che, and Yasuda (2011) find that students may prefer the Boston mechanism to the Deferred Acceptance mechanism in equilibrium. Troyan (2012), Akbarpour et al. (2022), and Akyol (2022) also examine and compare the two mechanisms in school choice with similar settings. Moreover, Bade (2015), Harless and Manjunath (2015), Immorlica et al. (2020), Chen and He (2021), Artemov (2021), and Noda (2022) address endogenous information acquisition for own preferences and investigate mechanism performance. Dasgupta (2022) explores information design problem and shows that carefully designed information structures could improve students' welfare. By contrast, the focus of our analysis is the priority uncertainty that is present in most real-life college admissions markets.

Compared to preference uncertainty, priority uncertainty in school choice has rarely been focused on despite its significance. In the context of the Chinese college admissions system, Lien, Zheng, and Zhong (2017) and Chen (2018) examine application timing and the properties of the resulting equilibrium outcomes under priority uncertainty. Pan (2019) conduct an experiment and shows that pre-exam applications could result in many mismatches. Chen and Pereyra (2019) provide empirical support for priority uncertainty affecting students' applications. Since different application timing lead to different information structures, our study also contributes to the literature exploring a general class of information structures including no and full information.

Another strand that this paper contributes is the growing literature on Bayesian persuasion and information design, pioneered by Kamenica and Gentzkow (2011) for a single-receiver case.[1] Bergemann and Morris (2016) and Taneva (2019) extend the model to incorporate multiple receivers with strategic interactions. Arieli and Babichenko (2019) investigate public and private information structures with and without strategic interactions. Furthermore, many studies have explored economic applications. Alonso and Câmara (2016) and Chan et al. (2019) consider voting models with information disclosure, while Bergemann et al. (2022a) and Bergemann et al. (2022b) examine optimal information design problems in auctions. Another recent application constitutes consumer search, as researched by Au and Whitmeyer (2018), Choi, Kim, and Pease (2019), Dogan and Hu (2022), and Ke, Lin, and Lu (2022). Our model and those in Noda (2022) and Dasgupta (2022) apply Bayesian persuasion in matching markets; however, our study is novel,

[1]Kamenica (2019) and Bergemann and Morris (2019) survey literature on Bayesian persuasion and information design.

in that we characterize the set of implementable outcomes and examine multiple-sender Bayesian persuasion.

Finally, this study contributes to the literature on robust information design. Mathevet, Perego, and Taneva (2020) and Sandmann (2020) investigate the implementation of worst-equilibrium outcomes that is robust to adversarial equilibrium selection. Hoshino (2022) explores a stronger implementation, which we refer to as *unique implementation* and which we also consider here. For applications, Halac, Lipnowski, and Rappoport (2020, 2022), Moriya and Yamashita (2020), and Halac, Kremer, and Winter (2021) examine unique implementation in team work production. Other applications include bank runs and regime change, which are studied in Goldstein and Pauzner (2005), Goldstein and Huang (2016), Li, Song, and Zhao (2019), and Inostroza and Pavan (2022). The base games of these studies, except for Hoshino (2022), are both binary-action and supermodular, and Morris, Oyama, and Takahashi (2020, 2022a,b) provide a comprehensive analysis of robust implementation in general binary-action supermodular games. In contrast to previous studies, our study considers robust implementation beyond binary-action supermodular games. Moreover, unlike Hoshino (2022), who assumes the existence of the generalized *risk-dominant* strategies of Kajii and Morris (1997), students in our model have no such strategies.

The remainder of this paper is organized as follows. Section 2 describes the model, and the main results are presented in Section 3. We illustrate applications in Section 4, and in Section 5, we discuss centralized markets and assumptions in the model. Appendix A describes the formal model, while all proofs are provided in Appendix B.

2 Model

2.1 College admissions systems with priority uncertainty

There are a finite set of schools $C = \{c_1, \dots, c_{|C|}\}$ and a continuum of students $I \subset \mathbb{R}$ that is a closed interval. Each school has a unit mass of capacity. Let λ be the Lebesgue measure defined on I and assume $\lambda(I) < \infty$. An outside option for students is denote by $\varnothing$. For convenience, we sometimes express the outside option as $c_{|C|+1} = \varnothing$. A *matching* $\mu : I \to C \cup \{\varnothing\}$ is a measurable function that assigns each student to at most one school, with the capacity constraint $\lambda(\mu^{-1}(c)) \leq 1$ for each school $c \in C$. Let $\mathcal{M}$ be the set of all matchings.

Students are classified into finite types, which we call *tiers*. The finite set of tiers is denote by $T = \{1, \ldots, |T|\}$. Let $I_t \subset I$ be the set of students with tier $t \in T$ such that $\lambda(I_t) > 0$. The collection of the sets $\{I_t\}_{t \in T}$ is assumed to be a partition over the set of all students I. That is, each student belongs to exactly one group I_t for some $t \in T$.

Unlike classic school choice models, our model assumes that priority profiles are stochastic. For each tier $t \in T$, each student $i \in I_t$ draws their *priority profile* $\theta(i) \in [t-1, t]^{\mathrm{C}}$, which is uniformly independently distributed across students and schools. The priority of i at each school c is given by $\theta(i)(c)$, where smaller values represent higher priorities. The *state space* Θ is the set of all the priority profiles $\theta = (\theta(i))_{i \in I}$ that are induced according to the above procedure. Note that students with higher tiers (i.e., smaller $t \in T$) always have higher priorities for all states, which we assume for analytical simplicity. Students have the common uniform prior on Θ, denoted by $p \in \Delta(\Theta)$. See Appendix A for the formal definition of the common prior.

For each $t \in T$, students in group I_t have a common vNM utility function denoted by $u_t : \mathrm{C} \cup \{\varnothing\} \to \mathbb{R}$, where $u_t(c_1) > \cdots > u_t(c_{|\mathrm{C}|}) > u_t(\varnothing)$. That is, students have common ordinal preferences, while students from different groups may have different preference intensities. Normalize $u_t(\varnothing) = 0$.

2.2 Decentralized college admissions systems

Here, we model the decentralized college admissions procedure. The formulation follows Hafalir et al. (2018) and builds on their observation that decentralized systems often allow students to apply only to a very limited number of schools. For example, in Japan, each public university makes students take institution-specific exams in addition to the centralized exam, called the *Common Test*. These institution-specific exams are held on the same day, which essentially prevents students from applying to multiple public universities. Many other decentralized admissions systems in which schools have their own unique exams, such as in a part of the United Kingdom and Russia, share more or less similar structures, in that required admissions examinations for different schools are often held on the same day.

Considering these observations, we also model the decentralized system as a matching mechanism in which students apply to at most one school, and schools subsequently accept the most preferred set of candidates. Since we also discuss the case of college admissions systems adopting centralized mechanisms in Section 5, here, we introduce a range of notations for later uses.

Let $\mathcal{R}_i$ be the set of all strict orderings, *rankings*, over $C \cup \{\emptyset\}$, the schools or the outside option. A school c is *acceptable* at ranking R_i if $cR_i\emptyset$ holds. If a school c is the only acceptable school at R_i, we may with a slight abuse of notation sometimes write $R_i = c$. Let $\mathcal{R} = (\mathcal{R}_i)_{i \in I}$ be the set of all ranking profiles and $\mathcal{R}^1 \subset \mathcal{R}$ be the set of ranking profiles such that at most one school is acceptable for each student.

A *mechanism* $\varphi : \mathcal{R} \times \Theta \to \mathcal{M}$ is a function that maps ranking profiles and priority profiles to matchings. Here, we can define the *decentralized mechanism*. See Appendix A for the formal description. For each input of ranking profile R and priority profile θ, the decentralized mechanism, denoted by φ^D, outputs a matching according to the following procedure:

- Each student i applies to the most preferred school at R_i, if any. Among the set of applicants, each school c accepts students up to the unit capacity, where preference is given to students i with smaller values of $\theta(i)(c)$. All remaining students are rejected.

- Each student not accepted by any school matches with the outside option.

Note that the decentralized mechanism only considers the first choices of students' rankings. Therefore, while the domain of φ^D is defined as the set of all ranking profiles solely for consistency, our specification essentially prevents students from submitting applications to multiple schools. We assume the decentralized market $\varphi = \varphi^D$ except for in Section 5, where we analyze centralized markets with celebrated centralized mechanisms.

2.3 Information structures and equilibrium outcomes

As explored in the Introduction, in typical college admissions systems, each student receives some additional private information on priority profiles from some source. This study aims to study the implications of information structures on resulting allocations. This subsection models these information structures and resulting allocations.

An *information structure* is defined as a pair $\Pi = (S, \pi)$, where $S = (S_i)_{i \in I}$ is the product of the set of *signals*, and $\pi : \Theta \to S$ is a measurable function called a *disclosure rule*. Unlike the general framework of Bayesian persuasion as in Kamenica and Gentzkow (2011), we assume that the disclosure rule is deterministic. Since there are uncountable states in our model, a deterministic information structure does not imply that each student is either fully informed or not informed at all for each state, and thus there remains a non-trivial Bayesian persuasion problem.

An information structure $\Pi = (S, \pi)$ induces a preference revelation game with incomplete information, where for each student $i \in I$, a pure strategy of the student is a measurable function $\sigma_i : S_i \to \mathcal{R}_i$. A pure strategy profile $\sigma = (\sigma_i)_{i \in I}$ is a *pure Bayes Nash equilibrium (BNE)* if all students play their best replies under updated beliefs: For any $s_i \in S_i$, $t \in T$, $i \in I_t$, and $R_i \in \mathcal{R}_i$, we have

$$\int_{\Theta \times S_{-i}} p_i(\theta, s_{-i} \mid s_i)[u_t(\varphi(\sigma(s), \theta)) - u_t(\varphi((R_i, \sigma_{-i}(s_{-i})), \theta))] d\theta ds_{-i} \geq 0,$$

where $p_i(\cdot \mid s_i)$ is the regular conditional probability measure on $\Theta \times S_{-i}$ conditional on the private signal $s_i \in S_i$, which represents students' posterior beliefs given private signals. Since we focus on deterministic information structures, the above incentive condition can be rewritten as follows: For any $s_i \in S_i$, $t \in T$, $i \in I_t$, and $R_i \in \mathcal{R}_i$, we have

$$\int_{\Theta} p_i(\theta \mid s_i)[u_t(\varphi(\sigma(s), \theta)) - u_t(\varphi((R_i, \sigma_{-i}(s_{-i})), \theta))] d\theta \geq 0,$$

where $p_i(\theta \mid s_i) = p_i(\theta, \pi_{-i}(\theta) \mid s_i)$ for each $\theta \in \Theta$.

An information structure and a BNE induce an equilibrium outcome that we name *matching outcome*. Formally, a *matching outcome* is a measurable function $\mu : \Theta \to \mathcal{M}$, which is a state-contingent matching. We assert that a signal Π and a strategy profile σ *induce* a matching outcome μ if we have $\mu = \varphi \circ \sigma \circ \pi$.

While we explore the equilibrium matching outcomes in this study, we note that students in the same group are ex-ante homogeneous. Therefore, we also examine the structures of the equilibrium distributions of student type. Let $\xi : \mathcal{M} \to \mathbb{R}^{|C| \times |T|}$ be the function that defines the *distributions* of matchings: For each $k = 1, \ldots, |C|$ and $t \in T$, the distribution is defined as

$\xi_k^t(\mu) = \lambda(\mu^{-1}(c_k) \cap I_t)$, the mass of students with tier t matched with school c_k. In other words, $\xi(\mu)$ represents a state-contingent distribution of student types for each school. We say that a signal Π and a strategy profile σ *induce* a distribution $\xi : \Theta \to \mathbb{R}^{|C|\times|T|}$ if we have $\xi = \xi \circ \varphi \circ \sigma \circ \pi$. A distribution is *state-independent* if there exists a vector $\xi \in \mathbb{R}^{|C|\times|T|}$ such that the distribution equals ξ almost surely. We may often view the associated vectors as state-independent distributions.

Definition 1 (Implementation)**.** A matching outcome μ is *implementable* if there exists an information structure and a BNE that induce μ. An information structure Π and a strategy profile σ *implement* a matching outcome μ if σ is a BNE and Π and σ induce μ.

Likewise, we define implementable distributions. A distribution ξ is *implementable* if there exists an implementable matching outcome that induces ξ. An information structure and a strategy profile *implement* a distribution ξ if they implement a matching outcome inducing ξ.

3 Cutoff signals in decentralized admissions systems

In this section, we present our main results, particularly focusing on *cutoff signals* that capture the essence of real-life information structures. The first main result shows that cutoff signals are comprehensive: Theorem 1 indicates that all implementable state-independent distributions are implementable with cutoff signals. Additional properties of cutoff signals and the resulting allocations are also provided.

3.1 Cutoff signals

In reality, information structures often have somewhat similar structures. For example, for Japan's university entrance examinations, many private tutoring schools conduct mock exams and then provide students information on whether they have certain success probabilities using five-grade evaluations for students' chosen schools. Most US universities set a minimum threshold for high school grades and SAT scores for applications. This can be thought of as sending each student information regarding the competitive ratio, in addition to limiting the number of students who are eligible for applications.

The real-life signals of the aforementioned examples allow students to see if their priorities are higher than some predetermined cutoffs. The next concept, *cutoff signals*, defines a class of information structures that have these features.

Definition 2 (Cutoff signal)**.** We say that an information structure $\Pi = (S, \pi)$ is a *cutoff signal*, if there is a *cutoff vector* $e \in \mathbb{R}^{|C|}$ such that $S_i = 2^{C \cup \{\emptyset\}}$ and $\pi(\theta)(i) = \{c_l \in C \mid \theta(i)(c_l) \leq e_l\} \cup \{\emptyset\}$ for all $\theta \in \Theta$ and $i \in I$. Given a cutoff signal, the *straightforward strategy* of student $i \in I$ is the strategy of applying to the most preferred choice among the set s_i of alternatives, for each signal realization $s_i \in S_i$.

In other words, for each student, a cutoff signal privately informs whether they has a higher priority than the schools' cutoffs. Which agents set cutoff signals depends on the context, which has not yet been described in the model. In Section 4, we construct two extended models that include agents who select information structures seeking their own interests and analyze the consequences therein.

3.2 Characterization

The first main result of this study is stated as follows. The cutoff signals, which are well-observed in practice, are comprehensive, in that they can implement all implementable state-independent distributions. Hence, this result provides a rationale for practical information structures. Furthermore, it is sufficient to recommend the straightforward strategies for implementation, which are simple strategies that students can easily follow.

Theorem 1 (Characterization)**.** *For any implementable state-independent distribution ξ, there exist a cutoff signal and the straightforward strategy profile implementing ξ.*

In the proof, we provide another characterization for the set of all state-independent distributions that are implementable with cutoff signals. From Theorem 1, this result characterizes the set of all implementable state-independent distributions, as well. We say that a state-independent distribution $\xi \in \mathbb{R}^{|C| \times |T|}$ is *non-wasteful*, if for any $k > 1$, $\sum_t \xi_k^t > 0$ implies $\sum_t \xi_{k-1}^t = 1$.

Proposition 1 (Lemma 8)**.** A cutoff signal and the straightforward strategy profile implement a state-independent distribution $\xi \in \mathbb{R}^{|C|\times|T|}$, if and only if, ξ is non-wasteful and there exists a labeling function $k : T \to \{1, \ldots, |C|\}$ that is increasing in $t \in T$ with $k(0) \equiv 1$, such that, for each $t \in T$,

$$\{c_k \in C \mid \xi_k^t > 0\} \subset \{c_k \in C \mid k(t-1) \leq k \leq k(t)\},$$

and, for each $t \in T$, we have

$$\sum_{k=k(t-1)}^{k(t)} \xi_k^t \cdot u_t(c_k) \geq \lambda(I_t) \cdot u_t(c_{k^*(t)}),$$

where $k^*(t) = k(t) + 1$ if $\xi_{k(t)}^t = 1$ and $k^*(t) = k(t)$ otherwise.

In other words, implementable state-independent distributions have two features in addition to non-wastefulness. First, the set of schools assigned to each group has a monotone structure so that for $h < t$, almost no student in group I_t matches with a school preferred to a school with which a student in group I_h matches. This is an intuitive structure, recalling that students in I_h always have higher priorities than those in I_t at all schools. Second, the aggregated ex-ante payoff to students in I_t must exceed the aggregated payoff when all of them match with a school that are filled with students with lower tiers. This follows from the equilibrium condition: If this inequality is violated, we can easily show that a student in I_t has an incentive to deviate from the straightforward strategy to one in which they applies to a less preferred school that the student can match with probability one.

3.3 Fairness properties

In this subsection, we present some additional properties of the induced outcomes under the cutoff signals. We focus on the implications of implementing cutoff signals on *fairness properties* of the resulting matching outcomes. Since our model differs from the classic college admissions models due to the existence of incomplete information, it is appropriate to examine two fairness concepts, *ex-post* and *ex-ante* fairness.

First, we define ex-post fair matching outcomes. Given a matching outcome μ, student $i \in I_t$

has an *ex-post justified envy* at θ, if there exists a student j and $c = \mu(j)$ such that $\theta(i)(c) < \theta(j)(c)$ and $u_t(c) > u_t(\mu(\theta)(i))$. Let $X(\mu) \subset I \times \Theta$ denote the maximal jointly measurable set such that $(i, \theta) \in X(\mu)$ implies i having an ex-post justified envy at θ.[2] Then, for each matching outcome μ, define the expected mass of students having an ex-post justified envy at μ:

$$E(\mu) = \int_\Theta \int_I \mathbb{I}(X(\mu)) d\lambda dp,$$

where $\mathbb{I}$ denotes the indicator function. A matching outcome μ is *ex-post fair* if $E(\mu) = 0$. We say that μ is *no less ex-post fair than* μ' if $E(\mu) \leq E(\mu')$, and *less ex-post fair than* μ' if $E(\mu) > E(\mu')$.

The definition of ex-ante fairness builds on the following observations: Students in a same group are ex-ante homogeneous at all schools and for two students $i \in I_t$ and $j \in I_h$ with $t < h$, i always has a strictly higher priority than j does at all schools. We say that a matching outcome μ is *ex-ante fair* if for any $i \in I_t$ and $j \in I_h$ with $t < h$, we have

$$\int_\Theta u_t(\mu(\theta)(i)) dp \geq \int_\Theta u_t(\mu(\theta)(j)) dp.$$

That is, with the two observations, a matching outcome is ex-ante fair if no student has an *ex-ante justified envy*.

The following proposition provides an additional property of cutoff signals regarding implementable matching outcomes. Provided that one seeks to implement a given state-independent distribution, cutoff signals are the optimal information structure to achieve the intended distribution in that the induced matching outcome is the closest to being ex-post fair. Moreover, the induced matching outcome is guaranteed to be ex-ante fair, as well. These results reveal the potential benefits of the practical information structures.

Proposition 2 (Fairness)**.** If a cutoff signal and the straightforward strategies implement a matching outcome $\mu^* : \Theta \to \mathcal{M}$ that induces a state-independent distribution ξ, then for any matching outcome μ that induces ξ, the matching outcome μ^* is no less ex-post fair than μ. Moreover, μ^* is ex-ante fair.

[2] Although this is just a technical remark, the σ-algebra over the product space $I \times \Theta$ is defined differently from the usual product σ-algebra, so that our proofs may rely on the *exact law of large numbers* of Sun (2006). See Appendix A for details.

The next result concerns one implication of the selection of cutoff vectors on resulting matching outcomes. According to Proposition 2, all equilibrium matching outcomes are ex-ante fair. Meanwhile, they are in general not ex-post fair. The following proposition states that loose cutoff vectors might lead to matching outcomes that are less ex-post fair. Note, however, that it does not necessarily mean tight cutoff vectors are always preferred, as it depends on the objectives of those who set the cutoff policy. We discuss optimal cutoff signals in Section 4.

Proposition 3 (Comparative statics)**.** Suppose that two cutoff signals with cutoff vectors e and e' and the straightforward strategy profile implement μ and μ', respectively. If $e \geq e'$, then μ' is no less ex-post fair than μ.

3.4 Unique implementation

This subsection describes unique implementable matching outcomes. Suppose a benevolent planner is about to choose an information structure to implement a desired outcome.[3] Then, our definition of *partial* implementation does not guarantee that the intended outcome is indeed achieved because there are potentially other equilibria inducing different outcomes: Partial implementation is not robust to *adversarial equilibrium selection*. To overcome this issue, existing studies have considered a natural refinement, *unique implementation*, which ensures all equilibria induce an outcome arbitrarily close to the intended one.

Intuitively, a matching outcome μ is *uniquely implementable* if there exists a matching outcome μ' arbitrarily close to μ such that an information structure induces a Bayesian game in which all equilibria induce μ'. For two matching outcomes μ and μ', define a distance-like value between the two,

$$\rho(\mu, \mu') = \int_{\theta \in \Theta} \int_{i \in I} \mathbb{I}(\mu(\theta)(i) \neq \mu'(\theta)(i)) d\lambda(i) dp(\theta),$$

as the aggregated mass of students whose placements differ between the two matching outcomes. Unique implementation is now formally defined as follows.

Definition 3 (Unique implementation)**.** A matching outcome μ is *uniquely implementable* if for any

[3]We discuss optimal information structures in this scenario in Section 4.2.

$\varepsilon > 0$, there exist a matching outcome μ' with $\rho(\mu, \mu') < \varepsilon$ and an information structure under which all BNE induce the matching outcome μ'. We also say that the distribution $\xi \circ \mu$ is *uniquely implementable* if μ is uniquely implementable.

In the following proposition, we prove that some information structures uniquely implement *all* matching outcomes that are implementable with cutoff signals and the straightforward strategies. Combined with Theorem 1, this characterizes all uniquely implementable state-independent distributions.

Proposition 4 (Robustness)**.** All matching outcomes that are implementable with cutoff signals and the straightforward strategies are uniquely implementable. Consequently, all implementable state-independent distributions are uniquely implementable.

The proof constructs a variant of the cutoff signals, namely a *public* cutoff signal with a state-dependent cutoff vector. Suppose that a cutoff signal and the straightforward strategies implement a matching outcome. Then, at each state, we can detect the mass of students applying to each school. The information structure in the proof publicly announces a recommendation profile at each state, which is constructed inductively as follows: For each school c_k, among students who are not recommended some schools preferred to c_k, recommend the school c_k to those who have the highest priorities at the school so that the mass of students recommended c_k equals the targeted mass of students applying to c_k. We prove that, for generic cases, the equilibrium of students following these recommendations induces the unique equilibrium outcome, which coincides with the targeted matching outcome.

It should be emphasized that the set of rationalizable outcomes is not unique in the public cutoff signals considered in our proof. In particular, there are no weakly dominated strategies except for a trivial one that applies to the outside option. Technically, this is the first study on an economic model to find that *all* implementable state-independent distributions are uniquely implementable, *without* recourse to the iterated eliminations of dominated strategies.

4 Applications

In this section, we present the possible economic implications of our main theorems. The following subsections focus on multi-sender Bayesian persuasion settings based on Gentzkow and Kamenica (2016) and information design from Bergemann and Morris (2016) and Taneva (2019): Each setting respectively assumes the existence of multiple senders and a single policy-maker who commit to an information structure. To begin, we define a class of implementable state-independent distributions in preparation, which plays a significant role in the subsequent analysis.

Recall that Proposition 1 (Lemma 8) characterizes the implementable state-independent distributions ξ in terms of a labeling function that is increasing in type. Specifically, the proof of Lemma 8 defines the labeling function $k(\xi)$ such that each $k(\xi)(t)$ is the maximal number k such that $\xi_k^t > 0$, which is uniquely pinned down. For each type $t \in T$, we construct implementable state-independent distributions according to the next class of algorithms that are parameterized by $t \in T$, which works as follows. Let $\mathcal{E}$ to be the set of all implementable state-independent distributions.

- Step 1. Let $\underline{\mathcal{E}}$ and $\overline{\mathcal{E}}$ be the set of distributions ξ in $\mathcal{E}$ that minimize and maximize $k(\xi)(1)$, respectively. If $t = 1$, define $\mathcal{E}(1)$ as the set of distributions maximizing $\xi^1_{k(\xi)(1)}$ among $\overline{\mathcal{E}}$. Otherwise, define $\mathcal{E}(1)$ as the set of distributions minimizing $\xi^1_{k(\xi)(1)}$ among $\underline{\mathcal{E}}$.

- Step h, $h \geq 2$. Let $\underline{\mathcal{E}}$ and $\overline{\mathcal{E}}$ be the set of distributions ξ in $\mathcal{E}(h-1)$ that minimize and maximize $k(\xi)(h)$, respectively. If $h \geq t$, define $\mathcal{E}(h)$ as the set of all distributions in $\overline{\mathcal{E}}$ that maximize $\xi^h_{k(\xi)(h)}$ among $\overline{\mathcal{E}}$. Otherwise, define $\mathcal{E}(h)$ as the set of distributions in $\underline{\mathcal{E}}$ minimizing $\xi^h_{k(\xi)(h)}$ among $\underline{\mathcal{E}}$.

- The above algorithm ends in T steps. It is not difficult to see that $\mathcal{E}(T)$ must be a singleton. Finally, denote $\xi(t) \in \mathcal{E}(T)$.

It is evident from Lemma 8 that this class of algorithms is well-defined because the set of implementable state-independent distributions is closed with a natural topology. Intuitively, for each parameter $t \in T$, the above algorithm outputs the implementable state-independent distribution so that students I_h with $h < t$ face the severest competition while students I_h with $h \geq t$ face the mildest competition for each school.

4.1 Multiple-sender Bayesian persuasion: revenue maximizing schools

One plausible interpretation of the cutoff signals that were introduced in Subsection 3.1 is that each school independently commits to their own cutoff without communicating with each other, which then induces a cutoff signal. In this subsection, we explore a property of equilibrium outcomes in this multiple-sender Bayesian persuasion problem.

Assume that each school c_k commits to a cutoff e_k so as to maximize the expected mass of applicants. They do this to maximize their expected revenue from application fees or to simply improve their reputation, for example.

The two-stage game explored here proceeds as follows: In the first stage, each school $c_k \in C$ simultaneously chooses and commits to a cutoff $e_k \in \mathbb{R}$. Immediately after the first stage, nature draws a state $\theta \in \Theta$ following the prior p. In the second stage after the nature draws a state, each student $i \in I$ simultaneously chooses which school to apply, only by observing the cutoff profile $e \in \mathbb{R}^{|C|}$ and a private signal realization $\pi(\theta)(i)$.

Let M^* be the set of all matching outcomes implementable with cutoff vectors and the straightforward strategies. We say that $\mu \in M^*$ is the *least ex-post fair matching outcome* in M^* if μ' is no less ex-post fair than μ for any $\mu' \in M^*$. To put it nicely, the least ex-post fair matching outcome may be the most diverse allocation, but it may simultaneously be the least meritocratic allocation.

To avoid non-essential complexity, our analysis focuses on *weak Perfect Bayesian equilibrium* as an equilibrium notion. Furthermore, we focus on equilibria such that students never update their beliefs and play strategies that do not depend on signal realizations off the path of plays. The next proposition indicates that, if each school pursues its own interests, the resulting outcome is the least ex-post fair matching outcome, which may not be preferred depending on the context.

Proposition 5 (Application-maximizing schools)**.** Consider the above multiple-sender two-stage game. Then, the least ex-post fair matching outcome in M^* is an equilibrium outcome. Conversely, for any equilibrium where students play the straightforward strategies on the path of plays, the equilibrium outcome is the least ex-post fair matching outcome in M^*.

4.2 Information design: welfare maximization

Another possible scenario is that a single policy-maker, for example a government, commits to a cutoff vector and the associated cutoff signal in an attempt to maximize their objective as modeled in information design problems in Bergemann and Morris (2016) and Taneva (2019). Here, we examine an optimal cutoff signal of a policy-maker with a simple objective of generalized utilitarian sum. Throughout this subsection, let us assume that all students have a common cardinal utility $u : C \cup \{\emptyset\} \to \mathbb{R}$, that is, we have $u = u_t$ for all $t \in T$.

For each cutoff vector $e \in \mathbb{R}^{|C|}$, let $\mu(e)$ and $\xi(e)$ be the induced matching outcome and the distribution under the straightforward strategies, respectively. We assume that the planner commits to a cutoff signal to maximize an expected weighted utilitarian sum at $\mu(e)$, subject to the constraint that the straightforward strategies constitute a BNE.

Fix a type $t^* \in T$. Formally, consider a single policy-maker choosing a cutoff signal that solves the following optimization problem:

$$\begin{aligned}\max_{e \in F} V(e) &= \int_{\theta \in \Theta} \left(\sum_{t \geq t^*} \int_{i \in I_t} u(\mu(e)(\theta)(i)) d\lambda(i) \right) dp(\theta) \\ &= \int_{\theta \in \Theta} \left(\sum_{t \geq t^*} \sum_{k=1}^{|C|} \xi_k^t(e)(\theta) \cdot u(c_k) \right) dp(\theta),\end{aligned}$$

where F is the set of all cutoff vectors such that the straightforward strategy profile is a BNE under the associated cutoff signal. In other words, they seeks to maximize the utilitarian sum of those who have ex-ante lower priorities than a predetermined threshold, subject to the constraint that students play the straightforward strategies as an equilibrium. In the case $t^* = 1$, the objective function reduces to the usual utilitarianism.

In general, we may interpret the objective as a preference for an affirmative action policy, provided that the low-priority groups consist of students classified as a minority group. For example, social planners may give favorable treatment to those with incomes below the poverty line. Low-income students tend to have low priority, and thus we may interpret that tiers are determined by income level. Unlike previous studies on affirmative action such as Kojima (2012) and Hafalir, Yenmez, and Yildirim (2013), the planner here does not design priorities; instead, they chooses an information structure. Recall that Proposition 2 ensures that any resulting matching outcome is ex-ante fair, while a matching outcome arising from a priority design might not be.

The following proposition shows an optimal cutoff signal. From Theorem 1, the induced state-independent distribution maximizes the objective among the set of all implementable state-independent distributions. Further, according to Proposition 4, we may instead consider a public cutoff signal with a state-dependent cutoff vector for implementation, which enables a unique implementation that is robust to adversarial equilibrium selection. Thus, there is no trade-off between robust implementation and the planner's payoff. Instead, the trade-off is between the planner's objective and the degree of ex-post fairness in optimal matching outcomes.

Proposition 6 (Generalized utilitarianism)**.** A cutoff vector $e(t^*)$ such that $\xi(e(t^*)) = \xi(t^*)$ is a solution to the optimization problem. Moreover, the matching outcome $\mu(e(t))$ is less ex-post fair than $\mu(e(t-1))$ for all $t > 1$.

5 Discussions and Extensions

5.1 Centralized college admissions: characterizations

Since the emergence of market design, many school districts have centralized market-clearing procedures and have adopted non-manipulable centralized mechanisms. The *Deferred Acceptance* mechanism of Gale and Shapley (1962) and the *Top Trading Cycles* mechanism of Shapley and Scarf (1974) are favored practical examples of *strategy-proof* mechanisms, among others. Here, we briefly discuss the implementable outcomes under these celebrated mechanisms.

A mechanism φ is said to be *strategy-proof* if for any student $i \in I_t$, the honest revelation of their true preference is a weakly dominant strategy, that is,

$$u_t(\varphi(R_i^*, R_{-i}, \theta)) \geq u_t(\varphi(R_i, R_{-i}, \theta))$$

for all $R_{-i} \in \mathcal{R}_{-i}$, $R_i \in \mathcal{R}_i$, and $\theta \in \Theta$, where we denote by R_i^* the true preference ordering. Note that the definition directly implies that the honest report is weakly dominant irrespective of students' belief on the space of priority profiles. Consequently, the set of implementable distributions is very restricted when compared to the decentralized systems. In particular, the next proposition implies that the selection of information structures has no effect on the resulting outcomes. We say that an information structure $\Pi = (S, \pi)$ is *fully-revealing* if $S_i = \Theta$ and $\pi(\theta)(i) = \theta$ for all $i \in I$ and $\theta \in \Theta$.

Proposition 7 (Strategy-proof mechanisms)**.** Let φ be strategy-proof. Then, a matching outcome μ is implementable under φ only if there exists an ex-post equilibrium $\sigma^* : \Theta \to \mathcal{R}$ under the fully-revealing information structure that induces μ almost surely. In addition, if a matching outcome is uniquely implementable under φ, it is induced by the honest reports.

Our proof for Proposition 7 does not rely on most of our detailed modeling assumptions. First, the conclusion still holds even if we allow for heterogeneous ordinal preferences for students and non-uniform distributions of priorities. Second, it is evident that the continuity of students also play no role; therefore, the same conclusion follows in the classic school choice models of Abdulkadiroğlu and Sönmez (2003) with priority uncertainty. Finally, note that the essential converse is true by definition. That is, if an ex-post equilibrium under the fully-revealing information structure induces a matching outcome, then it is implementable. In this sense, Proposition 7 is a general characterization result for the set of equilibrium outcomes under strategy-proof mechanisms in school choice models with priority uncertainty.

While many centralized school choice systems employ strategy-proof mechanisms, some still use manipulable mechanisms. Perhaps the most popular is the *Immediate Acceptance* mechanism, which is also referred to as the *Boston* mechanism. Here, we explore the implementable outcomes under the Boston mechanism.

Formally, the Boston mechanism, which we denote by φ^B, selects the output of the following *Boston algorithm* for each ranking profile and priority profile input.

- Step 1. Each student applies to their most preferred acceptable school at the reported ranking if there is any. Among the applicants, each school accepts students following the given priority up to the unit capacity. All remaining students are rejected.

- Step k, $k \geq 2$. Each student rejected in the previous step applies to their next-most preferred acceptable school at the reported ranking, if there is any. Among the applicants, each school additionally accepts students following the given priority up to the remaining capacity. All remaining students are rejected.

The Boston algorithm ends at the step in which no student is rejected. All students who match with no school match with the outside option.

Provided that there are sufficiently many students relative to the number of schools, which could often be the case, we find that the set of implementable state-independent distributions under the Boston mechanism coincides with that under the decentralized market.

Proposition 8. Suppose that $T > |C|$ and $\lambda(I_t) > 1$ for all types $t \in T$. Then, a state-independent distribution $\xi \in \mathbb{R}^{|C|\times|T|}$ is implementable under the Boston mechanism φ^B if and only if it is implementable under the decentralized mechanism φ^D. Consequently, all implementable state-independent distributions are implementable with cutoff signals and uniquely implementable.

The intuition for this proposition is simple. If there are sufficiently many students as assumed, then for any information structure and any equilibrium strategy profile, there exists a group of students such that all students in the group match with no school almost surely. Therefore, the Boston algorithm must terminate in the first round almost surely, or a student with zero interim payoff has a profitable deviation. Consequently, the induced matching outcomes under the Boston algorithm are equivalent with those under the decentralized market in which students apply to at most one school. The latter statements then follow from Theorem 1 and Proposition 4.

5.2 Common cardinal preferences within groups

It is assumed throughout this study that all students in one group are completely ex-ante homogeneous, in the sense that they have ex-ante symmetric priorities and common cardinal preferences.

The next example shows that Theorem 1 no longer holds if they have heterogeneous preference intensities. That is, if students in a group have different cardinal utility functions, there exists an implementable distribution that cannot be implemented with a cutoff signal and the straightforward strategies.

Example 1. Suppose that there are three schools, c_1, c_2, and c_3, and mass $\lambda(I) = 3$ of students such that $I = I_1$. Unlike the model described in Section 2, let us assume now that students I_1 are further divided into two subsets: A and B with $A \cup B = I$. Suppose $\lambda(A) = 2$ and $\lambda(B) = 1$. Returning to the original setting, the decentralized mechanism φ^D is employed.

A student in A and a student in B have different preference intensities. Denote by u_A and u_B utility functions of each student in A and B, respectively. Let $u_A(c_1) = 4 > u_A(c_2) = 2$ and $u_B(c_1) = 3 > u_B(c_2) = 2$. Assume that $u_A(c_3) = u_B(c_3) = 1.9$ and $u_A(\varnothing) = u_B(\varnothing) = 0$. Students in group A have higher valuations for school c_1.

Now, consider an information structure $\Pi = (S, \pi)$ that discloses nothing, namely $|S| = 1$. Students' strategies then reduce to simply choosing single applications. Then, consider the strategy profile in which students in A apply to school c_1 and students in B apply to school c_2. Simple calculations find that the expected payoff to each student in A and B equals $4/2 = 2$ and $2/1 = 2$, respectively, for this strategy profile. Further, noting that all students here have ex-ante symmetric priorities, if a student in A (resp. B) instead applies to c_2 (resp. c_1), then the expected payoff to the student becomes 2 (resp. $3/2$). Therefore, the specified strategy profile is an equilibrium. Note that no student will be placed at school c_3 with probability one.

We show that no cutoff signals with the straightforward strategies implement the above equilibrium outcome. Take any cutoff vector e. Under the straightforward strategies, we show in the appendices that the mass of applicants to each school is constant at almost all states $\theta \in \Theta$, and is given by

$$n_k(e) = \lambda(\{i \in I \mid \theta(i)(c_k) \leq e_k \wedge \theta(i)(c_l) > e_l, \forall l < k\}),$$

for each $k = 1, 2, 3$. Note that we must have $n_3(e) = 0$ to achieve the desired state-independent distribution. Since it must hold that $n_1(e) + n_2(e) + n_3(e) = \lambda(I)$ in equilibrium, we have $n_1(e) +$

$n_2(e) = 3$. Below, we confirm that there is no such cutoff vector.

Suppose on the contrary that the straightforward strategies constitute an equilibrium. Note that all students have ex-ante symmetric priorities. Therefore, each student in B receives a signal that includes both c_1 and c_2 with strictly positive probability. The equilibrium condition then implies $3/n_1(e) \geq 2/n_2(e)$, and thus $n_2(e) \geq 6/5$. Meanwhile, consider a student and a signal in which they plays c_2. Since $2/n_2(e) \leq 5/3 < 1.9$, this student has a profitable deviation c_3, which contradicts the incentive condition.

The above example shows that an effective utilization of preference heterogeneity might expand the set of implementable distributions compared to cutoff signals combined with the straightforward strategies, which treat students in one group symmetrically. By contrast, preference intensities are not observable in most practical scenarios. It remains open, extending the model so that any student may have heterogeneous intensities, whether cutoff signals together with the straightforward strategies implement all implementable *symmetric* matching outcomes (by symmetric, we mean that allocations to students in the same group are ex-ante symmetric).

5.3 Stochastic information structures

While the information structures considered in this study are restricted to those that send one signal for each state, i.e., deterministic, standard models of Bayesian persuasion/information design may allow for those that disclose randomized signals. The following example demonstrates that with such *stochastic* information structures, one may implement a broader class of outcomes.

Example 2. Suppose that there are two schools, c_1 and c_2, and two groups, I_1 and I_2, with $I_1 \cup I_2 = I$, such that $\lambda(I_1) = 2$ and $\lambda(I_2) = 1$, respectively. Let all students have a common cardinal utility function, $u(c_1) = 3 > u(c_2) = 2 > u(\varnothing) = 0$.

Now, we construct an information structure $\Pi = (S, \pi)$ that directly recommends to students which schools to apply to. For each state $\theta \in \Theta$, let $I_1(\theta) \subset I_1$ be the subset of students $i \in I_1$ such that $\theta(i)(c_1) < 1/2$. Then, the disclosure rule π is defined as follows: For each state θ, the nature first draws a point ω from a uniform distribution over the closed interval $[0, 1]$. If $\omega \leq 2/5$, all students i in I_1 (resp. I_2) receive $\pi(\theta)(i) = c_1$ (resp. $\pi(\theta)(i) = c_2$). If $\omega \geq 4/5$, all students i in I_1 (resp. I_2) receive $\pi(\theta)(i) = c_2$ (resp. $\pi(\theta)(i) = c_1$). Otherwise, for each $i \in I_1$, the student receives c_1 if $i \in I_1(\theta)$ and c_2 if $i \notin I_1(\theta)$. Students in I_2 receive $\varnothing$. According to the construction, the most preferred available option is recommended to students in I_2, provided that all students in I_1 follow the recommendations. The information structure constructed here is summarized in Figure 1:

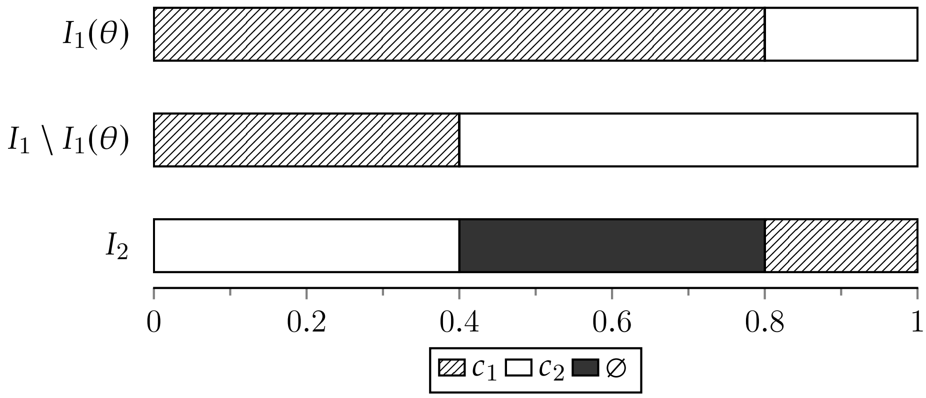

Figure 1: Stochastic private information structure

Notes: The horizontal line represents the values of ω. The top bar represents recommendations for students in $I_1(\theta)$, the middle bar for students in $I_1 \setminus I_1(\theta)$, and the bottom bar for students in I_2, where the following color code is used: school c_1 = lined, school c_2 = white, $\varnothing$ = solid black.

Under the specified information structure, the *obedient strategies*, $\sigma_i(s_i) = s_i$ for all signals $s_i \in S_i$, constitute a BNE. To verify this, suppose that all students follow the obedient strategies. It follows from Lemma 2 in Appendix A that the mass of students $I_1(\theta)$ equals $2/2 = 1$ for almost all $\theta \in \Theta$. Therefore, for each student $i \in I_1$, given the signal $s_i = c_1$ (resp. $s_i = c_2$), the interim expected payoff from the obedient strategy is given by 2 (resp. $3/2$), while the deviation to playing c_2 (resp. c_1) leads to the interim payoff of 2 (resp. $3/2$). In other words, all students in I_1 are indifferent between following the recommendations and not following them. Since it is obviously optimal for students in I_2 to follow the recommendations, the obedient strategy profile is an equilibrium.

In the obedient strategies under the above information structure, students in I_2 match with the most preferred school c_1 with positive probability at almost all states. It is worth noting that, from Lemma 8 and $\lambda(I_1) > 1$, no cutoff signal could match a student in I_1 with the most preferred school c_1. Moreover, no deterministic signal can implement a state-independent distribution that matches a positive mass of students in I_2 to c_1 almost surely. In fact, if such an information structure and a strategy profile exist, all students in I_1 could be almost sure that they could match with c_1 at any given signal. This is a contradiction because $\lambda(I_1) > 1$.

This example indicates that, using stochastic disclosure rules, we may match students having ex-ante low priorities with the most preferred school. The intuition behind this phenomenon is as follows. If recommendations are randomized and correlated, it is possible that all students in I_1 apply to the same second-preferred school c_2 with some positive probability, *without violating incentive compatibility*. This contrasts with implementable distributions specified in Lemma 8, which exhibit a monotone structure such that no student in a lower tier obtains an better assignment than those who belong to a higher tier and match with a school. It is an open question to characterize the set of distributions that are implementable with the general (stochastic) information structures.

Appendices

A Formal model

A.1 The state space and the product space

To incorporate heterogeneous priority uncertainty, we have considered the following situation: For each student in a group I_t, a priority profile is drawn according to the identical and independent uniform distribution over $[t-1, t]^C$. Here we present the formal model of the probability space.

The model description itself does not change: the set of students I and the state space Θ are equipped with the usual (product of) Borel σ-algebra, $\mathcal{B}$ and $\mathcal{B}^I$, and uniform measures λ and p, respectively. Rigorously speaking, the probability measure p over $\Theta = ([0, T]^C)^I$ is defined according to the standard Kolmogorov Extension Theorem, where for each coordinate $i \in I_t$ the probability measure p_i over the set of score profiles $[0, T]^C$ is the uniform distribution over the subset $[t-1, t]^C$. By the construction of the product measure p, it has the following intuitive property. The proof is trivial and thus be omitted.

Lemma 1. For any $t \in T$, $i \in I_t$ and $c \in C$, we have $p(\{\theta \in \Theta \mid \theta(i)(c) \leq e\}) = e - (t-1)$ for all $e \in \mathbb{R}$ with $t - 1 \leq e \leq t$.

What is unusual is the way we define a product measure space over $I \times \Theta$ so that it has additional intuitive and indispensable features. Sun (2006) proves, in a model that essentially includes our setting, the existence of a probability space $(\Theta, \mathcal{F}, p)$ whose σ-algebra extends that of $(\Theta, \mathcal{B}^I, p)$, and the product space

$$(I \times \Theta, \mathcal{B} \boxtimes \mathcal{F}, \lambda \boxtimes p),$$

where this probability space is called a *Fubini extension* of the usual product space. As the name suggests, all measurable functions in this space satisfy the statement of Fubini's Theorem. Moreover, the independent coordinate functions $f_i : \Theta \to \mathbb{R}$, defined as $f_i(\theta) = \theta(i)$, are all measurable functions. With this formulation, we get another crucial property for this extended probability space, which is referred to as *the Exact Law of Large Numbers*:

Lemma 2. For any $t \in T$, $c \in C$, and $e \in \mathbb{R}$ with $t - 1 \leq e \leq t$, for almost all $\theta \in \Theta$, it holds that $\lambda(\{i \in A_t \mid \theta(i)(c) \leq e\}) = \lambda(A_t) \cdot [e - (t-1)]$, where A_t is any measurable subset of students I_t.

Proof. This is almost a corollary of the results stated in Sun (2006), though we put the proof for our setting. Since the Fubini property holds in his extended product space, letting $\mathbb{I}$ be the indicator function, we have

$$\begin{aligned}
\lambda(\{i \in A_t \mid \theta(i)(c) \leq e\}) &= \int_{i \in I} \mathbb{I}(\{i \in A_t \mid f_i(\theta)(c) \leq e\}) d\lambda(i) \\
&= \int_{(i,\theta) \in I \times \Theta} \mathbb{I}(\{(i, \theta) \in A_t \times \Theta \mid f_i(\theta)(c) \leq e\}) d(\lambda \boxtimes p(i, \theta)) \\
&= \int_{i \in I} \left(\int_{\theta \in \Theta} \mathbb{I}(\{(i, \theta) \in A_t \times \Theta \mid \theta(i)(c) \leq e\}) dp(\theta) \right) d\lambda(i) \\
&= \lambda(A_t) \cdot [e - (t-1)],
\end{aligned}$$

where the second equality follows from Theorem 2.8. of Sun (2006) for almost all $\theta \in \Theta$ (one of the law of large numbers in his paper), the third equality follows from the Fubini's Theorem, and the last transformation follows from Lemma 1. $\square$

Intuitively, the *exact law of large numbers* implies that the fraction e of the students draws a priority no greater than e on each school, although slight modifications of this intuition are needed as in the statement due to the ex-ante heterogeneity of students.

A.2 The decentralized mechanism

In the decentralized mechanism φ^D, each school accepts the students up to the unit capacity according to the priority order. One problem is that there are possibly many ways to choose the set of accepted applicants, because there is no atom.

The schools' choices of the accepted students in our model are formally described as follows: Let $A \subset I$ be the set of applicants to a school c. If $\lambda(A) \leq 1$, the school accepts all the applicants. Otherwise, the school accepts the students in the set

$$\inf\{B \subset A \mid \lambda(B) \geq 1, \theta(i)(c) < \theta(j)(c), \forall i \in B, \forall j \in A \setminus B\},$$

where we let $\inf \emptyset = \emptyset$, and reject the remaining students. That is, each school accepts the minimal set of students up to the unit capacity, following the priority order. The formal description of the Boston mechanism φ^B also follows this choice rule.

For example, if the set of applicants to a school c is the set I of all students, the school accepts the students i with $\theta(c)(i) < \overline{e}$ and rejects the rest, where $\overline{e}$ is the unique solution e to the equation

$$\lambda(\{i \in I \mid \theta(c)(i) < e\}) = 1,$$

which exists if $\lambda(I) > 1$ and independent of θ almost surely from Lemma 2.

B Proofs

B.1 Proof of Theorem 1

In this section, we prove Theorem 1, which states that all implementable state-independent distributions are implementable with cutoff signals and the straightforward strategies. We begin with a technical preliminary proposition, which is also referred in the proof of Proposition 8. For this purpose, the next lemma is stated independent of underling mechanisms, while the rest of the proof assume implicitly the decentralized mechanism φ^D.

Lemma 3. Suppose that a state-independent distribution $\xi \in \mathbb{R}^{|C|\times|T|}$ is implementable under a mechanism φ with an information structure $\Pi = (S, \pi)$ such that each S_i is finite. Assume that $\xi_k^t < \lambda(I_t)$ for some k and $t \in T$. Then, there exists a positive mass of students $i \in I_t$ such that

$$p(\{\theta \in \Theta \mid \varphi_i(\sigma(\pi(\theta))) = c_k\} \mid s_i) < 1,$$

for some signal $s_i \in S_i$. In other words, there exists a positive mass of students i in the group I_t such that i does not match with the school c_k with some positive probability conditional on some signal.

Proof. For each student $i \in I_t$, let $\Theta(i)$ be the set of states in which i matches with c_k. Formally,

$$\Theta(i) = \{\theta \in \Theta \mid \varphi_i(\sigma(\pi(\theta))) = c_k\},$$

for each student $i \in I_t$. Then, by definition, we have

$$\begin{aligned}
\xi_k^t = \int_{\Theta} \xi_k^t dp &= \int_{\theta\in\Theta} \left(\int_{i\in I_t} \mathbb{I}(\{\theta \in \Theta(i)\}) d\lambda(i) \right) dp(\theta) \\
&= \int_{i\in I_t} \left(\int_{\theta\in\Theta} \mathbb{I}(\{\theta \in \Theta(i)\}) dp(\theta) \right) d\lambda(i) \\
&= \int_{i\in I_t} p(\Theta(i)) d\lambda(i),
\end{aligned}$$

where $\mathbb{I}$ is the usual indicator function. Therefore, there exists a positive mass of students $i \in I_t$ such that $p(\Theta(i)) < 1$, because otherwise $\xi_k^t = \lambda(I_t)$. Now, recall that the set of possible signal realizations S_i for each such student i is finite. Hence, there must exist a signal realization $s_i \in S_i$ under which $p(\Theta(i) \mid s_i) < 1$ holds. This is what we wanted. □

In the following, we prove sequences of lemmas that explore the crucial properties of implementable distributions. Before moving on to the lemmas, here let us explore some important preliminary observations. That is, for any information structure $\Pi = (S, \pi)$ implementing a state-independent distribution, it is without loss of generality that we assume throughout the proof of Theorem 1 each S_i being finite and each signal $s_i \in S_i$ being realized with a strictly positive probability.

First, if a state-independent distribution $\xi \in \mathbb{R}^{|C|\times|T|}$ is implementable, then the *revelation principle* argument of Bergemann and Morris (2016) and Taneva (2019) implies the existence of an information structure $\Pi = (S, \pi)$ with each $S_i = \mathcal{R}_i$ being finite that implements ξ, combined with the *obedient strategy profile* $\sigma(s) = s$. Therefore, without loss of generality, we assume throughout this proof that each S_i is finite. Note that the product $S = (S_i)_{i \in I}$ is still uncountably infinite.

Next, in this information structure, suppose $p(s_i) = p \circ \pi^{-1}(s_i) = 0$ for some $s_i \in S_i$. Then, replacing the signal realization of s_i with another signal with positive support, the resulting information structure still implements ξ with the obedient strategy profile, because the beliefs of the students at each signal do not change. To sum up, it is without loss of generality that we further assume $p(s_i) > 0$ for each $s_i \in S_i$.

The next lemma claims that all implementable state-independent distributions under the decentralized market are *non-wasteful*. Recall that a state-independent distribution $\xi \in \mathbb{R}^{|C|\times|T|}$ is *non-wasteful*, if $\sum_t \xi_k^t > 0$ implies that $\sum_t \xi_{k-1}^t = 1$ for all $k > 1$.

Lemma 4. Suppose that a state-independent distribution $\xi \in \mathbb{R}^{|C|\times|T|}$ is implementable. Then, the state-independent distribution ξ is non-wasteful.

Proof. Let $\Pi = (S, \pi)$ and σ implement ξ. Take any $k > 1$ with $\sum_t \xi_k^t > 0$. Then, there must exist a student $i \in I_t$ for some $t \in T$ and a signal $s_i \in S_i$ such that $\sigma_i(s_i) = c_k$. The interim payoff to the student i at the signal s_i is bounded from above by $u_t(c_k)$. Now, consider the deviation of the student i to the interim strategy c_{k-1}. If we have $\sum_t \xi_{k-1}^t < 1$, then the definition of state-independent distributions imply that, under this deviation, the student i matches with c_{k-1} almost surely, which leads to the interim expected payoff of $u_t(c_{k-1}) > u_t(c_k)$, a contradiction. Therefore, the state-independent distribution ξ must be non-wasteful. □

Next, we show that each implementable state-independent distribution exhibits a monotone structure. That is, for any $t < h$, no student $i \in I_h$ has any chance of matching with a school that is more preferred to a school with which some student $j \in I_t$ match.

Lemma 5. Suppose that a state-independent distribution $\xi \in \mathbb{R}^{|C|\times|T|}$ is implementable. Then, there exists a labeling function $k : T \to \{1, \ldots, |C|\}$, such that, for any $t \in T$, it holds that

$$\{c_k \in C \mid \xi_k^t > 0\} \subset \{c_k \in C \mid k(t-1) \leq k \leq k(t)\},$$

and $k(t) \geq k(t-1)$, where $k(0) = 1$.

Proof. For each $t \in T$, define $k(t)$ to be the maximal number k such that $\xi_k^t > 0$. If there is no such number k, let $k(t) = |C|$. Setting $k(0) = 1$, we prove in the following that the labeling function $k : T \to \{1, \ldots, |C|\}$ constructed this way satisfies the desired properties given in the statement. Let $\Pi = (S, \pi)$ and σ implement ξ.

Take any $t \in T$ and $c_k \in C$ with $\xi_k^t > 0$. Notice that we must have $k \leq k(t)$ by the above construction. To complete the proof, suppose on the contrary that $k < k(t-1)$. Again by the construction, we have $\xi_{k(t-1)}^{t-1} > 0$. Therefore, there exists a student $i \in I_{t-1}$ and a signal $s_i \in S_i$ such that $\sigma_i(s_i) = c_{k(t-1)}$ holds . It is clear that the interim payoff to i conditional on s_i is bounded from above by $u(c_{k(t-1)})$.

Meanwhile, consider a deviation of the student i to the strategy $R_i = c_k$. Then, the hypothesis $\xi_k^t > 0$ and the definition of state-independent distributions imply that i could match with c_k almost surely, resulting in the interim payoff of $u_t(c_k)$, which is strictly greater than $u_t(c_{k(t-1)})$ from $k < k(t-1)$. This is a contradiction. Therefore, the condition $k(t-1) \leq k \leq k(t)$ must hold, and thus we get the inclusions given in the statement.

Finally, we prove $k(t) \geq k(t-1)$ for any $t \in T$. Pick any $t > 1$. If $\xi_k^t = 0$ for all schools $c_k \in C$, then we have $k(t) = |C| \geq k(t-1)$ by the construction. Otherwise, $\xi_k^t > 0$ for some k. In this case, the inclusion in the statement shows $k(t-1) \leq k \leq k(t)$. This is the end of the proof. □

The labeling function k plays an important role in the subsequent results. For each implementable state-independent distribution $\xi \in \mathbb{R}^{|C|\times|T|}$, we denote by $k(\xi)$ be the *labeling function* constructed as in the proof of Lemma 5.

The next key lemma follows from the aggregation of the students' incentive conditions. Especially, they are necessary conditions for the students not replacing their choices in their equilibrium strategies with a school that admits some students with lower tiers. Recall that we write $c_{|C|+1} = \emptyset$ to be the outside option for convenience.

Lemma 6. Suppose that a state-independent distribution $\xi \in \mathbb{R}^{|C|\times|T|}$ is implementable. Take the labeling function $k(\xi)$ of the state-independent distribution. Define $k^*(\xi) : T \to \{1, \ldots, |C| + 1\}$ to be a function such that $k^*(\xi)(t) = k(\xi)(t) + 1$ if $\xi^t_{k(t)} = 1$ and $k^*(\xi)(t) = k(\xi)(t)$ otherwise. Then, we have

$$\sum_{k=k(\xi)(t-1)}^{k(\xi)(t)} \xi^t_k \cdot u_t(c_k) \geq \lambda(I_t) \cdot u_t(c_{k^*(\xi)(t)}),$$

for any $t \in T$.

Proof. Let $\Pi = (S, \pi)$ and σ implement ξ. For each student $i \in I$, let $\mathbb{I}_i : C \times \Theta \to \{0, 1\}$ be an indicator function that satisfies $\mathbb{I}_i(c, \theta) = 1$ if and only if $(\varphi^D \circ \sigma \circ \pi)(\theta)(i) = c$, for each $c \in C$ and $\theta \in \Theta$. Also, for each student $i \in I$ and school $c \in C$, denote by $S_i(c) \subset S_i$ the set of signals $s_i \in S_i$ such that $\sigma_i(s_i) = c$, that is, the set of signals in which the application equals the school c. Denote by $k = k(\xi)$ and $k^* = k^*(\xi)$ for convenience.

Now, take any $t \in T$. Then, for each student $i \in I_t$, consider an interim strategy $R_i = c_{k^*(t)}$ of the student, which may or may not be different from the equilibrium strategy $\sigma_i(s_i)$ for some signals. From Lemma 5, almost no student $i \in I_h$ with $h < t$ applies to the school $c_{k^*(t)}$. Moreover, with probability one, the mass of applicants in I_t to $c_{k^*(t)}$ is strictly less than one, which follows from the definition of the function k^*. The deviation R_i therefore yields the interim expected utility of $u_t(c_{k^*(t)})$ to the student. Since the interim incentive compatibility condition for each student $i \in I_t$ shows that this deviation is not profitable, we must have the inequality

$$\frac{1}{p_i(s_i)} \int_{\theta \in \pi^{-1}(s_i)} \mathbb{I}_i(c, \theta) u_t(c) dp(\theta) \geq u_t(c_{k^*(t)}),$$

for each $s_i \in S_i(c)$. Summing them up over $S_i(c)$ and then over C, we further get

$$\sum_{c \in C} \int_{\theta \in \Theta} \mathbb{I}_i(c, \theta) u_t(c) dp(\theta) = \sum_{c \in C} \sum_{s_i \in S_i(c)} \int_{\theta \in \pi^{-1}(s_i)} \mathbb{I}_i(c, \theta) u_t(c) dp(\theta)$$

$$\geq \sum_{c\in C} \sum_{s_i \in S_i(c)} p_i(s_i) u_t(c_{k^*(t)}) = u_t(c_{k^*(t)}),$$

where the first equivalence follows from the simple observation that $\mathbb{I}_i(c,\theta) = 0$ must hold for all states $\theta \in \Theta$ with $\pi_i(\theta) \notin S_i(c)$. Finally, aggregating the above inequalities for all students in I_t, we get the following inequality:

$$\sum_{k=k(t-1)}^{k(t)} \xi_k^t \cdot u_t(c_k) = \int_{\theta\in\Theta} \left(\int_{i\in I_t} \sum_{k=1}^{|C|} \mathbb{I}_i(c_k,\theta) d\lambda(i) \right) u_t(c_k) dp(\theta)$$

$$= \int_{i\in I_t} \sum_{k=1}^{|C|} \int_{\theta\in\Theta} \mathbb{I}_i(c_k,\theta) u_t(c_k) dp(\theta) d\lambda(i) \geq \lambda(I_t) \cdot u_t(c_{k^*(t)}),$$

where the first equality follows from the definition of state-independent distributions and a statement in Lemma 5: $\xi_k^t > 0$ only if $k(t-1) \leq k \leq k(t)$. Notice that this is exactly the inequality we want, completing the proof. □

The last preliminary lemma below is used for constructing a cutoff vector for a cutoff signal that implements a targeted state-independent distribution. In the main proof of Theorem 1, for each *target vector* $n(\xi) \in \mathbb{R}^{|C|\times|T|}$ in the statement of the next lemma, we construct a cutoff vector so that the mass of applicants to each school equals the target almost surely, under the associated cutoff signal along with the straightforward strategy profile.

It is also worth noting that the next proposition no longer focuses on implementable state-independent distributions. In fact, we show in Lemma 8 that the set of statements in Lemma 5 and Lemma 6 is a sufficient condition for implementation of state-independent distributions with cutoff signals. For each vector $n \in \mathbb{R}^{|C|\times|T|}$, we write $n \wedge \mathbf{1} = (\min\{n_k^t, [1 - \sum_{h<t} n_k^h]_+\})_{k,t} \in \mathbb{R}^{|C|\times|T|}$, where $[a]_+ = \max\{a, 0\}$ for each $a \in \mathbb{R}$. In other words, if the distribution of the applicants equals n almost surely, then the resulting state-independent distribution coincides with $n \wedge \mathbf{1}$.

Lemma 7. Let $\xi \in \mathbb{R}^{|C|\times|T|}$ be a non-wasteful state-independent distribution that has functions $k = k(\xi)$ and $k^* = k^*(\xi)$ each of which satisfies the conditions given in Lemma 5 and Lemma 6, respectively. Then, there exists a *target vector* $n(\xi) \in \mathbb{R}^{|C|\times|T|}$ with $\xi = n(\xi) \wedge \mathbf{1}$ that has each of the following three properties:

1. For each $t \in T$ with $\sum_k \xi_k^t > 0$, we have $\sum_{k=1}^{|C|} n_k^t(\xi) = \lambda(I_t)$.

2. For each $t \in T$, the value $\xi_k^t \cdot u_t(c_k)/n_k^t(\xi)$ is weakly decreasing in $k^*(t-1) \leq k < k^*(t)$.

3. For each $t \in T$, we have $\xi_k^t \cdot u_t(c_k)/n_k^t(\xi) \geq u_t(c_{k^*(t)})$ for $k^*(t-1) \leq k < k^*(t)$.

Proof. The proof is basically by construction. First of all, define a function $n : \mathbb{R}_{++}^{|C|\times|T|} \to \mathbb{R}^{|C|\times|T|}$ as follows: Take any $\delta \in \mathbb{R}_{++}^{|C|\times|T|}$. Pick any coordinate k and $t \in T$. If $k < k(t-1)$ or $k > k(t)$, define $n_k^t(\delta) = 0$. Notice that we must have $k(t-1) \leq k \leq k(t)$ otherwise. Then, for each k satisfying $k(t-1) \leq k \leq k(t)$ and each $t \in T$, define $n_k^t(\delta) = \xi_k^t$ if we have $k = k(t) = k^*(t)$, and define

$$n_k^t(\delta) = \xi_k^t \frac{u_t(c_k)}{\delta_k^t u_t(c_{k^*(t)})}$$

for other cases. Observe that the function n is continuous in $\delta > \mathbf{0}$. Let $A \subset \mathbb{R}_{++}^{|C|\times|T|}$ be the set of inputs $\delta \in \mathbb{R}_{++}^{|C|\times|T|}$ such that $n(\delta) \in \mathbb{R}^{|C|\times|T|}$ satisfies $\xi = n(\delta) \wedge \mathbf{1}$ plus all the three requirements given in the statement except for the first condition.

First, note that $n(\delta) \in \mathbb{R}^{|C|\times|T|}$ satisfies the last two conditions in the statement if it holds that $\delta = \mathbf{1} \in \mathbb{R}_{++}^{|C|\times|T|}$, because the direct calculation implies equality

$$\xi_k^t \cdot u_t(c_k)/n_k^t(\mathbf{1}) = u_t(c_{k^*(t)}),$$

for each t and k with $k^*(t-1) \leq k < k^*(t)$. Since it is trivial from the construction of the function n that $\xi = n(\mathbf{1}) \wedge \mathbf{1}$ is also met, we have $\mathbf{1} \in A$, and hence $A \neq \emptyset$. Second, Lemma 5 implies that, for $\hat{\delta} \in \mathbb{R}_{++}^{|C|\times|T|}$ defined as $\hat{\delta}_k^t = u_t(c_k)/u_t(c_{k^*(t)}) > 0$ for each t and k, the last two conditions and $\xi = n(\hat{\delta}) \wedge \mathbf{1}$ are again satisfied, because for each t and k with $k^*(t-1) \leq k < k^*(t)$, we have

$$\xi_k^t \cdot u_t(c_k)/n_k^t(\hat{\delta}) = u_t(c_k),$$

which is strictly decreasing in k. Thus, $\hat{\delta} \in A$ holds. Third, it is not difficult to see by analogous calculations that the set A is convex.

Now, for each type $t \in T$ such that $\sum_k \xi_k^t > 0$ holds, consider a continuous function $f_t : A \to \mathbb{R}$,

$$f_t(\delta) = \sum_{k=1}^{|C|} n_k^t(\delta) - \lambda(I_t),$$

for each $\delta \in A$, where the continuity of each f_t follows from the function $n(\delta)$ being continuous in $\delta \in A$. From Lemma 6, we have $f_t(\mathbf{1}) \geq 0$, because

$$f_t(\mathbf{1}) \geq 0 \Longleftrightarrow \sum_{k=1}^{|C|} \xi_k^t u_t(c_k) \geq \lambda(I_t) u_t(c_{k^*(t)}) \Longleftrightarrow \sum_{k=k(t-1)}^{k(t)} \xi_k^t u_t(c_k) \geq \lambda(I_t) u_t(c_{k^*(t)}),$$

where the last equivalence follows from Lemma 5. Meanwhile, from $n_k^t(\hat{\delta}) = \xi_k^t \leq 1$ for all k, we get $f_t(\hat{\delta}) \leq 0$. Therefore, the standard Intermediate Value Theorem shows that there exists $\delta^t \in A$, which takes a form $\delta^t = a^t \mathbf{1} + (1 - a^t)\hat{\delta}$ for some $0 \leq a^t \leq 1$, such that $f_t(\delta^t) = 0$.

Finally, define the target vector $n(\xi) \in \mathbb{R}^{|C| \times |T|}$ such that for each k, $n_k^t(\xi) = n_k^t(\delta^t)$ for each $t \in T$ with $\sum_k \xi_k^t > 0$, and $n_k^t(\xi) = 0$ for other $t \in T$. Since $\delta^t \in A$ for each $t \in T$, the last two conditions and $\xi = n(\xi) \wedge \mathbf{1}$ are satisfied. Moreover, $f_t(\delta^t) = 0$ for each $t \in T$ implies the target vector satisfying the first condition as well. This is the end of the proof. □

At last, we are ready to prove Theorem 1. The following lemma serves as the proof, which is restated in the subsection 3.1. Since this lemma no longer focuses on implementable state-independent distributions and the converse statement is immediate from previous statements, it provides a characterization for state-independent distributions that are implementable with cutoff signals and the straightforward strategies.

Lemma 8. Let $\xi \in \mathbb{R}^{|C| \times |T|}$ be a non-wasteful state-independent distribution that has functions $k(\xi)$ and $k^*(\xi)$ each of which satisfies the conditions in Lemma 5 and Lemma 6, respectively. Then, there exists a cutoff signal that implements ξ along with the straightforward strategy profile.

Proof. Let $\xi \in \mathbb{R}^{|C| \times |T|}$ be a non-wasteful state-independent distribution that has functions $k = k(\xi)$ and $k^* = k^*(\xi)$ that satisfy the conditions in Lemma 5 and Lemma 6. It follows from Lemma 7 that there exists a target vector $n = n(\xi)$ that has the four properties listed in the lemma.

To begin with, we construct a cutoff vector $e \in \mathbb{R}^{|C|}$. For each k, let $t \in T$ be a number that solves $k^*(t-1) \leq k < k^*(t)$, which is unique if exists. If there is no such t, set $e_k = |T|$. Then, define the cutoff level for the school c_k as follows:

$$e_k = \frac{n_k^t}{\lambda(I_t) - \sum_{h<k} n_h^t} + (t-1).$$

Write $e = (e_k)_k$ to be the cutoff vector constructed according to the above definition. Notice that it is not difficult to check by mathematical induction along with Lemma 2 that we have

$$\lambda(\{i \in I_t \mid \theta(i)(c_k) > e_k, \forall k < h\}) = \lambda(I_t) - \sum_{k<h} n_k^t,$$

for each h almost surely. Our goal is to show that the cutoff signal $\Pi = (S, \pi)$ with this cutoff vector e implements the state-independent distribution ξ.

Take $\sigma : S \to \mathcal{R}$ to be the straightforward strategy profile of the students. That is, for each student $i \in I_t$ and each signal $s_i \subset C \cup \{\varnothing\}$, the student applies to the most preferred school among the set of recommended schools, $\sigma_i(s_i) = \arg\max_{c \in s_i} u_t(c)$ for each i and s_i.

First, we need to check that the cutoff signal with the cutoff vector e and the straightforward strategies σ induce the desired outcome ξ almost surely. To see this, take any school $c_h \in C$. It follows from Lemma 2 and the definition of the straightforward strategies σ that for almost all states $\theta \in \Theta$, and for $t \in T$ such that $t - 1 \leq e_h \leq t$, the mass of students in I_t applying to the school c_h is calculated as follows:

$$\begin{aligned}
\lambda(\{i \in I_t \mid \sigma_i(\pi_i(\theta)) = c_h\}) &= \lambda(\{i \in I_t \mid \theta(i)(c_h) \leq e_h \wedge \theta(i)(c_k) > e_k, \forall k < h\}) \\
&= \left(\lambda(I_t) - \sum_{k<h} n_k^t\right) \cdot (e_h - (t-1)) \\
&= \left(\lambda(I_t) - \sum_{k<h} n_k^t\right) \cdot \left(\frac{n_h^t}{\lambda(I_t) - \sum_{k<h} n_k^t}\right) \\
&= n_h^t.
\end{aligned}$$

Therefore, since it follows from Lemma 7 that we must have $\xi = n \wedge \mathbf{1}$, it holds that Π and σ induce the state-independent distribution $\xi \in \mathbb{R}^{|C| \times |T|}$ at almost all states $\theta \in \Theta$.

Second, it remains to show that the straightforward strategy profile σ is indeed a BNE. Take any student $i \in I_t$ and any signal $s_i \in S_i$. Consider an interim deviation $R_i = c_k$ of the student at the signal s_i for some $c_k \in C$.

The deviation is not profitable if $c_k \notin s_i$ holds. In fact, this means $\theta(i)(c_k) > e$ for all θ with $\pi_i(\theta) = s_i$, while we also have $\theta(j)(c_k) \leq e$ for the mass $n_k^h \geq \xi_k^h$ of students in I_h for all $h \in T$, almost surely. Since $c_k \notin s_i$ also implies $k < k^*(t)$ from the monotonicity in Lemma 5, the above relation and non-wastefulness of the state-independent distribution ξ shows that, conditional of the signal realization s_i, the student i matches with c_k with probability zero, which cannot be strictly higher than the interim payoff from the interim strategy $\sigma_i(s_i)$.

Finally, suppose that $c_k \in s_i$. Note that we must have $k^*(t-1) \leq h \leq k^*(t)$ by the construction of the cutoff vector, where $\sigma_i(s_i) = c_h$. Additionally, from the definition of the strategies σ, we must have $h < k$ for the school c_h. Suppose $k \geq k^*(t)$. Here, by the above equality, Lemma 1, and Lemma 2, we may calculate the student's interim expected payoff from σ conditional on the signal s_i as follows:

$$\begin{aligned}
\int_{\theta\in\Theta} \mathbb{I}_i(c_h,\theta)u_t(c_h)dp(\theta \mid s_i) &= \frac{u_t(c_h)}{p(s_i)}\int_{\theta\in\pi^{-1}(s_i)} \mathbb{I}_i(c_h,\theta)dp(\theta) \\
&= u_t(c_h)\cdot p(\{\theta\in\Theta \mid \lambda(\{j\in A(\theta) \mid \theta(j)(c_h) \leq \theta(i)(c_h)\}) < \xi_h^t\}) \\
&= u_t(c_h)\cdot p(\{\theta\in\Theta \mid n_h^t\cdot(\theta(i)(c_h)-(t-1)) \leq \xi_h^t\}) \\
&= u_t(c_h)\cdot p(\{\theta\in\Theta \mid \theta(i)(c_h) \leq \xi_h^t/n_h^t + (t-1)\}) \\
&= \frac{\xi_h^t u_t(c_h)}{n_h^t},
\end{aligned}$$

where $A(\theta)$ is the set of those who apply to c_h at the state θ in I_t. From Lemma 7, the last term is no less than $u(c_{k^*(t)})$, which is the largest possible interim payoff from the deviation to $R_i = c_k$. Therefore, the deviation is not profitable.

By contrast, suppose that $k < k^*(t)$ holds. Then, we must have $h < k$ from the definition of the straightforward strategies, and thus the second property listed in Lemma 7 implies the inequality

$$\frac{\xi_h^t u(c_h)}{n_h^t} \geq \frac{\xi_k^t u(c_k)}{n_k^t},$$

where the last term is, by an analogous calculation, exactly the interim payoff of the student i conditional on s_i from the interim deviation $R_i = c_k$. Therefore, the deviation is again not profitable. In conclusion, the straightforward strategy $\sigma_i(s_i)$ maximizes the interim payoff to the student i conditional on the signal s_i, hence an equilibrium. □

Proof of Theorem 1. It follows immediately from Lemma 4, Lemma 5, Lemma 6, Lemma 7, and Lemma 8. □

B.2 Proof of Proposition 2 and Proposition 3

The proofs of Proposition 2 and Proposition 3 are basically by the enumerations of the mass of students having ex-post justified envies.

Proof of Proposition 2. First, we can count the expected mass of unmatched students for each type,

$$\lambda(I_t) - \sum_k \xi_k^t \geq 0,$$

for each tier $t \in T$. Second, take any state $\theta \in \Theta$ and matched student $i \in I_t$ with $c_i = \mu^*(\theta)(i) \neq \varnothing$. For each school $c_l \in C$, if $u_t(c_l) > u_t(c_i)$ holds, then the definition of the straightforward strategy implies that $c_l \notin \pi(\theta)(i)$, and thus $e_l < \theta(i)(c_l)$. Therefore, for any student $j \in I$ with $\mu(\theta)(j) = c_l$, we must have $\theta(j)(c_l) \leq e_l < \theta(i)(c_l)$, which means i has no ex-post justified envy at θ, under the matching outcome μ^*.

For any matching outcome inducing a state-independent distribution $\xi \in \mathbb{R}^{|C|\times|T|}$, each unmatched student $i \in I_t$ has an ex-post justified envy at a state θ, if there exists a student $j \in I_h$ with $h > t$ who matches with a school at the state θ. In conclusion, we have

$$E(\mu^*) = \sum_{t<t^*(\xi)} \left(\lambda(I_t) - \sum_k \xi_k^t \right) \leq E(\mu),$$

for any μ that induces ξ almost surely, where $t^*(\xi)$ is the maximal number t such that $\sum_k \xi_k^t > 0$. Hence, μ^* is no less ex-post fair than μ.

To see that μ^* is ex-ante fair, take any $t \in T$. Then, for any student $i \in I_t$, Lemma 8 and the definitions of cutoff signals and the straightforward strategies result in the existence of two monotone labeling functions $k = k(\xi)$ and $k^* = k^*(\xi)$ such that

$$\int_{\theta\in\Theta} u_t(\mu^*(\theta)(i))dp(\theta) = \sum_{k=k(t-1)}^{k(t)} \frac{\xi_k^t}{\lambda(I_t)} \cdot u_t(c_k) \geq u_t(c_{k^*(t)}) \geq \int_{\Theta} u_t(\mu^*(\theta)(j))dp,$$

for any $j \in I_h$ with $h > t$. Hence, μ^* is ex-ante fair, completing the proof. □

Proof of Proposition 3. Let $\xi = \xi \circ \mu$ and $\tilde{\xi} = \xi \circ \mu'$ almost surely, where $\xi \in \mathbb{R}^{|C|\times|T|}$ and $\tilde{\xi} \in \mathbb{R}^{|C|\times|T|}$.

As seen in the proof of Proposition 2, we have

$$E(\mu) = \sum_{t<t^*(\xi)} \lambda(I_t) - \sum_{t<t^*(\xi)} \sum_k \xi_k^t,$$

and the corresponding equation for μ' by replacing ξ with $\tilde{\xi}$.

Note first that we must have $t^*(\xi) \leq t^*(\tilde{\xi})$. To see this, take the least preferred school c_k such that $\xi_k^t > 0$, where $t = t^*(\xi)$. Then, we have $e_k > t - 1$ by the definition of straightforward strategies. It also implies, under e, that the mass of students $i \in I_h$ with $h < t$ applying to c_k is strictly less than the unit capacity almost surely. Since all students with tier $h < k$ applying to c_k under e' must apply to c_k under $e \geq e'$ at the straightforward strategies, we also know that, at e', the mass of students in I_h with $h < t$ admitted to c_k is strictly less than the unit capacity almost surely. From Lemma 4, therefore, we must have $t^*(\tilde{\xi}) \geq t = t^*(\xi)$.

Next, we prove $\sum_{t<t^*} \sum_k \xi_k^t \leq \sum_{t<t^*} \sum_k \tilde{\xi}_k^t$, where $t^* = t^*(\tilde{\xi})$. Take labeling functions $k = k(\xi)$ and $\tilde{k} = k(\tilde{\xi})$ for the distributions ξ and $\tilde{\xi}$ respectively, according to Lemma 8. Set $k = k(t^*)$ and $\tilde{k} = \tilde{k}(t^*)$. If $k < \tilde{k}$, then from Lemma 8 the non-wastefulness condition immediately shows the desired inequality. Thus, assume $k \geq \tilde{k}$. Since $e \geq e'$ holds by assumption, we have $k = \tilde{k}$.

Now, the inequalities $e_k \geq e'_k > t^* - 1$ show that all students in I_h with $h < t^*$ receive a signal including c_k at all states. Since $e \geq e'$, the mass of students in I_h with $h < t^*$ applying to c_k is always larger in e' than in e, from the definition of the straightforward strategies. Therefore, we have $\sum_{t<t^*} \xi_k^t \leq \sum_{t<t^*} \tilde{\xi}_k^t$. Since Lemma 8 implies that this inequality holds with equality for all other schools, we finally get $\sum_{t<t^*} \sum_k \xi_k^t \leq \sum_{t<t^*} \sum_k \tilde{\xi}_k^t$.

We have shown that the two inequalities $t^*(\xi) \leq t^*(\tilde{\xi})$ and $\sum_{t<t^*(\tilde{\xi})} \sum_k \xi_k^t \leq \sum_{t<t^*(\tilde{\xi})} \sum_k \tilde{\xi}_k^t$ hold. Therefore, we finally have the following inequalities:

$$\begin{aligned}
\sum_{t<t^*(\xi)} \lambda(I_t) - \sum_{t<t^*(\xi)} \sum_k \xi_k^t &= \sum_{t<t^*(\tilde{\xi})} \lambda(I_t) - \sum_{t<t^*(\tilde{\xi})} \sum_k \xi_k^t + \sum_{t=t^*(\tilde{\xi})}^{t^*(\xi)-1} \left(\lambda(I_t) - \sum_k \xi_k^t \right) \\
&\geq \sum_{t<t^*(\tilde{\xi})} \lambda(I_t) - \sum_{t<t^*(\tilde{\xi})} \sum_k \xi_k^t \\
&\geq \sum_{t<t^*(\tilde{\xi})} \lambda(I_t) - \sum_{t<t^*(\tilde{\xi})} \sum_k \tilde{\xi}_k^t,
\end{aligned}$$

which implies $E(\mu) \geq E(\mu')$. This is exactly what we wanted. $\square$

B.3 Proof of Proposition 4

The proof of Proposition 4 constructs variants of cutoff signals that uniquely implement almost all matching outcomes that are implementable with cutoff signals and the straightforward strategies. While the signal realizations of constructed signals are also based upon cutoff vectors, there are two crucial differences. First, they are *public* signals, that is, the set of recommended schools for the students are publicly available. Second, cutoff vectors in the subsequent proof will become state-dependent, although it is not allowed in the definition of the original cutoff vectors. Whether the original cutoff signals could uniquely implement almost all distributions is a non-trivial question, although answering this is beyond the scope of this paper.

Before proceeding to the proof, let us introduce some intuitive notations and their fundamental properties, which will be useful for the subsequent proofs in addition to that of Proposition 4. Take a cutoff vector $e \in \mathbb{R}^{|C|}$ that along with the straightforward strategies implements an distribution. Then, from the exact law of large numbers of Lemma 2, we can derive the exact mass of students of each type applying to each school. More precisely, we can define the *associated target*, denoted by $n(e) \in \mathbb{R}^{|C| \times |T|}$, where

$$n_k^t(e) = \lambda(\{i \in I_t \mid \theta(i)(c_k) \leq e_k \wedge \theta(i)(c_l) > e_l, \forall l < k\}),$$

for each k and t. Note that the right hand side does not depend on the selection of states θ almost surely due to Lemma 2. Notice also that $n_k^t(e)$ is independent of e_l with $l > k$ and is continuous in e, which can be observed from the direct calculation.

If the straightforward strategy profile is a BNE at a cutoff signal with the vector $e \in \mathbb{R}^{|C|}$, then we may find the resulting equilibrium state-independent distribution. This is denoted by $\xi(e) \in \mathbb{R}^{|C| \times |T|}$, and is given explicitly by $\xi(e) = n(e) \wedge \mathbf{1}$. If every student plays the straightforward strategy, each student in I_t with an application to c_k earns the interim expected payoff of $\xi_k^t(e) u_t(c_k) / n_k^t(e)$.

With these notations at hand, we have the following preliminary lemma. This statement is analogous to Proposition 1 (Lemma 8) and characterizes when a cutoff signal implements an distribution with the straightforward strategies.

Lemma 9. The straightforward strategies constitute a BNE under a cutoff signal with a cutoff vector $e \in \mathbb{R}^{|C|}$, if and only if the following three conditions are satisfied: (i) $\xi(e)$ is non-wasteful; (ii) there exists a monotone labeling $k(e)$ satisfying

$$\{c_k \in C \mid \xi_k^t(e) > 0\} \subset \{c_k \in C \mid k(e)(t-1) \leq k \leq k(e)(t)\},$$

and; (iii) for all k with $k(e)(t-1) \leq k < k(e)(t)$, we have

$$\frac{\xi_k^t(e)u_t(c_k)}{n_k^t(e)} \geq \frac{\xi_{k+1}^t(e)u_t(c_{k+1})}{n_{k+1}^t(e)} \geq u_t(c_{k^*(e)(t)}),$$

where $k^*(e)(t) = k(e)(t) + 1$ if $\xi_{k(e)(t)}^t = 1$ and $k^*(e)(t) = k(e)(t)$ otherwise.

Proof. The necessity of the conditions is trivial from Proposition 1. For sufficiency, take any student and a signal realization under which the student applies to a school c_k. For any $l < k$, the interim payoff from the application to c_l equals zero from the definitions of the straightforward strategies and the non-wastefulness of the distribution. Meanwhile, for any $l > k$, the application to c_l is not profitable, which follows immediately from the last inequality conditions. □

Let F be the set of all cutoff vectors such that the straightforward strategies constitute an equilibrium under the associated cutoff signals. Then, define $F^\circ \subset F$ to be the subset of the cutoff vectors in F each of which satisfies the last inequalities in Lemma 9 with strict inequalities for all $t \in T$ and k with $k(e)(t-1) \leq k < k(e)(t)$. With these preliminaries, we have the first crucial lemma for the proof of Proposition 4.

Lemma 10. Suppose that a matching outcome μ is implementable with a cutoff signal with a cutoff vector $e \in F^\circ$ and the straightforward strategies. Then, the matching outcome μ is uniquely implementable.

Proof. Consider a public signal $\Pi = (S, \pi)$ that for each $\theta \in \Theta$ publicly announces the profile of recommendations $c = (c_i)_{i \in I}$ according to the following procedure:

- Step 1. Take a set N_1 of the mass $\sum_{t\in T} n_1^t(e)$ of students who have the highest priorities at the school c_1. Set $I^2 = I \setminus N_1$.

- Step $k, k \geq 2$. Given that the set I^k is defined, take a set $N_k \subset I^k$ of the mass $\sum_{t\in T} n_k^t(e)$ of students who have the highest priorities at the school c_k among the remaining students I^k. Set $I^{k+1} = I^k \setminus N_k$.

Finally, let $\pi(\theta)(i) = c = (c_j)_{j\in I}$ for each $i \in I$, where $c_i = c_k$ if and only if $i \in N_k$. For the remaining students $i \in I^{|C|+1}$, let $c_i = \emptyset$. Note that this information structure is well-defined, because by the construction the set N_k is disjoint from the set N_h for each $k \neq h$. We show that all BNE under this information structure induce matching outcomes that are equivalent to the matching outcome μ almost surely.

Pick a public signal realization $c = (c_i)_{i\in I}$. Let N_k be the set of students $i \in I$ such that $c_i = c_k$. Take also any BNE σ, and let M_k be the set of students $i \in I$ such that $\sigma_i(c) = c_k$, that is, the set of students applying to the school c_k at a public signal realization $c = (c_j)_{j\in I}$.

We prove here that $M_k = N_k$ must hold for all k. Seeking a contradiction, assume the existence of k such that $M_k \neq N_k$, and take such minimal k, which is denoted by $\overline{k}$. Then, we have the unique solution t to the inequalities $k^*(e)(t-1) \leq \overline{k} < k^*(e)(t)$.

First of all, no student $i \in I_t \setminus N_{k^*(e)(t)}$ applies to c_h for some $h \geq k^*(e)(t)$. Suppose on the contrary that $\sigma_i(c) = c_h$ for some $i \in N_k \cap I_t$ with $k^*(e)(t-1) \leq k < k^*(e)(t)$. Since the strategy profile σ is a BNE, we must have

$$\lambda\left(M_k \setminus \bigcup_{l>k} N_l\right) > \lambda(N_k) = \sum_{t\in T} n_k^t(e) \geq n_k^t(e),$$

because the interim expected payoff from playing $R_i = c_k$ must be no greater than that of playing $\sigma_i(c_h)$, which is bounded from above by $u_t(c_{k^*(e)(t)})$: if the above inequality does not hold, then the definition of $e \in F^\circ$ leads to a contradiction.

Meanwhile, this implies the existence of a student $j \in N_l \cap I_t$ with $l < k$ such that $j \in M_k$. His or her interim expected payoff from playing $\sigma_j(c) = c_k$ is, by the above inequality, bounded from above by $\xi_k^t(e)u_t(c_k)/n_k^t(e)$. Therefore, again, for the equilibrium condition at the signal realization $c = (c_j)_{j \in I}$ to hold, we must have

$$\lambda\left(M_l \setminus \bigcup_{p>l} N_p\right) > \lambda(N_l) = \sum_{t \in T} n_l^t(e) \geq n_l^t(e),$$

which follows from the definition of $e \in F^\circ$, because otherwise the strategy $R_j = c_l$ leads to a strictly higher interim expected utility. Repeating the same argument, we eventually have

$$\lambda(N_{\overline{k}}) \geq \lambda\left(M_{\overline{k}} \setminus \bigcup_{k>\overline{k}} N_k\right) > \lambda(N_{\overline{k}}),$$

where the first inequality follows from the definition of $\overline{k}$, that is, $N_k = M_k$ for all $k < \overline{k}$. This is a contradiction, however.

Second, we prove by mathematical induction that the relations $\lambda(M_k) \geq 1$ and $M_k \subset \bigcup_{l \leq k} N_l$ hold, for all k with $k^*(e)(t-1) \leq k < k^*(e)(t)$. Note that $M_l = N_l$ for each $l < k^*(e)(t-1) \leq \overline{k}$. Assume that both $\lambda(M_h) \geq 1$ and $M_h \subset \bigcup_{l \leq h} N_l$ hold for each $h < k$. Notice that the non-wastefulness condition of the distribution $\xi(e)$ from Lemma 9 shows that $\lambda(N_k) \geq 1$ for $k < k^*(e)(t)$. Then, we must have $\lambda(M_k \cap N_k) \geq 1$, because otherwise there exists a student $i \in N_k \setminus M_k$ who has a profitable deviation $R_i = c_k$, which provides i with the highest possible interim payoff of $u_t(c_k)$: the student i can match with a school c_h such that $h < k$ with zero probability, due to the inductive hypothesis. Given this, $i \in M_k$ implies $i \in N_l$ for some $l \leq k$, because otherwise the interim payoff to i is zero, so that i has an obvious profitable deviation. Since the case $k = k^*(e)(t-1)$ could be proven analogously, we complete the proof.

Finally, we prove that $N_{\overline{k}} = M_{\overline{k}}$ holds, which results in a contradiction. From the second argument and the definition of $\overline{k}$, we have $M_{\overline{k}} \subset N_{\overline{k}}$. Let $i \in I_t$ be a student in $N_{\overline{k}}$ such that $i \in M_h$ for some $h \neq \overline{k}$. Note that $h > \overline{k}$ holds again by the definition of $\overline{k}$. Moreover, we have $h < k^*(e)(t)$ from the first argument. Now, with the strategy $R_i = c_{\overline{k}}$, they obtains the interim payoff of

$$\frac{\xi_{\overline{k}}^t(e) u_t(c_{\overline{k}})}{\lambda(M_{\overline{k}} \cap I_t)} \geq \frac{\xi_{\overline{k}}^t(e) u_t(c_{\overline{k}})}{\lambda(N_{\overline{k}} \cap I_t)} = \frac{\xi_{\overline{k}}^t u_t(c_{\overline{k}})}{n_{\overline{k}}^t(e)},$$

and thus the equilibrium condition and Lemma 9 imply $\lambda(M_h \cap I_t) < \lambda(N_h \cap I_t) = n_h^t(e)$, which then implies the existence of a student $j \in I_t$ with $j \in N_h$ such that $j \in M_l$ for some $l \neq h$. Note that we must have $l > h$ from the second argument. With the similar rationale, therefore, the definition of $e \in F^\circ$ shows $\lambda(M_l \cap I_t) < \lambda(N_l \cap I_t)$. Repeating this procedure, we eventually have $\lambda(M_{k^*(e)(t)-1} \cap I_t) < \lambda(N_{k^*(e)(t)-1} \cap I_t)$. However, this implies that, from the second argument, some student $i \in I_t$ with $i \in N_{k^*(e)(t)-1}$ applies to a school c_h such that $h \geq k^*(e)(t)$, which is negated by the first claim, hence a contradiction.

Summarizing, if a strategy profile σ is a BNE, it must be the obedient strategy profile. It remains to check that the obedient strategy profile σ indeed constitutes a BNE. Fix a signal of recommendations $c = (c_i)_{i \in I}$ and a student $i \in I_t$. Suppose $c_i = c_k$. Then, the interim expected payoff to the student is exactly calculated as $\xi^t_k(e)u_t(c_k)/n^t_k(e) > 0$ under the strategies σ.

First, any deviation $R_i = c_h$ with $h < k$ yields the interim expected payoff of zero, because i realizes for sure that the applicants to c_h have higher priorities than i has at c_h, conditional on the public signal realization. Second, any deviation $R_i = c_h$ with $k < h < k^*(e)(t)$ yields the interim expected payoff of at most $\xi^t_h(e)u_t(c_h)/n^t_h(e)$ under the regular conditional probability, because conditional on the signal, the priority $\theta(j)(c_h)$ of any student j with $c_i = c_h$ is uniformly distributed over $[t-1, e_h]$ with $e_h < t$ while that of the student i is uniformly distributed over $[t-1, t]$. This is again not profitable from the definition of $e \in F^\circ$ and Lemma 9. Finally, any deviation $R_i = c_h$ with $h > k^*(t)$ yields the interim expected payoff of $u_t(c_h)$, which is not profitable from the definition of $e \in F^\circ$ and Lemma 9. The obedient strategy profile σ is therefore a BNE.

In conclusion, the obedient strategy profile is an essentially unique BNE. It is obvious to see that the obedient strategy profile induces the matching outcome μ almost surely, and thus μ is uniquely implementable. □

For each $e \in F$, recall that $\mu(e)$ is the matching outcome induced by the cutoff signal with the cutoff vector e and the straightforward strategies. According to Lemma 10, it remains to show that all $\mu(e)$ are uniquely implementable, for all $e \in F$ that are not in F°. for each such $e \in F \setminus F^\circ$, let $h(e)$ be the minimal number h such that

$$\frac{\xi^t_h(e)u_t(c_h)}{n^t_h(e)} = \frac{\xi^t_{h+1}(e)u_t(c_{h+1})}{n^t_{h+1}(e)},$$

where t solves $k(t-1) \leq h \leq k(t)$. For each $e \in F^\circ$, set $h(e) = |C|$. Note that for each $e \in F$, we have $h(e) = |C|$ if and only if $e \in F^\circ$, which follows from Lemma 9. Define the usual sup-norm $|\cdot|$ over the space of cutoff vectors $\mathbb{R}^{|C|}$.

Lemma 11. Take any cutoff vector $e \in F \setminus F^\circ$. Then, for any $\delta > 0$, there exists a cutoff vector $\hat{e} \in F$ with $|e - \hat{e}| < \delta$ such that $h(e) < h(\hat{e})$.

Proof. Recall that the cutoff vector $e \in F \setminus F^\circ$ satisfies the conditions in Lemma 9. In the following, we define a new cutoff vector $\hat{e}$ that has the designated properties. Take sufficiently small $\varepsilon > 0$. To begin with, define $\hat{e}_l = e_l$ for all $l < h(e)$, and let $\hat{e}_{h(e)} = e_{h(e)} - \varepsilon$.

Note here that, from Lemma 2, we may calculate the target vector $n(e)$ for each $e \in F$ explicitly as follows: pick any k and t. If $e_k < t-1$, then it is obvious that $n_k^t(e) = 0$ holds. If $t-1 \le e_k \le t$, then Lemma 2 implies that

$$n_k^t(e) = \left(\lambda(I_t) - \sum_{l<k} n_l^t(e)\right)(e_k - (t-1)),$$

which follows from the definition of the target vectors $n(e)$. For the other case $e_k > t$, it holds that $n_k^t(e) = \lambda(I_t) - \sum_{l<k} n_l^t(e)$.

Therefore, using the above formula, we can inductively define the remaining coordinates of the cutoff vector $\hat{e} \in \mathbb{R}^{|C|}$ so that the associated target vector satisfies the following conditions: for each $k > h(e)$ with $h(e) < k < k^*(e)(t)$, pick the cutoff e_k, given $\hat{e}_l$ for $l < k$, such that

$$\frac{\xi_k^t(\hat{e})u_t(c_k)}{n_k^t(\hat{e})} = \frac{\xi_k^t(e)u_t(c_k)}{n_k^t(e)},$$

where we must have the strict inequality $\hat{e}_k > e_k$ due to the above formula and $n_{h(e)}^t(\hat{e}) < n_{h(e)}^t(e)$, for any $\hat{e} \in \mathbb{R}^{|C|}$ that is consistent with this construction. For $k = k^*(e)(t)$, select the cutoff e_k that solves the equality

$$\frac{\xi_k^{t+1}(\hat{e})u_{t+1}(c_k)}{n_k^{t+1}(\hat{e})} = \frac{\xi_k^{t+1}(e)u_t(c_k)}{n_k^{t+1}(e)},$$

where we must have the inequality $n_k^{t+1}(\hat{e}) \le n_k^{t+1}(e)$ from $\xi_k^{t+1}(\hat{e}) \le \xi_k^{t+1}(e)$, which follows from Lemma 9 and $\xi_k^t(\hat{e}) \ge \xi_k^t(e)$. Moreover, we have $|\hat{e}_k - e_k| < \delta$ for sufficiently small $\varepsilon > 0$, which follows from the continuity of the target $n(e)$. Analogously, provided that we have defined $\hat{e}_l$ for each $l < k$ such that $k^*(e)(t-1) \le k < k^*(e)(t)$ for some $t \in T$, define the target $\hat{e}_k$ so that the interim payoffs to the students in I_t under the straightforward strategies coincide between in the cutoff signals with the cutoff vectors e and $\hat{e}$.

Here, we have defined a cutoff vector $\hat{e} \in \mathbb{R}^{|C|}$. It follows from the construction that $k^*(\hat{e}) = k^*(e)$ holds for small enough $\varepsilon > 0$. Therefore, it is not difficult to see that the cutoff vector $\hat{e}$ satisfies the incentive conditions of Lemma 9, which implies $\hat{e} \in F$. Moreover, the construction of $\hat{e}$ implies

$$\frac{\xi_h^t(\hat{e})u_t(c_h)}{n_h^t(\hat{e})} = \frac{\xi_h^t(e)u_t(c_h)}{n_h^t(\hat{e})} > \frac{\xi_{h+1}^t(e)u_t(c_{h+1})}{n_{h+1}^t(e)},$$

and thus we have $h(e) < h(\hat{e})$ by the definition, completing the proof. □

Proof of Proposition 4. First, Lemma 10 shows that a matching outcome $\mu(e)$ is uniquely implementable if $e \in F^\circ$. Second, suppose $e \in F \setminus F^\circ$. It is not difficult to see that the induced matching outcome $\mu(e)$ has a continuity in $e \in F$ with our distance-like function defined in the subsection 3.4. That is, for any $\varepsilon > 0$, the iterative applications of Lemma 11 with small enough $\delta > 0$ eventually imply the existence of a cutoff vector $\hat{e} \in F^\circ$ such that $\rho(\mu(e), \mu(\hat{e})) < \varepsilon$. Then, Lemma 10 again implies that $\mu(\hat{e})$ is uniquely implementable, and thus so is $\mu(e)$ by the definition of unique implementation. □

B.4 Proof of Proposition 5 and Proposition 6

To begin with, we explore an intuitive property for the outputs of the parameterized algorithms introduced at the beginning of Section 4.

Lemma 12. Take any $t \in T$. Then, for any implementable state-independent distribution $\xi \in \mathcal{E}$,

$$k(\xi(t))(h) \leq k(\xi)(h) \text{ and } k(\xi(t))(g) \geq k(\xi)(g),$$

for all $h < t$ and $g \geq t$.

Proof. We show the case $h < t$, and the analogous argument shows the case $g \geq t$. Suppose on the contrary that $k(\xi(t))(h) > k(\xi)(h)$ for some $h < t$. Take such minimal h. Then, we have

$$k(\xi(t))(h-1) \leq k(\xi)(h-1) \leq k(\xi)(h) < k(\xi(t))(h),$$

where the second inequality follows from Proposition 1.

Now, consider the following state-independent distribution $\overline{\xi}$: for each $g \in T$, define $\overline{\xi}_k^g = \xi_k^g(t)$ if $g < h$, and define $\overline{\xi}_k^g = \xi_k^g$ if $g > h$, for each k. If it holds that $k(\xi(t))(h-1) \leq k \leq k(\xi)(h-1)$, then define $\overline{\xi}_k^h = \xi_k^h(t)$. For the remaining cases, define $\overline{\xi}_k^h = \xi_k^h$. Intuitively, the distribution $\overline{\xi}$ coincides with $\xi(t)$ for schools c_k with $k \leq k(\xi)(h-1)$ and with ξ for the remaining schools. Notice that by the construction, we have $k(\overline{\xi})(g) = k(\xi(t))(g)$ if $g < h$ and $k(\overline{\xi})(g) = k(\xi)(g)$ otherwise.

It suffices to show that $\overline{\xi}$ is implementable, which leads to a contradiction with the definition of $\xi(t)$, because $\overline{\xi} \in \mathcal{E}(h-1)$ and $k(\overline{\xi})(h) = k(\xi)(h) < k(\xi(t))(h)$. First of all, the distribution $\overline{\xi}$ is non-wasteful by the construction. Second, the labeling function is increasing in T and satisfies the inclusions in Proposition 1. Finally, we prove that the last condition of Proposition 1 is met.

Take any $g \in T$. The distribution $\overline{\xi}$ trivially satisfies the inequality if $g \neq h$. Thus, suppose $g = h$. Note that we must have $\overline{\xi}_k^h = \xi_k^h(t) \geq \xi_k^h$ if $k = k(\xi(t))(h-1) = k(\xi)(h-1)$, due to the definition of the algorithm. Therefore, it follows from the non-wastefulness that $\overline{\xi}_k^h \geq \xi_k^h$ for all k. Since the distribution ξ is implementable, it satisfies the last inequality in Proposition 1, and thus the distribution $\overline{\xi}$ satisfies the inequality as well. In conclusion, Proposition 1 shows that $\overline{\xi}$ is implementable. □

Lemma 13. For each distribution $\xi(t)$ defined with the parameterized algorithm, let $e(t) \in F$ be a cutoff vector such that $\xi(e(t)) = \xi(t)$. Then, $\mu(e(t))$ is no less ex-post fair than $\mu(e(t+1))$.

Proof. By the construction of the parameterized algorithms in Section 4, we have $E(\mu(e(t+1)) \geq E(\mu(e(t))$ if and only if we have

$$\sum_{h=t}^{t^*(e(t+1))-1} \left(\lambda(I_h) - \sum_k \xi_k^h(t+1)\right) \geq \sum_{h=t}^{t^*(e(t))-1} \left(\lambda(I_h) - \sum_k \xi_k^h(t)\right).$$

It is not difficult to see from the construction of the distribution $\xi(t)$ that $\lambda(I_h) = \sum_k \xi_k^h(t)$ must hold for all $h \geq t$ with $h < t^*(e(t))$. The analogous relation holds for the distribution $\xi(t+1)$. Therefore, the above inequality reduces to the inequality $\lambda(I_{t+1}) \geq \sum_k \xi_k^{t+1}(t+1)$, which is trivially satisfied. This means that $\mu(e(t))$ is no less ex-post fair than $\mu(e(t+1))$ by definition, completing the proof. □

The next result constructs a cutoff vector that induces the least ex-post fair matching outcome in M^* along with the straightforward strategy profile, and explores its crucial properties.

Lemma 14. There exists a cutoff vector e^* under which the straightforward strategy profile is a BNE that satisfies

$$\frac{\xi_k^t(e^*)u_t(c_k)}{n_k^t(e^*)} = \frac{\xi_{k+1}^t(e^*)u_t(c_{k+1})}{n_{k+1}^t(e^*)} = u_t(c_{k^*(e^*)(t)}),$$

for all t and k with $k(e^*)(t-1) \leq k < k(e^*)(t)$, and $e^*_{k(e^*)(|T|)} = |T|$. Conversely, for any cutoff vector e satisfying this condition, we must have $e_k = e^*_k$ for all k with $k < k(e^*)(|T|)$. The cutoff signal with the cutoff vector e^* and the straightforward strategies induce the least ex-post fair matching outcome in M^*.

Proof. For the first claim, consider an implementable state-independent distribution $\xi(|T|)$ that is the output of the algorithm at the parameter $|T| \in T$ described in Section 4. Theorem 1 shows the existence of a cutoff signal with a cutoff vector e^* and the straightforward strategies that implement the distribution $\xi(|T|)$, i.e., $\xi(e^*) = \xi(|T|)$ holds. Further, we assume that $e^*_{k(e^*)(|T|)} = |T|$ because otherwise we may redefine this coordinate, which does not affect the induced distribution $\xi(|T|)$. We show that this cutoff vector e^* meets the desired property. Recall that the cutoff vector e^* satisfies the three conditions in Lemma 9.

Seeking a contradiction, let k be the minimal number such that

$$\frac{\xi_k^t(e^*)u_t(c_k)}{n_k^t(e^*)} > \frac{\xi_{k+1}^t(e^*)u_t(c_{k+1})}{n_{k+1}^t(e^*)},$$

where t solves $k(e^*)(t-1) \leq k < k(e^*)(t)$. Then, as in the proof of Lemma 11, we can take another cutoff vector $e \in F$ under which the straightforward strategies constitute a BNE, such that $e_l = e^*_l$ for $l < k$ and $e_k = e^*_k + \varepsilon$ for sufficiently small $\varepsilon > 0$, and that $n_l^t(e) = n_l^t(e^*)$ for all $l > k$ with $l < k(e^*)(t)$. Note that we can also assume $k(e)(t) = k(e^*)(t)$ for small enough ε due to the continuity of $\xi(e)$. However, it follows from the construction that $\xi^t_{k(e)(t)}(e) < \xi^t_{k(e^*)(t)}(e^*)$ holds, which contradicts the definition of $\xi(e^*) = \xi(|T|)$. The first claim thus holds true.

To see the second argument, let e be another cutoff vector that satisfies the stated conditions. Then, we must have

$$n_k^t(e) = \frac{\xi_k^t(e)u_t(c_k)}{u_t(c_{k^*(e)(t)})},$$

for each t and k with $k(e)(t-1) \leq k < k(e)(t)$. The definition of $\xi(|T|)$ and $\xi(|T|) = \xi(e^*)$ first imply that $k(e)(1) \geq k(e^*)(1)$ and $k^*(e)(1) \geq k^*(e^*)(1)$. From $\sum_{k \leq k(e)(1)} n^1_k(e) = \sum_{k \leq k(e^*)(1)} n^1_k(e^*) = \lambda(I_1)$ by the definition of target vectors, it is not difficult to see that we have both $k(e)(1) = k(e^*)(1)$ and $k^*(e)(1) = k^*(e^*)(1)$ by the above equality and the non-wastefulness of $\xi(e) = n(e) \wedge \mathbf{1}$. Repeating this, we eventually get $k(e)(t) = k(e^*)(t)$ and $k^*(e)(t) = k^*(e^*)(t)$ for each $t \in T$. Therefore, we must have $n(e) = n(e^*)$ by the above equality, which holds only if $e_k = e^*_k$ for all $k < k(e^*)(|T|)$.

Finally, we prove that e^* induces the least ex-post fair matching outcome in M^*. Recall in the proof of Proposition 2 that the mass of students with an ex-post justified envy in the induced matching outcome can be calculated just by looking at the induced distribution $\xi(e^*)$, as follows:

$$E(e^*) = \sum_{t<t^*(e^*)} \lambda(I_t) - \sum_{t<t^*(e^*)} \sum_{k} \xi^t_k(e^*) = \sum_{t=1}^{t^*(e^*)-1} \lambda(I_t) - \sum_{t=1}^{t^*(e^*)-1} \sum_{k=k(e^*)(t-1)}^{k(e^*)(t)} \xi^t_k(e^*),$$

where $t^*(e^*)$ is the maximal number h of types T such that $\sum_k \xi^h_k(e^*) > 0$, and the last equality follows from Proposition 1.

Take any $e \in F$. Write $k(e) = k(\xi(e))$. From Proposition 1 and Lemma 12, we have $k(e^*)(t) \leq k(e)(t)$ for each $t \in T$ and $t^*(e^*) \geq t^*(e)$ because $\xi(e) = \xi(|T|)$. Therefore, if $t = t^*(e^*) = t^*(e)$, since the distributions are non-wasteful from Proposition 1, we must have $E(e^*) \geq E(e)$, meaning that $\mu(e)$ is no less ex-post fair than $\mu(e^*)$.

Suppose next that $t^*(e^*) > t^*(e)$. Since $\xi(e)$ is non-wasteful, this inequality implies that

$$\sum_{t=1}^{t^*(e)-1} \sum_{k=k(e)(t-1)}^{k(e)(t)} \xi^t_k(e) = |C| - \sum_{k=k(e)(t^*(e)-1)}^{k(e)(t^*(e))} \xi^{t^*(e)}_k(e) \geq |C| - \lambda(I_{t^*(e)}),$$

where Proposition 1 implies the last inequality. Note here that we use the fact that the total capacities of the schools equals $|C|$. From $t^*(e^*) > t^*(e)$, we get $t^*(e^*) - 1 \geq t^*(e)$, and thus

$$\begin{aligned}
E(e^*) &\geq \sum_{t=1}^{t^*(e)} \lambda(I_t) - \sum_{t=1}^{t^*(e^*)-1} \sum_{k=k(e^*)(t-1)}^{k(e^*)(t)} \xi^t_k(e^*) \\
&\geq \sum_{t=1}^{t^*(e)-1} \lambda(I_t) + \lambda(I_{t^*(e)}) - |C| \\
&\geq \sum_{t=1}^{t^*(e)-1} \lambda(I_t) - \sum_{t=1}^{t^*(e)-1} \sum_{k=k(e)(t-1)}^{k(e)(t)} \xi^t_k(e) = E(e).
\end{aligned}$$

In conclusion, we get $E(e^*) \geq E(e)$. Therefore, the matching outcome $\mu(e)$ is no less ex-post fair than $\mu(e^*)$. This is the end of the proof. $\square$

To prove Proposition 5, we first need to pin down the equilibrium signal-independent strategies of the second stage game without updated beliefs. The next lemma answers this question. A strategy profile in the second-stage game described in Section 4 is a *weak Perfect Bayesian equilibrium*, if every player plays a best reply at all information sets, and every player forms an updated belief at all information sets on the path of play according to the Bayes' rule. The former optimality condition is often referred to as *sequential rationality*.

Lemma 15. Consider the second-stage game induced by a cutoff vector e. Provided that all students at all signals have the common prior as their belief, a strategy profile σ that is independent of signals satisfies sequential rationality in this sub-game, only if for all states θ, we have

$$\lambda(\{i \in I_t \mid \sigma_i(\pi(\theta)(i)) = c_k\}) = n_k^t(e^*),$$

for all t and k with $k(e^*)(t-1) \leq k < k(e^*)(t)$. Consequently, for each k and t, the expected mass of applicants off any equilibrium path is no less than $n_k^t(e^*)$.

Proof. Since the left hand side of the condition in the statement is independent of θ from the selection of strategies, we denote it by n_k^t. Suppose on the contrary that there exists t and k with $k(e^*)(t-1) \leq k < k(e^*)(t)$ such that $n_k^t \neq n_k^t(e^*)$. Among these pairs of t and k, take minimal k.

Now, since each student in I_t is completely homogeneous, i.e., they also share the same belief, all strategies in the support $\{c_k \in C \mid n_k^t > 0\}$ must be indifferent for the students I_t. Note that a strategy $c_{k^*(e^*)(t-1)} \in C$ always produce a strictly positive expected payoff under the non-updated belief, and thus we get $\sum_k n_k^t = \lambda(I_t)$, that is, no student in I_t applies to the outside option. Also, if it holds that $n_{k+1}^t > 0$ for some k, then we must have $n_k^t > 1$, because otherwise a student $i \in I_t$ with $\sigma_i = c_{k+1}$ has an obviously profitable deviation c_k. Consequently, from the construction of k, there exists some $k(t)$ such that we have

$$\frac{\xi_h^t(e^*)u_t(c_h)}{n_h^t} = \frac{\xi_{h+1}^t(e^*)u_t(c_{h+1})}{n_{h+1}^t} = u_t(c_{k(t)}),$$

for each h with $k(e^*)(t-1) \leq h < k(t)$. Since $\sum_h n_h^t = \lambda(I_t) = \sum_h n_h^t(e^*)$, the above equality and Lemma 14 together imply $n_k^t = n_k^t(e^*)$, a contradiction. □

It is not difficult to see that for any sub-game off the path of plays, there does exist a strategy profile of the students that satisfy sequential rationality under their beliefs being the common prior. In fact, neglecting signal realizations and viewing each student type as a representative player, the sub-game reduces to a normal form game with finite players and strategies. Since the payoff function of each representative player, defined as the aggregated payoffs of the students with the corresponding type, is linear in mixed strategy profiles, there exists a mixed-strategy Nash equilibrium. Then, the corresponding strategy profile that does not depend on signal realizations satisfies sequential rationality condition.

Now we are ready to prove Proposition 5 and Proposition 6.

Proof of Proposition 5. It is immediate from Lemma 15 that the pair of e^* and the straightforward strategies constitute an equilibrium. Lemma 14 then indicates that the equilibrium outcome equals the least ex-post fair matching outcome in M^*.

Take any equilibrium where the students play straightforward strategies on the path of plays. Let e be the profile of cutoffs on the equilibrium path. From Lemma 14, it is sufficient to prove $e_k = e_k^*$ for each $k < k(e^*)(|T|)$. To see this, suppose on the contrary that $e_k \neq e_k^*$ for some k. Take such minimal k. Let t solve $k(e^*)(t-1) \leq k < k(e^*)(t)$.

First, assume $e_k > e_k^*$ holds. Then, by the construction of k, we must have $n_k^t(e) > n_k^t(e^*)$. Since the straightforward strategy profile is a BNE in the cutoff signal e, Lemma 9 and Lemma 14 show that $k^*(e^*)(t) \neq k+1$ and that $n_{k+1}^t(e) > n_{k+1}^t(e^*)$. Repeating the process results in $k^*(e)(t) > |C|+1$, which is a contradiction. Next, assume that $e_k < e_k^*$ holds. It follows from the construction of k that $n_k(e) < n_k(e^*)$ holds, and thus any deviation of the school c_k is profitable from Lemma 15. This is again a contradiction.

Therefore, we must have $e_k = e_k^*$ for each $k < k(e^*)(T)$. Then, the cutoff vector e and the straightforward strategies induce a matching outcome, which from Lemma 14 and $\xi(e) = \xi(e^*)$ is the least ex-post fair matching outcome in M^*. This is the end of the proof. □

Proof of Proposition 6. Let e^* solve $\xi(e^*) = \xi(t^*)$. Take any cutoff vector $e \in F$. We write $k(e) = k(\xi(e))$ and $k(e^*) = k(\xi(e^*))$ in this proof for convenience. First note that Lemma 12 implies $k(e^*)(t^* - 1) \leq k(e)(t^* - 1)$. Let $K(e)$ be the maximal number k such that $\sum_t \xi_k^t(e) > 0$. Then, Proposition 1 and Lemma 12 show that $K(e) \leq K(e^*)$ must hold.

Now, since the distributions are non-wasteful from Proposition 1, the objective values under respective cutoff vectors are calculated as follows:

$$V(e) = \left(\sum_{t\in T} \xi^t_{k(e)(t^*-1)}(e)\right) u(c_{k(e)(t^*-1)}) + \sum_{k=k(e)(t^*-1)+1}^{K(e)-1} u(c_k) + \left(\sum_{t\in T} \xi^t_{K(e)}(e)\right) u(c_{K(e)}).$$

If it holds that $k = k(e^*)(t^*-1) = k(e)(t^*-1)$, then the definition of the algorithm shows that we must have $\xi_k^{t^*}(e) \leq \xi_k^{t^*}(e^*)$. Analogous rationale shows that $k = K(e) = K(e^*)$ implies $\xi_k^{t^*}(e) \leq \xi_k^{t^*}(e^*)$. Therefore, again from the non-wastefulness condition, we get

$$V(e) \leq \sum_{k=k(e^*)(t^*-1)}^{k(e)(t^*-1)} \left(\sum_{t\in T} \xi^t_k(e)\right) u(c_k) + \sum_{k=k(e)(t^*-1)+1}^{K(e)-1} u(c_k) + \sum_{k=K(e)}^{K(e^*)} \left(\sum_{t\in T} \xi^t_k(e)\right) u(c_k),$$

where the right hand side of this inequality equals $V(e^*)$ from Proposition 1. Thus, the cutoff vector e^* is a solution to the given problem. The second argument follows immediately from Lemma 13, completing the proof. □

B.5 Proof of Proposition 7 and Proposition 8

Proposition 7 relies on the observation that the honest reports are always optimal strategies, even under incomplete information: if a strategy profile does not constitute an ex-post equilibrium under the fully-revealing information structure, some students do not play best replies at some states. This in turn implies that the honest report provides the students with strictly higher payoffs at these states, without harming payoffs earned from other states.

Proof of Proposition 7. The last claim is trivial, once we notice that the profile of honest reports always constitutes a BNE for any information structure. Thus, we only prove the first claim. Take any signal $\Pi = (S, \pi)$ and BNE $\sigma : S \to \mathcal{R}$. It suffices to show $\sigma^* = \sigma \circ \pi$ is an *essentially* ex-post equilibrium under the fully-revealing information structure: for each $i \in I$, the strategy $\sigma_i^*(\theta)$ is a best reply to $\sigma^*_{-i}(\theta)$ in the preference revelation game induced from $\varphi(\cdot, \theta)$ for almost all $\theta \in \Theta$.

Suppose, on the contrary, that there exists a student $i \in I$ who does not play best replies in a set Θ^* of states such that $p(\Theta^*) > 0$. Then, there must exist a signal realization $s_i \in S_i$ of the student i such that $p(\Theta^* \mid s_i) > 0$.

Now, the definition of Θ^* leads to the existence of state-wise best replies on the restricted domain, $\bar{\sigma}_i : \Theta^* \to \mathcal{R}_i$, such that $\bar{\sigma}_i(\theta) \neq \sigma^*_i(\theta)$ for all $\theta \in \Theta^*$. Then, since φ is strategy-proof, we have the following:

$$\begin{aligned}
&\int_{\theta\in\Theta} p_i(\theta \mid s_i) u(\varphi(R^*, \sigma^*_{-i}(\theta), \theta)) d\theta \\
\geq &\int_{\theta\in\Theta^*} p_i(\theta \mid s_i) u(\varphi(\bar{\sigma}_i(\theta), \sigma^*_{-i}(\theta), \theta)) d\theta + \int_{\theta\in\Theta\setminus\Theta^*} p_i(\theta \mid s_i) u(\varphi(\sigma^*(\theta), \theta)) d\theta \\
> &\int_{\theta\in\Theta} p_i(\theta \mid s_i) u(\varphi(\sigma^*(\theta), \theta)) d\theta.
\end{aligned}$$

This contradicts with the assumption that σ is a BNE. Therefore, σ^* is an essentially ex-post equilibrium, which completes the proof. □

The proof of Proposition 8 relies on the following simple observation: The Boston algorithm at any realization of any equilibrium strategy profile must terminate at the first round, given the sufficient diversity of student types. Before formally proving this conclusion, we need one preparation. The next lemma claims that implementable distributions under the Boston mechanism have a property that is a variant of non-wastefulness. Building on the observation in the proof of Theorem 1, we assume without loss of generality that for any information structure $\Pi = (S, \pi)$ considered in the subsequent proofs, the set S_i is finite and $p(s_i) > 0$ for each $s_i \in S_i$.

Lemma 16. Assume that $|T| > |C|$ and $\lambda(I_t) > 1$ for all $t \in T$. Suppose that a state-independent distribution $\xi \in \mathbb{R}^{|C|\times|T|}$ is implementable under the Boston mechanism φ^B. Then, it holds that $\sum_{t=1}^{k} \xi^t_k = 1$ for all $k \in T$ with $k \leq |C|$.

Proof. Let $\Pi = (S, \pi)$ and σ implement ξ under φ^B. Seeking a contradiction, take the minimal $k \in T$ that violates the condition given in the statement. Then, the assumption implies

$$\xi^k_k \leq \sum_{t=1}^{k} \xi^t_k < 1 < \lambda(I_k).$$

Therefore, from Lemma 3, there exists a set N of students in I_k with $\lambda(N) > 0$ such that each student $i \in N$ does not match with the school c_k under a signal $s_i \in S_i$. Meanwhile, the definition of $k \in T$ implies that

$$0 = \int_{\Theta} \sum_{h<k} \xi_h^k dp \geq \int_{\theta\in\Theta} \int_{i\in N} \sum_{h<k} \mathbb{I}_i(c_h, \theta) d\lambda(i) dp(\theta) = \int_{i\in N} \int_{\theta\in\Theta} \sum_{h<k} \mathbb{I}_i(c_h, \theta) dp(\theta) d\lambda(i),$$

where $\mathbb{I}_i$ is the indicator function that is defined in the proof of Lemma 6. Since the left hand side of the inequality equals zero, for any small $\varepsilon > 0$, there must exist a student $i \in N$ such that

$$\int_{\theta\in\Theta} \sum_{h<k} \mathbb{I}_i(c_h, \theta) dp(\theta) \leq \varepsilon,$$

that is, the student $i \in N$ matches with a school preferred to the school c_k with probability zero. Take such student $i \in N$. Note also that i matches with some school c_h with $h < k$ with probability zero conditional on the signal s_i as well, because $p(s_i) > 0$.

Now, consider a deviation $R_i^1 = c_k$ for the student i at s_i. Since $\sum_{t=1}^{k} \xi_k^t < 1$, the student matches with the school c_k with probability one, which makes i strictly better off, compared to the original strategy σ_i under which the interim expected payoff is at most $\varepsilon \cdot u_k(c_1) + (1-\varepsilon) \cdot u_k(c_{k+1})$. This contradicts the strategy profile σ being a BNE, completing the proof. □

Likewise, the above conclusion follows under the decentralized mechanism φ^D. The proof needs no technical modification and thus is omitted.

Lemma 17. Assume that $|T| > |C|$ and $\lambda(I_t) > 1$ for all $t \in T$. Suppose that a state-independent distribution ξ is implementable under the decentralized mechanism φ^D. Then, it holds that $\sum_{t=1}^{k} \xi_k^t = 1$ for all $k \in T$ with $k \leq |C|$.

The next proposition parallels the observations stated in the literature: under complete information environment with finite students, all Nash equilibrium matchings are achieved with strategy profiles that list at most one school. See Pathak and Sönmez (2008), Haeringer and Klijn (2009), and Kojima (2012), among others. Although we need a set of assumptions for this result to hold, it could often be the case in which there are sufficiently many students with different types in reality.

Lemma 18. Suppose $|T| > |C|$ and $\lambda(I_t) > 1$ for all $t \in T$. If an information structure $\Pi = (S, \pi)$ implements a state-independent distribution $\xi \in \mathbb{R}^{|C|\times|T|}$ under the Boston mechanism φ^B, then there exists a strategy profile $\sigma : S \to \mathcal{R}^1$ such that Π and σ implement the distribution ξ under φ^B. In particular, for any BNE, the Boston algorithm ends in the first round almost surely. Consequently, the information structure Π and σ implement ξ under the decentralized mechanism φ^D as well.

Proof. Let $\Pi = (S, \pi)$ and σ implement ξ. The proof consists of two parts. First, we prove that all students match with their first choice almost surely whenever they could match. Second, we prove the *truncation* of the original strategy profile σ, which makes all schools except for the first choice unacceptable while keeping other relative orderings being equal, still constitutes a BNE.

We start with introducing some notations as a preparation. Let $\Theta^1 \subset \Theta$ be the set of all states under which the Boston algorithm terminates in the first round. Also, define $\Theta^1(k)$ to be the set of all states under which the school c_k fills its unit capacity in the first round of the Boston algorithm. Note that we have $\bigcap_{k=1}^{|C|} \Theta^1(k) \subset \Theta^1$ by their definitions.

We first prove that $p(\Theta^1) = 1$. Suppose on the contrary that $p(\Theta^1) < 1$. Then, there must exist a school c_k such that $p(\Theta^1(k)) < 1$, because otherwise

$$p(\Theta \setminus \Theta^1) \leq p\left(\Theta \setminus \bigcap_{k=1}^{|C|} \Theta^1(k)\right) \leq p\left(\bigcup_{k=1}^{|C|} \Theta \setminus \Theta^1(k)\right) \leq \sum_{k=1}^{|C|} p(\Theta \setminus \Theta^1(k)) = 0$$

holds, hence a contradiction. Now, take any $i \in I_T$. The student i matches with no school almost surely from Lemma 16, and thus their interim expected utility equals zero at all signal realizations. Meanwhile, from the finiteness of S_i and the relation $p(\Theta \setminus \Theta^1(k)) > 0$, for some $s_i \in S_i$, their updated posterior belief satisfies $p(\Theta \setminus \Theta^1(k) \mid s_i) > 0$. It is easy to see that the deviation $R_i^1 = c_k$, which makes i match with c_k for a strictly positive probability, is profitable for i at s_i. Again, this contradicts the assumption that σ is a BNE, and thus $p(\Theta^1) = 1$ must hold.

The above discussion shows that the Boston algorithm must terminate in the first round almost surely, implying that all students earn positive utilities only from their first choices. Finally, we show that the truncation of σ is also a BNE. To define the truncation formally, let $c_i(s_i)$ be the most preferred school (that can be the outside option) under the ranking $\sigma_i(s_i)$, for each $i \in I$ and

$s_i \in S_i$. Then, the truncation $\overline{\sigma} : S \to \mathcal{R}^1$ of the strategies σ is defined as $\overline{\sigma}_i(s_i) = c_i(s_i)$ for each i and s_i. It is clear that the truncation $\overline{\sigma}$ induces the state-independent distribution ξ almost surely.

For proving $\overline{\sigma}$ being a BNE, take any student $i \in I$, signal realization $s_i \in S_i$, and ranking $R_i \in \mathcal{R}_i$. Let c_i be the first choice at R_i. Since the Boston algorithm ends in the first round almost surely at σ, so does at $\overline{\sigma}$. Moreover, the student i is negligibly small; therefore, applications except for the first round do not yield any positive interim payoff to the student, even after the deviation.

Finally, if the deviation R_i is profitable at $\overline{\sigma}$, they gains a higher interim utility from the first application c_i. Therefore, the proceeding discussions imply that R_i is also a profitable interim deviation at σ. Summarizing, the truncation $\overline{\sigma}$ must also be a BNE, completing the proof. □

Proof of Proposition 8. If a state-independent distribution $\xi \in \mathbb{R}^{|C|\times|T|}$ is implementable under the Boston mechanism φ^B, then Lemma 18 implies that the distribution ξ is implementable under the decentralized mechanism φ^D. To see the converse, suppose that an information structure Π and a strategy profile σ implement a state-independent distribution $\xi \in \mathbb{R}^{|C|\times|T|}$. With an analogous argument in the proof of lemma 18, we may prove from Lemma 17 that the Boston algorithm under Π and σ must terminate in the first round almost surely.

It is sufficient to show that σ is a BNE under the Boston mechanism φ^B as well. Suppose on the contrary that there exists a profitable interim deviation R_i for a student i at a signal s_i. Since the Boston algorithm ends in the first round, the truncation $\overline{R}_i$ of R_i yields the same interim payoff to the student, which is thus a profitable deviation as well. This implies, however, that $\overline{R}_i$ is a profitable deviation also under the decentralized mechanism φ^D, contradicting the strategy profile σ being an equilibrium under φ^D. □

References

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

Abdulkadiroğlu, Atila and Tayfun Sönmez. 2003. "School choice: A mechanism design approach." *American economic review* 93 (3):729–747.

Akbarpour, Mohammad, Adam Kapor, Christopher Neilson, Winnie van Dijk, and Seth Zimmerman. 2022. "Centralized school choice with unequal outside options." *Journal of Public Economics* 210:104644.

Akyol, Ethem. 2022. "Ex-Ante Welfare Superiority of the Boston Mechanism Over the Deferred Acceptance Mechanism." *Dynamic Games and Applications* :1–32.

Alonso, Ricardo and Odilon Câmara. 2016. "Persuading voters." *American Economic Review* 106 (11):3590–3605.

Arieli, Itai and Yakov Babichenko. 2019. "Private bayesian persuasion." *Journal of Economic Theory* 182:185–217.

Artemov, Georgy. 2021. "Assignment mechanisms: common preferences and information acquisition." *Journal of Economic Theory* 198:105370.

Au, Pak Hung and Mark Whitmeyer. 2018. "Attraction Versus Persuasion." *arXiv preprint arXiv:1802.09396* .

Avery, Christopher, Soohyung Lee, and Alvin E Roth. 2014. "College admissions as non-price competition: The case of South Korea." Tech. rep., National Bureau of Economic Research.

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

Bergemann, Dirk, Paul Duetting, Renato Paes Leme, and Song Zuo. 2022a. "Calibrated Click-Through Auctions." .

Bergemann, Dirk, Tibor Heumann, Stephen Morris, Constantine Sorokin, and Eyal Winter. 2022b. "Optimal information disclosure in auctions." .

Bergemann, Dirk and Stephen Morris. 2016. "Bayes correlated equilibrium and the comparison of information structures in games." *Theoretical Economics* 11 (2):487–522.

———. 2019. "Information design: A unified perspective." *Journal of Economic Literature* 57 (1):44–95.

Chan, Jimmy, Seher Gupta, Fei Li, and Yun Wang. 2019. "Pivotal persuasion." *Journal of Economic theory* 180:178–202.

Chen, Li. 2018. "Timing of preference submissions under the Boston mechanism." *Available at SSRN 3017156* .

Chen, Li and Juan Sebastián Pereyra. 2019. "Self-selection in school choice." *Games and Economic Behavior* 117:59–81.

Chen, Yan and Yinghua He. 2021. "Information acquisition and provision in school choice: an experimental study." *Journal of Economic Theory* 197:105345.

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

Choi, Michael, Kyungmin Kim, and Marilyn Pease. 2019. "Optimal information design for search goods." In *AEA Papers and Proceedings*, vol. 109. 550–56.

Dasgupta, Sulagna. 2022. "Information Design in One-sided Matching Markets." .

Dogan, Mustafa and Ju Hu. 2022. "Consumer search and optimal information." *The RAND Journal of Economics* .

Gale, David and Lloyd S Shapley. 1962. "College admissions and the stability of marriage." *The American Mathematical Monthly* 69 (1):9–15.

Gentzkow, Matthew and Emir Kamenica. 2016. "Competition in persuasion." *The Review of Economic Studies* 84 (1):300–322.

———. 2017. "Bayesian persuasion with multiple senders and rich signal spaces." *Games and Economic Behavior* 104:411–429.

Goldstein, Itay and Chong Huang. 2016. "Bayesian persuasion in coordination games." *American Economic Review* 106 (5):592–96.

Goldstein, Itay and Ady Pauzner. 2005. "Demand–deposit contracts and the probability of bank runs." *the Journal of Finance* 60 (3):1293–1327.

Haeringer, Guillaume and Flip Klijn. 2009. "Constrained school choice." *Journal of Economic theory* 144 (5):1921–1947.

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

Hafalir, Isa E, M Bumin Yenmez, and Muhammed A Yildirim. 2013. "Effective affirmative action in school choice." *Theoretical Economics* 8 (2):325–363.

Halac, Marina, Ilan Kremer, and Eyal Winter. 2021. "Monitoring Teams."

Halac, Marina, Elliot Lipnowski, and Daniel Rappoport. 2020. "Rank Uncertainty in Organizations." *Forthcoming in American Economic Review* .

———. 2022. "Addressing Strategic Uncertainty with Incentives and Information." In *AEA Papers and Proceedings*, vol. 112. 431–37.

Harless, Patrick and Vikram Manjunath. 2015. "The importance of learning in market design." Tech. rep., working paper, University of Rochester.

Hoshino, Tetsuya. 2022. "Multi-Agent Persuasion: Leveraging Strategic Uncertainty." *International Economic Review* 63 (2):755–776.

Immorlica, Nicole, Jacob Leshno, Irene Lo, and Brendan Lucier. 2020. "Information acquisition in matching markets: The role of price discovery." *Available at SSRN 3705049* .

Inostroza, Nicolas and Alessandro Pavan. 2022. "Adversarial Coordination and Public Information Design." .

Kajii, Atsushi and Stephen Morris. 1997. "The robustness of equilibria to incomplete information." *Econometrica: Journal of the Econometric Society* :1283–1309.

Kamenica, Emir. 2019. "Bayesian persuasion and information design." *Annual Review of Economics* 11:249–272.

Kamenica, Emir and Matthew Gentzkow. 2011. "Bayesian persuasion." *American Economic Review* 101 (6):2590–2615.

Ke, T Tony, Song Lin, and Michelle Y Lu. 2022. "Information design of online platforms." *Available at SSRN* .

Kojima, Fuhito. 2012. "School choice: Impossibilities for affirmative action." *Games and Economic Behavior* 75 (2):685–693.

Kojima, Fuhito and M Utku Ünver. 2014. "The "Boston" school-choice mechanism: an axiomatic approach." *Economic Theory* 55 (3):515–544.

Li, Fei, Yangbo Song, and Mofei Zhao. 2019. "Global Manipulation by Local Obfuscation: Information Design in Coordination Games." *Available at SSRN* .

Lien, Jaimie W, Jie Zheng, and Xiaohan Zhong. 2017. "Ex-ante fairness in the Boston and serial dictatorship mechanisms under pre-exam and post-exam preference submission." *Games and Economic Behavior* 101:98–120.

Mathevet, Laurent, Jacopo Perego, and Ina Taneva. 2020. "On information design in games." *Journal of Political Economy* 128 (4):1370–1404.

Moriya, Fumitoshi and Takuro Yamashita. 2020. "Asymmetric-information allocation to avoid coordination failure." *Journal of Economics & Management Strategy* 29 (1):173–186.

Morris, Stephen, Daisuke Oyama, and Satoru Takahashi. 2020. "Implementation via Information Design in Binary-Action Supermodular Games." Tech. rep.

———. 2022a. "Implementation via Information Design using Global Games." *Available at SSRN* .

———. 2022b. "On the Joint Design of Information and Transfers." *Available at SSRN* .

Noda, Shunya. 2022. "Strategic experimentation with random serial dictatorship." *Games and Economic Behavior* 133:115–125.

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

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

Sandmann, Christopher. 2020. "Recursive Information Design." Tech. rep., Toulouse School of Economics.

Shapley, Lloyd and Herbert Scarf. 1974. "On cores and indivisibility." *Journal of mathematical economics* 1 (1):23–37.

Sun, Yeneng. 2006. "The exact law of large numbers via Fubini extension and characterization of insurable risks." *Journal of Economic Theory* 126 (1):31–69.

Taneva, Ina. 2019. "Information design." *American Economic Journal: Microeconomics* 11 (4):151–85.

Troyan, Peter. 2012. "Comparing school choice mechanisms by interim and ex-ante welfare." *Games and Economic Behavior* 75 (2):936–947.