

# School Choice with Asymmetric Information: Priority Design and the Curse of Acceptance

Andrew Kloosterman  
Department of Economics  
University of Virginia

Peter Troyan\*  
Department of Economics  
University of Virginia

April 23, 2018

## Abstract

An implicit assumption in most of the matching literature is that all participants know their preferences. If there is variance in the effort agents spend researching options, some may know more about their preferences, while others may know less. When this is true, strategizing is complex, (ex-post) stable outcomes need not exist, and informed agents gain at the expense of less informed agents, outcomes we attribute to a *curse of acceptance*. However, when priorities are designed so that all agents have a secure school, we recover positive results: equilibrium strategies are simple, the outcome is ex-post stable, and less informed students are protected from the curse of acceptance, which makes them better off. Our results have potential policy implications for the current debate in school choice over how priority design affects outcomes.

## 1 Introduction

*“Please accept my resignation. I refuse to join any club that would have me as a member.”*

*-Groucho Marx*

In the last decade, school choice has rapidly expanded across the United States and around the world,<sup>1</sup> which has led to a vast and rapidly expanding economics literature, encompassing a wide array of theoretical, practical, and empirical papers. On the mechanism

---

\*Authors' names appear in random order. We would like to thank (also in random order) Al Roth, Assaf Romm, and Fuhito Kojima, for helpful comments, as well as the audiences at the University of Zurich and the 14th Matching in Practice workshop. We gratefully acknowledge the Bankard Fund for Political Economy for supporting this research. Address: P.O. Box 400182, Charlottesville, VA, 22904. Emails: ask5b@virginia.edu and troyan@virginia.edu.

<sup>1</sup>In a 2012 survey, almost 40% of parents in the United States reported having access to public school choice, a number which has certainly only increased since (National Center For Education Statistics, 2015).

design side of this literature, the standard modeling approach emanates from the seminal paper of Abdulkadiroğlu and Sönmez (2003), and is the analogue of what the broader mechanism design/auctions literature refers to as *private values*: each student is perfectly informed about her own (usually strict) ordinal preference relation over all schools and is asked to report these preferences to a mechanism, which determines the assignment.<sup>2</sup> While this simplifies the analysis, it abstracts away from differentially informed students that likely exists in practice. For example, a key component of most school choice markets is a pre-assignment information acquisition stage, during which some parents may spend considerable time researching and visiting schools to learn about their quality.<sup>3</sup> However, not all parents engage in such activities, perhaps due to a lack of time or other resources, as captured by the following article from *US News & World Report*:

School choice policies are predicated on the assumption that parents have enough information to make an informed decision on where to send their children. However, the vast majority of school districts do not provide families with robust school quality information – in part because school quality is difficult to measure and portray clearly to parents and students. Even districts with large choice programs, like Denver and New Orleans, often struggle to provide families with the information they need to make optimal decisions. Research also shows that districts don’t always do a good job of disseminating information and explaining options to certain groups – low-income parents and those whose children are first-generation Americans, for example. (McEachin et al., 2015)

The standard private values framework is unable to capture the types of preference interdependencies and informational asymmetries suggested by the above. In fact, a closer theoretical analogue might be the other important model routinely studied in auctions and mechanism design, *common values*: every agent has the same value for the object, but agents may have different information about what this value is. This manifests in school choice if all parents care about a school’s quality, but some are more informed about which schools are higher quality than others. More generally, quality might be one of several factors that

---

<sup>2</sup>Mechanism design and auctions also usually consider independence in this setting. This has not previously been seen as relevant in most of the matching literature, as the underlying distribution from which preferences are drawn is not usually modeled. While this works with private values, in order to understand and analyze correlations in preferences, as we will do in this paper, some type of distributional assumptions will obviously be necessary.

<sup>3</sup>The importance of information acquisition is underscored by the existence of popular websites such as Greatschools.org and Insideschools.org, which aim to provide independent reviews of public and private schools to help parents make their choices. The latter, which is focused exclusively on evaluating schools in New York City, states that they send staff to “visit hundreds of schools each year and interview thousands of people—principals, teachers, students, and parents” because “test scores don’t tell the whole story”.

influence student preferences (other examples include things like location, sibling attendance, language programs, etc.), and so student preferences are likely correlated, although perhaps not perfectly so. The goal of this paper is to re-examine school choice in a more realistic, interdependent values framework.<sup>4</sup>

Perhaps the most well-known result from the literature on auctions with interdependent values is the so-called winner’s curse. We start by identifying a related “curse” in our environment, which we call the *curse of acceptance*: upon observing their assignment, less informed students update their belief about the quality of their assigned school downwards. Intuitively, the more informed students will submit preferences with high quality schools at the top of their lists, leaving more empty seats at the low quality schools for the less informed students. While the less informed may not know which schools are high quality or low quality, they are at least aware that there are others who do have such knowledge, and so they infer that any school to which they are eventually assigned is likely to be of lower quality on average. In short, just like Groucho, less informed students would prefer not to enroll at the school that admits them.

This environment leads to the failure of many standard properties of well-known assignment mechanisms, such as strategy-proofness and stability.<sup>5</sup> First, strategy-proofness as conventionally understood is not feasible (for any mechanism), as some students do not know their true preferences.<sup>6</sup> In fact, as we show, determining an optimal strategy can be very complex for the less informed students, and dominant strategies generally will not exist. Second, deferred acceptance (DA), the standard stable mechanism in the usual model and the most popular mechanism in applications, may no longer produce a stable match. Indeed, the standard definition of stability no longer applies when students may not know their preferences. We first introduce a definition of *ex-post* stability that requires that no student’s expected utility from successfully rematching, conditional on her information, is higher than her expected utility from remaining at her assigned school. We show that ex-post stability

---

<sup>4</sup>Much of the mechanism design literature also assumes that all agents get different signals but from the same underlying information structure. In school choice, it is likely that some parents might be more or less informed than others, as highlighted by the above quote. Thus, it is natural to think of students as getting signals from different information structures, and so this is the assumption we will pursue.

<sup>5</sup>Strategy-proofness means that in the game induced by the assignment mechanism, it is a dominant strategy for every student to report her true preferences. Stability means that no student prefers some school to her assignment and has higher priority than a student who is assigned to that school. More formal definitions can be found below.

<sup>6</sup>Pathak and Sönmez (2008) argue that one of the main advantages of strategy-proof mechanisms is that they “level the playing field” between what they refer to as sophisticated students (who strategize optimally) and sincere students (who always report their true preferences, whether it is in their interest or not). Note, however, that they still follow the standard modeling approach and assume that all sincere students know their own preferences. When these assumptions do not hold, strategy-proof mechanisms may no longer “level the playing field”. See also Pathak and Sönmez (2013).

may not hold in general, and attribute this to the curse of acceptance: ex-post, less informed students are able to infer that other schools are likely better than their assignment, and they may have high enough priority at one of these schools to block the original match.

Our negative results rely on showing that desirable properties fail for some specific priority structure. In the second half of the paper, we show these negative results motivate looking at the problem in a new light, as one of priority design, and show that if priorities are designed in an appropriate way, positive results can be recovered. In particular, we introduce the idea of a *secure school*, which is a school  $s$  such that a student  $j$ 's priority ranking at  $s$  is below the capacity of  $s$ . In other words, it is a school that  $j$  can be certain she will receive (under any stable mechanism), so long as she ranks it first.<sup>7</sup> We consider what happens when the priority structure is designed so that each student has a secure school in two models, perfectly correlated preferences in the standard finite economy model and imperfectly correlated preferences in a model with a continuum of students. In both cases, we show that ex-post stable matchings exist, and the deferred acceptance mechanism finds one. Additionally, while DA will still not be strategy-proof, the equilibrium strategies are simple and focal.

Last, and perhaps most importantly, our informational environment highlights another new issue not found in the earlier school choice literature relating to the welfare consequences of school choice. An important goal of school choice in general is to provide all students (and in particular, those in poorer neighborhoods) with a fair chance of attending a good school. However, if students are differentially informed, exactly the opposite of this intended effect may occur: that is, the informed students will, on average, be assigned to better schools at the expense of the less informed, who may be worse off. Indeed, in our continuum model, introducing school choice is actually a Pareto *dis*-improvement for the less informed students when there are no secure schools: in other words, because of the curse of acceptance, less informed students may actually be worse off *than if there were no school choice at all*. However, when all students have a secure school, this is no longer true, and introducing school choice is a Pareto improvement. In particular, we show that the informed students can use their information to improve their own assignment, but not at the expense of the uninformed students. Thus, the use of secure schools can be seen as a practical compromise: it allows for choice for students who know their preferences, which increases efficiency, while still protecting the less informed students from the unintended consequences of the curse of

---

<sup>7</sup>The idea of a secure school is reminiscent of “neighborhood schools” or “walk-zone schools” that are often discussed in the literature. However, we emphasize that secure schools *need not* be neighborhood schools, and can be based on many other criteria besides geography. For example, a school district may divide the schools into certain quality levels, and allocate priorities based on this metric (as has been done recently in Boston). This is discussed in further detail at the end of this introduction.

acceptance.

While there is a rich school choice literature comparing different mechanisms (Boston vs. DA vs. Top Trading Cycles, etc.), there is far less work devoted to the topic of how to design the priority structure that the mechanism will use, which is arguably just as important (if not moreso) as the choice of mechanism.<sup>8</sup> Most formal models simply take priorities as given, but in reality, many school districts have the ability to design the priorities to achieve certain goals. For instance, Boston recently undertook a major redesign of the menu of schools from which students are able to choose. Previously, the district had been divided geographically, and a student’s menu of possible choices she could list depended only on where she lived. After intense debate, the city eventually settled on a new plan proposed by Shi (2015). Roughly, schools are divided into quality tiers, and a student’s menu consists of the two closest Tier 1 schools, the four closest Tier 2 schools, etc.; if a school is not in a student’s menu, then she is prohibited from listing it as a choice at all.

While Boston’s redesign was done with many objectives in mind,<sup>9</sup> note that restricting the set of students who can apply to a school effectively gives high priority to those who are allowed to list it, which is similar in spirit to giving each student a “good” (i.e., Tier 1) secure school in our model. In order to provide a broader theoretical insight, we abstract away from particular institutional details that apply only to certain cities. Rather, our formal model is intended to capture salient features that are present in many real-world markets, yet are absent from the standard framework, while still remaining tractable enough to produce clear results. Boston’s priority redesign plan was undertaken for a variety of other reasons, but our paper shows that there may be an additional benefit if parents are differentially informed, which is likely true in practice: it mitigates the curse of acceptance for less informed students and more evenly distributes the efficiency gains from introducing school choice. This is something that could not be captured in standard models that take the priority structure as a given and assume private values and perfectly informed students. We hope that our model brings attention to this important informational issue and its distributional consequences that have heretofore been mostly overlooked, and will be a useful step towards better-designed school choice markets in practice.

---

<sup>8</sup>In a survey of the school choice literature, Pathak (2016) argues that much of research comparing different mechanisms has turned out to be not as important for practical implementation as it was thought at the time, and that “the criteria used to allocate seats [the priorities] - taken as given by much of the literature - are just as important as how the mechanism processes applicants’ claims”. While most school districts seem to have converged on DA as the choice mechanism, the question of how to optimally design the priorities is far less settled.

<sup>9</sup>See Shi (2015) and Pathak and Shi (2017) for further details.

## Related Literature

The only other papers we are aware of that take a similar “mechanism design” approach to studying matching under interdependent values are Chakraborty, Citanna, and Ostrovsky (2010) (hereafter, CCO) and a follow-up paper by the same authors, Chakraborty, Citanna, and Ostrovsky (2015).<sup>10</sup> The main thrust of these papers is that stability is difficult to achieve, which is shown using several possible definitions of stability and providing impossibility results for each. However, there are several key distinctions between our work and theirs. First, their model assumes agents report an arbitrary vector of signals and that the mechanism designer has knowledge of the signal space and agents’ expected utilities as a function of everyone’s signals, and all of this information can be used in determining the allocation. We, on the other hand, only consider mechanisms that ask students to report ordinal preferences, which is the case in most real-world school choice markets. Second, even though CCO similarly find that stability may fail under several different definitions, even under their weak stability (which is the closest-related definition to our ex-post stability), the reasoning is distinct from that in our paper. Specifically, their example exhibits failure of stability by having an informed party who wants to lie about his preferences to mislead an uninformed party into matching with an undesirable agent. In our model, informed students have a dominant strategy to truthfully reveal, and stability fails due to the curse of acceptance outlined above. Last, CCO are only able to recover positive results by imposing a strict assumption, which in our environment would translate to a common priority relation that all schools use to admit students. Under this assumption, they find that a serial dictatorship is stable in their sense. Our positive results allow for a larger class of admissible priority structures, and, as we will argue, this has important implications for the design of priorities in school choice markets and the resulting welfare of the students.

Another line of literature consists of papers that have looked at interdependent values in specific decentralized matching games. For instance, Chade (2006) identifies a phenomenon he calls the “acceptance curse effect” in a dynamic marriage model where men and women randomly meet each period and must decide whether to accept their current partner and leave the market or wait and get a new draw in the next period. While somewhat related to our curse of acceptance, the details of its manifestation, as well as his overall motivation and model, are quite different from ours, and thus the results are not formally comparable. Lee (2009) analyzes early acceptance programs in a model of decentralized college admissions, and

---

<sup>10</sup>Che et al. (2015) study efficiency and incentives in a pure object allocation problem (i.e., there are no priorities on the object side) with interdependent values, and show that in general, no Pareto efficient and ex-post incentive compatible mechanism exists. The absence of priorities in their model is a significant difference from ours, as they are not at all concerned about stability.

argues that such programs can actually increase efficiency if college valuations of students are interdependent. Chade et al. (2014) also study college admissions with noisy signals about student quality for the colleges and costly applications for the students, and the effect introducing these features has on application decisions and admissions standards in equilibrium. Finally, Che and Koh (2016) study college admissions when colleges have a limited capacity but there is aggregate uncertainty about how popular a college is amongst the students, and thus how likely it is that a student offered admission will accept. While they do assume preference interdependencies, all of these papers work in models of two colleges and consider decentralized admissions games, and so are quite different from our work.

## 2 Preliminaries

### 2.1 Model

Let  $J = \{j_1, \dots, j_N\}$  be a set of students and  $S = \{s_1, \dots, s_M\}$  be a set of schools. Each school  $s$  has capacity  $q_s$  and a priority relation  $\succ_s$ , where  $\succ_s$  is a strict, complete, and transitive binary relation over  $J$ . We write  $q = (q_s)_{s \in S}$  and  $\succ = (\succ_s)_{s \in S}$  to denote the capacity vector and priority profile for all schools, respectively. We assume that  $\sum_{s \in S} q_s \geq N$ .<sup>11</sup>

We model preference interdependencies by having the students' preferences depend on an underlying state  $\omega$ . The state space  $\Omega$  is finite, with an associated probability distribution  $\Pr : \Omega \rightarrow [0, 1]$  such that  $\sum_{\omega \in \Omega} \Pr(\omega) = 1$ . Every student has a (state-dependent) utility function, where  $u_j^\omega(s)$  is student  $j$ 's utility for school  $s$  in state  $\omega$ . We assume that preferences are strict, i.e., given a state  $\omega$ ,  $u_j^\omega(s) \neq u_j^\omega(s')$  for all  $s \neq s'$ . Students may have different levels of information about the underlying state (and therefore about their own ex-post preferences over schools). Where relevant, agents evaluate lotteries using von Neumann-Morgenstern preferences. In particular, a student's expected utility from matching to school  $s$  given information that the true state lies in some subset of the state space  $\mathcal{I}' \subseteq \Omega$ , is written

$$E(u_j^\omega(s) | \mathcal{I}') = \sum_{\omega \in \Omega} u_j^\omega(s) \Pr(\omega | \mathcal{I}'),$$

where  $\Pr(\omega | \mathcal{I}')$  is the posterior probability of the true state being  $\omega$  conditional on information  $\mathcal{I}'$ , and is obtained via Bayes' rule.

After Nature draws  $\omega$ , each student receives a signal  $\mathcal{I}_j(\omega) \subseteq \Omega$ , where  $\mathcal{I}_j(\omega)$  denotes the

---

<sup>11</sup>This is a natural assumption in school choice, where all students must legally be offered a seat in some school.

subset of states that are possible given  $j$ 's signal when the true state is  $\omega$ . In particular, the agents are partitioned into a set of informed students,  $I$ , and a set of uninformed students,  $U$ , where  $J = I \cup U$ . All students  $j \in I$  receive a perfectly informative signal of the state, while all students  $j \in U$  receive no signal (or receive a completely uninformative signal). That is, for the informed students  $\mathcal{I}_j(\omega) = \{\omega\}$  for all  $\omega \in \Omega$ , while for the uninformed students  $\mathcal{I}_j(\omega) = \Omega$  for all  $\omega \in \Omega$ .<sup>12</sup>

The state space  $\Omega$  defined here is quite flexible, but in what follows we will concretely define state spaces where the students' utility for a school is determined by an intrinsic quality, which is common to all students, and, later in the paper, also include an idiosyncratic component which is individual to each student. Informed students should be thought of as those whose parents have invested the time and energy to learn which schools are the high quality schools. Uninformed students also prefer schools with high intrinsic quality, though they are uncertain which schools these are at the time they will be asked to submit their preferences. While it is also possible with this model to define finer levels of "partial informativeness" on the part of the students, we use the model presented because it is the most parsimonious formulation that exhibits our main theoretical insights.

## 2.2 Matchings and mechanisms

A **matching** is a function  $\mu : J \cup S \rightarrow 2^{J \cup S}$  such that: (i)  $\mu(j) \in S$  for all  $j \in J$ , (ii)  $\mu(s) \subseteq J$  and  $|\mu(s)| \leq q_s$  for all  $s \in S$ , and (iii)  $\mu(j) = s$  if and only if  $j \in \mu(s)$ . Hereafter, we use the shorthands  $\mu_j$  for  $\mu(j)$  and  $\mu_s$  for  $\mu(s)$ . In words,  $\mu_j$  is the school assigned to student  $j$  and  $\mu_s$  is the set of students assigned to school  $s$ . Let  $\mathcal{M}$  denote the set of all possible matchings.

In order to implement a "good" matching, market organizers (e.g., school districts) must elicit the private information of the agents. In most real-world settings, the way this is done is by asking the agents to submit an ordinal preference relation over the set of schools, and then applying a particular mechanism to these preferences to determine the final matching. To model this formally, let  $\mathcal{P}$  denote the set of all strict ordinal preference relations over  $S$ . Given a preference relation  $P_j \in \mathcal{P}$ , we write  $sP_j s'$  to denote that  $s$  is strictly preferred to  $s'$ , and we write  $sR_j s'$  if either  $sP_j s'$  or  $s = s'$ . We will sometimes use the notation  $\text{rank}_{P_j}(s) = |\{s' : s'R_j s\}|$  to denote the ranking of  $s$  in the preference relation  $P_j$ . Let  $P = (P_j)_{j \in J}$  denote a profile of preference relations, one for each student. A **mechanism**

---

<sup>12</sup>Note that we model the informed students as observing the true state  $\omega$  precisely, which includes information about the preferences of all other agents. This is for notational convenience only; the mechanisms we consider will be "strategy-proof" as long as an agent is perfectly informed about her own preferences, and so whether she knows the preferences of the other students or not will be irrelevant.

is a function  $\psi : \mathcal{P}^N \rightarrow \mathcal{M}$ . We use the notation  $\psi_j(P)$  to denote student  $j$ 's assignment under mechanism  $\psi$  when the submitted reports are  $P$ ; analogously,  $\psi_s(P)$  denotes school  $s$ 's assignment.

A mechanism  $\psi$  induces a game in which the action space for each player is  $\mathcal{P}$ . A *strategy* for student  $j$  in this game is a mapping  $\sigma_j : \Omega \rightarrow \mathcal{P}$  that is measurable with respect to her information.<sup>13</sup> A profile of strategies  $\sigma = (\sigma_1, \dots, \sigma_N)$  is a **(Bayesian) equilibrium** of the game induced by mechanism  $\psi$  if

$$E[u_j^\omega(\psi_j(\sigma(\omega))) | \mathcal{I}_j(\omega)] \geq E[u_j^\omega(\psi_j(\sigma'_j(\omega), \sigma_{-j}(\omega))) | \mathcal{I}_j(\omega)]$$

for all other strategies  $\sigma'_j$ , all  $j \in J$ , and all  $\omega \in \Omega$ .

In standard matching models (i.e.,  $U = \emptyset$  and  $J = I$ ) a mechanism is said to be *strategy-proof* if, for each student, reporting her true preferences is a dominant strategy of the induced preference revelation game.<sup>14</sup> Strategy-proofness is a desirable property because it makes the mechanism simple for students to play. Strategy-proof mechanisms will continue to be “strategy-proof” for the informed students in our model, in that it will still be dominant for them to report their true ordinal preferences in every state. This is obvious as truthful reporting is optimal against all other students’ reports and so other students’ information, which determines their report, is irrelevant. For uninformed students, however, following the truthful strategy is not possible; indeed, the concept of a truthful strategy is not a meaningful one for uninformed students. More formally, the *truthful strategy*  $\sigma_j(\omega) = P_j(\omega)$ , where  $P_j(\omega)$  is  $j$ 's true ordinal preference in state  $\omega$ , is not measurable for uninformed students. Thus, we must use Bayesian equilibrium as our solution concept, which, as we shall soon show, can be far more complicated.

While there are many possible mechanism choices, for the remainder of this paper we focus in particular on the student-proposing deferred acceptance (DA) mechanism. There are several reasons for this. First, DA has a rich history in the literature, going back to the foundational work of Gale and Shapley (1962), and in their seminal paper, Abdulkadiroğlu and Sönmez (2003) propose DA as a promising solution for school choice markets. Second, in private values settings, DA satisfies many desirable properties; namely, it is (i) strategy-proof (ii) stable and (iii) Pareto dominates any other stable mechanism. Third, and most importantly, because of its appealing theoretical properties, DA is one of the most widely used mechanisms in practice, including in such large school districts as New York City,

---

<sup>13</sup>That is, for any two states  $\omega$  and  $\omega'$  where  $\mathcal{I}_j(\omega) = \mathcal{I}_j(\omega')$ , it must be that  $\sigma_j(\omega) = \sigma_j(\omega')$ . In particular, informed students can condition their reports on the state while uninformed students can not.

<sup>14</sup>Formally, a mechanism  $\psi$  is *strategy-proof* if, for any  $j$ , there do not exist preferences  $P_j, P'_j, P_{-j}$  such that  $\psi_j(P'_j, P_{-j}) P_j \psi_j(P_j, P_{-j})$ .

Boston, and New Orleans, as well as many others in the United States and around the world (see, for example, Abdulkadiroğlu et al. (2005a) and Abdulkadiroğlu et al. (2005b)). As most of the prior work and positive results for DA uses the standard private values model, it becomes particularly important to understand what happens to the appealing features of the DA mechanism when these assumptions are relaxed.

Here, we give a brief definition of the standard deferred acceptance mechanism as it applies to standard school choice problems (e.g., Abdulkadiroğlu and Sönmez, 2003). The inputs to the mechanism consist of an ordinal preference relation  $P_j$  for each student  $j$ .

## Deferred Acceptance

**Step 1** Each student  $j$  applies to her most preferred school according to  $P_j$ . Each school  $s$  considers the set of students who have applied to it, and tentatively admits the  $q_s$ -highest priority students according to  $\succ_s$  (if there are less than  $q_s$  applicants, it admits all students). All students not tentatively admitted to a school are rejected.

**Step  $k, k > 1$**  Any student who was rejected in the previous step applies to her most preferred school according to  $P_j$  that has not yet rejected her. Each school considers the pool of students tentatively held from the previous step plus any new applicants, and tentatively admits the  $q_s$ -highest priority students according to  $\succ_s$ . All students not tentatively admitted to a school are rejected.

The algorithm ends after the first step where no students are rejected.

## 2.3 Full matchings and ex-post stability

Suppose the state is  $\omega$ . In the classical matching literature, given a matching  $\mu$ , a student and a school  $(j, s)$  are called a **blocking pair** if (in our notation): (i)  $u_j^\omega(s) > u_j^\omega(\mu_j)$  and (ii) either  $|\mu_s| < q_s$  or  $j \succ_s j'$  for some  $j' \in \mu_s$ . A matching  $\mu$  is then called **stable** if there are no blocking pairs.<sup>15</sup> In the school choice literature, stability is sometimes called *fairness*, and is given a normative interpretation in that a stable outcome is desirable because no student will justifiably envy the school assignment of a student over which she has higher priority at that school. This is a crucial concern to many school districts, because they must be able to justify why one student is admitted to a school and another is rejected, or else be vulnerable

---

<sup>15</sup>An additional component of stability is that every student prefers her assigned school to being unmatched, and no school wants to unilaterally drop one of its assigned students. We assume that students find all schools acceptable, and schools find all students acceptable, so this part of stability is not an issue. We assume this only for ease of exposition; none of our results are driven by this assumption.

to legal action.<sup>16</sup> Since our main ideas may also apply to two-sided matching models, and the terminology is more familiar, we stick with the word “stability”.

The classical definition should still apply for the informed students. For the uninformed students, they do not know the state, and so part (i) of the definition of a blocking pair needs to be modified to an expected utility. One might think this should be an unconditional expectation, as the uninformed students know nothing about the state. However, blocking is a process that happens after the matching has been implemented and so uninformed students do actually have some information, namely, they know what school they were assigned to. Part (ii) also needs modification for the uninformed students, since having higher priority than some other student at a school is no longer a simple yes or no, but rather yes in some states and no in others. Both of these features can be folded into the conditional expected value for an uninformed student. In particular, the relevant expected utility for a student who is assigned to school  $s'$  and is considering blocking with school  $s$  should condition on states where the uninformed student is matched to the school  $s'$  and they have higher priority than some other student at school  $s$  (or  $s$  is not at full capacity).<sup>17</sup>

To capture this formally, we first introduce the concept of a full matching, which is a mapping from states to observed matchings. This is an important object in our setting, because strategies in the preference revelation game induce state-dependent (full) matchings. We then introduce the definition of ex-post stability, which is defined on full matchings, and extends the classical notion of stability to our environment.

**Definition 1.** A **full matching** is a function  $\mu : \Omega \rightarrow \mathcal{M}$  that assigns a matching to each state  $\omega \in \Omega$ .

We will use bold face type,  $\mu$ , to denote full matchings, and write  $\mu^\omega \in \mathcal{M}$  for the matching in state  $\omega$ , and  $\mu_j^\omega$  for the assignment of student  $j$  in matching  $\mu^\omega$ .

There are two pieces of information that a student can use to update her information and form a potential blocking pair: first, the original school to which she is assigned, and second, whether the school she wishes to form a blocking pair with will accept her. Given a

---

<sup>16</sup>The fact that deferred acceptance (which is stable, but inefficient) is preferred across the board to the top trading cycles (TTC) mechanism (which is efficient, but unstable) provides further evidence of the importance of fairness as a goal. See Abdulkadiroğlu et al. (2005a) and Abdulkadiroğlu et al. (2005b) for discussions of this decision in New York and Boston, respectively. Abdulkadiroğlu et al. (2017) discuss the example of New Orleans, who used TTC for one year before abandoning it in favor of DA.

<sup>17</sup>We take the traditional “game theory” approach here, meaning that the students play an equilibrium and so know the strategies and therefore final matchings that will occur in each state. Hence, the uninformed students know exactly in which states they will match to school  $s$  and exactly in which states they will have higher priority than some student at  $s'$ . Obviously, this may not be exactly true in the real world, but is the best way to formalize the ideas and, as shall be seen, does very well in highlighting the key issues that arise in the presence of information asymmetry.

full matching  $\mu$ , define the following two sets (to avoid notational clutter, we suppress the dependence of these sets on  $\mu$ ):

$$\begin{aligned}\mathcal{A}_j(s) &= \{\omega \in \Omega : \mu_j^\omega = s\} \\ \mathcal{B}_j(s) &= \{\omega \in \Omega : \mu_j^\omega \neq s \text{ and } |\mu_s^\omega| < q_s \text{ or } j \succ_s j' \text{ for some } j' \in \mu_s^\omega\}\end{aligned}$$

The set  $\mathcal{A}_j(s)$  is the set of states in which student  $j$  is assigned to  $s$ , while the set  $\mathcal{B}_j(s)$  is the set of states in which  $j$  has high enough priority to block with school  $s$ , either because she has higher priority over one of the current students of  $s$ , or  $s$  is not filled to capacity. Let  $\mathcal{C}_j(s', s) = \mathcal{A}_j(s') \cap \mathcal{B}_j(s)$ . In words,  $\mathcal{C}_j(s', s)$  is the set of states in which student  $j$  is assigned to  $s'$ , but could block with school  $s$ .

**Definition 2.** Given a full matching  $\mu$ , student-school pair  $(j, s)$  are an **ex-post blocking pair** if there exists a state  $\tilde{\omega}$  such that  $\mathcal{C}_j(\mu_j^{\tilde{\omega}}, s) \neq \emptyset$  and

$$E[u_j^\omega(s) | \mathcal{C}_j(\mu_j^{\tilde{\omega}}, s) \cap \mathcal{I}_j(\tilde{\omega})] > E[u_j^\omega(\mu_j^{\tilde{\omega}}) | \mathcal{C}_j(\mu_j^{\tilde{\omega}}, s) \cap \mathcal{I}_j(\tilde{\omega})].$$

Full matching  $\mu$  is **ex-post stable** if there are no ex-post blocking pairs.

To understand this definition, fix a state  $\tilde{\omega}$  and notice that the right side of the inequality is student  $j$ 's expected utility from her current assignment,  $\mu_j^{\tilde{\omega}}$ , given her information that the true state lies in the set  $\mathcal{C}_j(\mu_j^{\tilde{\omega}}, s) \cap \mathcal{I}_j(\tilde{\omega})$  (the expected values are taken over all states  $\omega \in \mathcal{C}_j(\mu_j^{\tilde{\omega}}, s) \cap \mathcal{I}_j(\tilde{\omega})$ ); within each expected value, the school remains fixed, while the utility function varies with  $\omega$ ). The left side of the inequality is  $j$ 's expected utility from any other school  $s$  with which she may potentially want to form a blocking pair. If the latter is greater than the former, then  $(j, s)$  is an ex-post blocking pair, and the full matching  $\mu$  is not ex-post stable.

Informed students know the state exactly,  $\mathcal{I}_j(\tilde{\omega}) = \{\tilde{\omega}\}$  for all  $\tilde{\omega}$ , and so for them  $\mathcal{C}_j(\mu_j^{\tilde{\omega}}, s) \cap \mathcal{I}_j(\tilde{\omega}) = \tilde{\omega}$  when  $\mathcal{C}_j(\mu_j^{\tilde{\omega}}, s) \neq \emptyset$ . Therefore, the inequality reduces to  $u_j^{\tilde{\omega}}(s) > u_j^{\tilde{\omega}}(\mu_j^{\tilde{\omega}})$  when  $j$  has high enough priority to form a blocking pair with  $s$ . That is, the definition reduces to that of a classical blocking pair for informed students. For uninformed students,  $\mathcal{I}_j(\tilde{\omega}) = \Omega$  (they initially have no information), and so for them  $\mathcal{C}_j(\mu_j^{\tilde{\omega}}, s) \cap \mathcal{I}_j(\tilde{\omega}) = \mathcal{C}_j(\mu_j^{\tilde{\omega}}, s)$ . The information  $\mathcal{C}_j(\mu_j^{\tilde{\omega}}, s)$  is student  $j$ 's updated information given both the observation of her current assignment *and* the knowledge of the set of states in which she has high enough priority at school  $s$  to form a blocking pair.

It should be clear that given a mechanism  $\psi$ , any strategy profile  $\sigma$  induces a full matching  $\mu(\sigma)$  defined by  $\mu^\omega(\sigma) = \psi(\sigma(\omega))$ . This leads naturally to the next definition.

**Definition 3.** Mechanism  $\psi$  is **ex-post stable** if there exists an equilibrium  $\sigma$  such that the induced full matching  $\mu(\sigma)$  is ex-post stable.

*Remark 1.* This is not the only way one could think of to define ex-post stability. One alternative considered by Chakraborty et al. (2010) is to allow students to observe the entire final matching, rather than just their own assignment. However, besides arguably being less realistic, as they show, under such a definition, almost nothing will be ex-post stable. The definitions of stability in Chakraborty et al. (2010) also combine incentives for reporting with anticipated rematching, i.e., they allow students to engage in “joint manipulations” where they may first lie about their preferences to make themselves worse off at the initial matching, but do so because they anticipate being able to re-match ex-post after the lie. Most of the (vast) literature on stable matching treats these two issues separately, and, indeed, as shown by Kojima (2011), even in the standard private values model, essentially no mechanism is immune to these types of joint manipulations (see also Afacan, 2012).<sup>18</sup>

The following example is instructive for understanding the concepts of full matchings and ex-post stability.

**Example 1.** There are 3 students,  $I = \{j_1\}$  and  $U = \{j_2, j_3\}$ , and 3 schools  $S = \{A, B, C\}$  each with capacity one. The state space  $\Omega$  is the set of all permutations of  $S$  (and therefore has size  $|\Omega| = M! = 6$ ), and the probability distribution  $\Pr$  over  $\Omega$  is uniform. Each state  $\omega \in \Omega$  is identified with an ordinal preference relation over the schools, and, given the state, each student has the same ordinal preferences. For shorthand, we will write  $\omega = ABC$  to refer to the state where  $A$  is the most preferred school,  $B$  is the second most preferred school, and  $C$  is the least preferred school for all students and likewise for the other 5 states. For concreteness, we define utility as  $u_j^\omega(s) = M - \text{rank}_\omega(s)$ , where  $\text{rank}_\omega(s)$  borrows the notation defined above for preference profiles and is the ranking of school  $s$  in permutation  $\omega$  (also, this functional form is not important; any utility function that preserves the common ordinal preferences would give the same results). We refer to this model as the *common ordinal preference model*.

The priority structure is as follows:

$\succ_A$	$\succ_B$	$\succ_C$
$j_1$	$j_1$	$j_2$
$j_2$	$j_3$	$j_3$
$j_3$	$j_2$	$j_1$

---

<sup>18</sup>Kojima (2011) provides a necessary and sufficient condition for a *robustly stable* (his terminology for a mechanism that is immune to the joint manipulations described) mechanism to exist, which is that the priority structure of the schools must be acyclic (Ergin, 2002). Acyclicity is an extremely demanding condition that is unlikely to be satisfied in most real-world markets.

In this example, a full matching  $\mu$  is a function that gives a matching  $\mu^\omega$  for each of the 6 possible states  $\omega \in \Omega$ , and so, in this example, there are 36 possible full matchings (6 possible matchings per state  $\times$  6 states). The following table presents one possible full matching  $\mu$ :

	$\mu_{j_1}^\omega$	$\mu_{j_2}^\omega$	$\mu_{j_3}^\omega$
$\omega = ABC$	$A$	$C$	$B$
$\omega = ACB$	$A$	$C$	$B$
$\omega = BAC$	$B$	$C$	$A$
$\omega = BCA$	$B$	$C$	$A$
$\omega = CAB$	$A$	$C$	$B$
$\omega = CBA$	$B$	$C$	$A$

We claim that this full matching is ex-post stable. To see why, start with the informed student  $j_1$ . As  $\mathcal{I}_{j_1}(\omega) = \{\omega\}$ ,  $j_1$  simply conditions potential blocks on the state to determine if they deliver a higher expected utility. In states  $ABC$ ,  $ACB$ ,  $BAC$ , and  $BCA$ ,  $j_1$  gets the best school, and so does not want to block. In state  $CAB$  and  $CBA$ ,  $j_1$  gets the second-best school, but has lower priority at the best school  $C$  than  $j_2$ , who is matched to it. So the stability constraints hold for  $j_1$ .

The stability constraints for the uninformed students are where we diverge from the standard model of stability. Consider  $j_2$  first. It may be tempting to see that  $j_2$  gets  $C$  in every state, so  $\mathcal{A}_{j_2}(C) = \Omega$ , and therefore conclude that  $j_2$  learns nothing and so the stability constraints are satisfied. But one must be careful to also condition on whether potential blocks are possible or not. In particular, suppose  $j_2$  is considering blocking at  $A$ . Then,  $\mathcal{B}_{j_2}(A) = \{BAC, BCA, CBA\}$  as these are the three states where  $j_3$  is matched to  $A$  so  $j_2$  would have higher priority than the student matched to  $A$ . Combining these, we have  $\mathcal{C}_{j_2}(C, A) = \mathcal{B}_{j_2}(A) = \{BAC, BCA, CBA\}$ . Hence we compare  $E[u_{j_2}^\omega(A)|\mathcal{C}_{j_2}(C, A)] = 1/3$  to  $E[u_{j_2}^\omega(C)|\mathcal{C}_{j_2}(C, A)] = 1$  and see that  $j_2$  would prefer to stay at  $C$ . Alternatively,  $j_2$  does not have high enough priority to ever get school  $B$ , and so  $\mathcal{B}_{j_2}(B) = \emptyset$  and  $\mathcal{C}_{j_2}(C, B) = \emptyset$ . Furthermore, as  $j_2$  is never matched to  $A$  or  $B$ , we have that  $\mathcal{A}_{j_2}(A) = \mathcal{A}_{j_2}(B) = \emptyset$  and so  $\mathcal{C}_{j_2}(A, B) = \mathcal{C}_{j_2}(A, C) = \mathcal{C}_{j_2}(B, A) = \mathcal{C}_{j_2}(B, C) = \emptyset$  as well. We stress that it is true that the ex-post stability constraints are satisfied for  $j_2$ , but we must be careful as perhaps  $j_2$ , even though they learn nothing from observing their match, may find that they would only be accepted to a school when it is very good and therefore wish to block with that school.

Moving to  $j_3$ , consider first when  $j_3$  is matched to  $B$ . Here,  $\mathcal{A}_{j_3}(B) = \{ABC, ACB, CAB\}$ . On the other hand,  $j_3$  does not have high enough priority to block at  $A$ , as  $j_3$  has the lowest priority there, or to block with  $C$ , as  $j_2$  is matched to  $C$  in all 6 states. Hence

$\mathcal{B}_{j_3}(A) = \mathcal{B}_{j_3}(C) = \emptyset$  and so  $\mathcal{C}_{j_3}(B, A) = \mathcal{C}_{j_3}(B, C) = \emptyset$  as well. Similarly, when  $j_3$  is matched to  $A$ , we have  $\mathcal{A}_{j_3}(A) = \{BAC, BCA, CBA\}$  but  $\mathcal{B}_{j_3}(B) = \mathcal{B}_{j_3}(C) = \emptyset$  and so  $\mathcal{C}_{j_3}(A, B) = \mathcal{C}_{j_3}(A, C) = \emptyset$  too. Finally,  $\mathcal{A}_{j_3}(C) = \emptyset$  and so  $\mathcal{C}_{j_3}(C, A) = \mathcal{C}_{j_3}(C, B) = \emptyset$ . Thus, the full matching  $\mu$  is ex-post stable.

### 3 The Curse of Acceptance

As stated above, in the standard model where each student is informed about their own preferences, DA is strategy-proof and stable. In this section, we highlight some of the issues these properties present in our model with uninformed students, and how they can be attributed to a so-called curse of acceptance. We begin with stability. Example 2 shows that when there is even one student that is not perfectly informed about her preferences, DA is no longer (ex-post) stable.

**Example 2.** There are 4 students,  $I = \{j_1, j_2, j_3\}$  and  $U = \{j_4\}$ , and 4 schools,  $S = \{A, B, C, D\}$ , each with capacity 1.<sup>19</sup> We continue to work in the common ordinal preference model of Example 1. There are four schools now, so we write, for example,  $\omega = ABCD$  to refer to the state where  $A$  is the most preferred school,  $B$  is the second most preferred school, etc., for all students, and define utility as  $u_j^\omega(s) = 4 - \text{rank}_\omega(s)$ .<sup>20</sup> The priority structure is as follows:

$\succ_A$	$\succ_B$	$\succ_C$	$\succ_D$
$j_1$	$j_1$	$j_3$	$j_2$
$j_2$	$j_3$	$j_1$	$j_1$
$j_3$	$j_2$	$j_4$	$j_4$
$j_4$	$j_4$	$j_2$	$j_3$

After receiving their signals, each student must submit a preference list, and the final matching is computed using the DA algorithm. First, consider the informed students. For these students, playing the truthful strategy  $\sigma_j(\omega) = \omega$  is weakly dominant. This means, for example, in state  $\omega = DABC$ , students  $j_1, j_2$ , and  $j_3$  will all report  $\sigma_j(\omega) = DABC$ .

It has been shown (McVitie and Wilson, 1971; Dubins and Freedman, 1981) that an equivalent way to run DA is to at each step arbitrarily choose one unmatched student and have them apply to their most-preferred school where they have not yet been rejected. The order in which students are chosen to apply does not matter; in particular, we consider

<sup>19</sup>As with many similar impossibility results, this market can easily be embedded in other markets of arbitrary size and preference structure.

<sup>20</sup>Again, any utility values that preserve the common ordinal preference will give the same result.

running DA with only students  $j_1, j_2$ , and  $j_3$  first and then, when they are all tentatively matched, have  $j_4$  apply to her most-preferred school in her report and continue until DA ends.

Now, note that before  $j_4$  makes her first application in this method of DA,  $j_1, j_2$ , and  $j_3$  are tentatively matched to the best three schools in every state. Student  $j_4$  does not have very good priority at any school, and so will usually get the worst school regardless of the preference they have submitted. In particular, she only has a chance to match to a better (than the worst) school if, before she enters the market, either (i)  $j_2$  is matched to  $C$  or (ii)  $j_3$  is matched to  $D$ . Calculating DA on  $\{j_1, j_2, j_3\}$  shows that this only happens in 2 of the 24 states: in state  $\omega' = ABCD$ ,  $j_2$  is matched to  $C$  and in state  $\omega'' = BADC$ ,  $j_3$  is matched to  $D$ . The final matching in these 2 states will depend on what preference relation  $j_4$  reports. If  $j_4$  reports a preference relation  $P_{j_4}$  with  $CP_{j_4}D$  then  $j_4$  will match to  $C$  in *both*  $\omega'$  and  $\omega''$ , whereas if  $j_4$  reports a preference relation  $P'_{j_4}$  with  $DP'_{j_4}C$  then  $j_4$  will match to  $D$  in both  $\omega'$  and  $\omega''$ .<sup>21</sup> The key point to notice is that no matter what she submits,  $j_4$  will receive the same assignment in  $\omega'$  and  $\omega''$ . In one of these states, this assignment will be the third-best school, and in the other, it will be the worst. Combining this with the previous discussion, we see that  $j_4$  will get the worst school in 23 states and the third-best school in 1 state, regardless of what she reports.

We claim that for either report ( $P_{j_4}$  or  $P'_{j_4}$ ), the resulting full matching will not be ex-post stable. Suppose  $j_4$  reports  $CP_{j_4}D$  and is matched to  $C$ . Now, she is matched to  $C$  in 7 states: the 6 states where  $C$  is worst and  $\omega' = ABCD$ . But  $j_3$  is only matched to  $D$  in  $\omega'' = BADC$ , so if she proposes a block with  $D$  it will only be accepted in this state. Her utility from  $D$  in state  $\omega''$  is 1, while her utility from  $C$  is 0, and so she wishes to propose the block.<sup>22</sup> That is, the match is not ex-post stable. If  $j_4$  were instead to report  $P'_{j_4}$  with  $DP'_{j_4}C$ , a symmetric argument shows that she will want to propose blocking with  $C$  when she matches to  $D$ .

In summary, we have shown the following.

**Proposition 1.** *The deferred acceptance mechanism is not ex-post stable.*<sup>23</sup>

<sup>21</sup>The locations of  $A$  and  $B$  in  $j'_4$ 's report are irrelevant as she will just be immediately rejected from either if she applies.

<sup>22</sup>Using our notation,  $\mathcal{A}_{j_4}(C) = \{ABCD, ABDC, ADBC, BADC, BDAC, DABC, DBAC\}$  and  $\mathcal{B}_{j_4}(D) = \{BADC\}$  and so  $\mathcal{C}_{j_4}(C, D) = \{BADC\}$ . Thus,  $1 = E[u_{j_4}^\omega(D) | \mathcal{C}_{j_4}(\mu_{j_4}^\omega, D) \cap \mathcal{I}_{j_4}(\omega)] > E[u_{j_4}^\omega(\mu_{j_4}^\omega) | \mathcal{C}_{j_4}(\mu_{j_4}^\omega, D) \cap \mathcal{I}_{j_4}(\omega)] = 0$ , where  $\mu$  is the full matching induced by DA.

<sup>23</sup>While DA is not ex-post stable, it is also possible to ask whether there exists an ex-post stable mechanism. In the appendix, we show that there is in fact an ex-post stable mechanism we call the *state-learning mechanism*. Intuitively, this mechanism works by learning the state from the reports of the informed students, and uses the information to assign everybody to a stable matching. However, this mechanism, and any other ex-post stable mechanism, requires the mechanism to ignore the stated preferences of the uninformed.

In fact, there are even larger problems with deferred acceptance other than the fact that it is not ex-post stable. In the last example, the uninformed student gets the worst possible school in 23 of the 24 states. Of course, part of the reason for this is that she has relatively low priority at each school. But, as we will show more carefully in later, even when her priority is higher, she will always do worse than average.

These two features, no ex-post stability and generally poor outcomes for the uninformed students, can be attributed to a so-called *curse of acceptance*: uninformed students do not know which schools are the good schools, but upon seeing their assignment, they update their beliefs about the quality of the school they were assigned downwards. The reason is that they know that the more informed students know which schools are the good schools and will end up taking them for themselves. This leaves the less informed students with the low-quality schools, which not only makes them worse off from a welfare perspective, but also leads to ex-post instability.

This still is not the end of the story for how poorly we think DA performs in this environment, as we have yet to consider the incentives for preference-reporting to the mechanism. In the standard model, one of the most desirable features of DA is its strategy-proofness. However, as argued previously, strategy-proofness is no longer a feasible property in our model, and it turns out that without it, determining the equilibrium can be quite complicated. In the previous example, this was not so much of a problem, as in fact every strategy for  $j_4$  yielded the same expected utility and therefore every strategy was an equilibrium. This is just an artifact of the example, however, and usually determining an optimal strategy will be far more complicated and depend very much on the details of the full priority structure. This can make playing the mechanism very difficult for some parents, which is unappealing to many school districts (see Pathak and Sönmez, 2008).

The next example clearly illustrates this point. It actually considers two different problems, which barely differ in their primitives but have vastly different equilibria. Furthermore, the example shows that a student's own priority is not sufficient for determining their equilibrium strategy (i.e., reporting a preference list that orders schools by one's priorities at them is generally *not* an equilibrium strategy for the uninformed students).

**Example 3.** The students, schools, and states are the same as in Example 2. We consider two more examples with slightly different priority structures. The first is as follows:

---

For this and other reasons, which we discuss further in the appendix, we do not believe such mechanisms are relevant in practice.

$\succ_A$	$\succ_B$	$\succ_C$	$\succ_D$
$j_3$	$j_2$	$j_2$	$j_2$
$j_4$	$j_1$	$j_1$	$j_1$
$j_2$	$j_4$	$j_3$	$j_3$
$j_1$	$j_3$	$j_4$	$j_4$

Following the same argument as in the previous example, after running DA on the three informed students,  $j_4$  can only get  $A$  when it is not worst if  $j_2$  or  $j_1$  is matched to  $A$  and can only get  $B$  when it is not worst if  $j_3$  is matched to  $B$ . There are no states where  $j_2$  or  $j_1$  are matched to  $A$ . However,  $j_3$  is matched to  $B$  in both  $CDBA$  and  $DCBA$ . So if  $j_4$  reports a preference  $P_{j_4}$  with  $AP_{j_4}B$ , then  $j_4$  will get matched to the worst school in every state, but if  $j_4$  reports a preference  $P_{j_4}$  with  $BP_{j_4}A$ , then  $j_4$  will get matched to  $B$ , the third best school in states  $CDBA$  and  $DCBA$ . That is, the equilibrium strategies for  $j_4$  are any reports where  $BP_{j_4}A$  (the rankings of  $C$  and  $D$  are irrelevant).

But now let's slightly alter the priority structure by switching the priorities of  $j_1$  and  $j_3$  at school  $D$  (note in particular that the priorities of student  $j_4$  have not changed):

$\succ_A$	$\succ_B$	$\succ_C$	$\succ_D$
$j_3$	$j_2$	$j_2$	$j_2$
$j_4$	$j_1$	$j_1$	$j_3$
$j_2$	$j_4$	$j_3$	$j_1$
$j_1$	$j_3$	$j_4$	$j_4$

With this priority structure,  $j_1$  is matched to  $A$  in states  $CDAB$  and  $BDAC$  and  $j_3$  is matched to  $B$  in  $DCBA$ . So if  $j_4$  reports a preference  $P_{j_4}$  with  $AP_{j_4}B$ , then  $j_4$  will get matched to  $A$ , the third best school, in states  $CDAB$  and  $BDAC$ . If  $j_4$  reports a preference  $P_{j_4}$  with  $BP_{j_4}A$ , then  $j_4$  will get matched to  $B$ , the third best school, in state  $DCBA$ . Therefore, the equilibrium strategies for  $j_4$  are any reports where  $AP_{j_4}B$ . So, just switching the priorities of two students at a school (who both have higher priority than  $j_4$ , and to which  $j_4$  is never matched unless it is the worst school) completely reverses the equilibrium strategies. In one case,  $j_4$  should favor the school at which they have second-highest priority, but in the other case, they should favor the school at which they have third-highest priority. There is no way for  $j_4$  to determine their equilibrium strategy without considering the full priority structure of all students, and we think this challenge is rather unappealing.<sup>24</sup>

---

<sup>24</sup>Also, even though truthful reporting is not feasible, one may argue that DA is fine if there is some dominant strategy equilibrium. Our examples demonstrate that this is clearly not the case. For example, if the informed students reported the same list in every state and put distinct schools at the top of their respective lists, then the uninformed student  $j_4$  could list the remaining school at the top of their list and always be matched to it, which would be better.

To summarize, this section has shown that when preferences are interdependent and students have different levels of information, DA fails to be strategy-proof or ex-post stable. Furthermore, the uninformed students end up with generally poor outcomes on average. One way to proceed would be to look at different mechanisms; however, we do not think this would be a fruitful approach, for two reasons. First, the general logic of the curse of acceptance will hold independently of the mechanism used, and second, DA is the most widely used mechanism in practice, and so it seems important to understand whether we can recover positive results in the context of the DA mechanism. Thus, rather than investigating alternative mechanisms, the next section re-frames the problem from one of mechanism design to one of *priority* design, and argues that school districts can alleviate all of these issues by choosing an appropriate priority structure.

## 4 Secure Schools

The importance of the problems identified in the preceding section clearly depended on the underlying priority structure. In school choice settings, school districts generally have some control over the priority structure, and may be able to design priorities to achieve particular objectives. For example, Dur et al. (2017) and Pathak and Shi (2017) explore consequences of different design decisions relating to the design of walk zones in Boston. These papers still work in the standard framework of perfectly informed students. In this section and the next two, we show that when this assumption is relaxed, the design of priorities can have important consequences for incentives, ex-post stability, and welfare of the uninformed students. In particular, in the following sections, we show that designing priorities such that all students have a so-called “secure school” makes DA ex-post stable, the equilibrium strategies simple, and improves the welfare of the uninformed students by protecting them from the curse of acceptance.

We say that school  $s$  is a **secure school** for student  $j$  if  $|\{j' \in J : j' \succ_s j\}| \leq q_s$ . In words, a secure school is one with enough seats for  $j$  and every student who has higher priority than  $j$ . The rest of this section shows how secure schools allow uninformed students to guard themselves against the curse of acceptance. They allow for a natural and intuitive equilibrium that is ex-post stable and guarantees the uninformed students receive at least an average utility, rather than always being left with the worst schools in every state.

Consider again Example 2. Note that in this example, student  $j_4$  does not have a secure school. What happens if we modify the priority structure to give her a secure school? The priority structure below does so, where we have raised  $j_4$ 's priority at school  $D$ .

$\gamma_A$	$\gamma_B$	$\gamma_C$	$\gamma_D$
$j_1$	$j_1$	$j_3$	$j_4$
$j_2$	$j_3$	$j_1$	$j_2$
$j_3$	$j_2$	$j_4$	$j_1$
$j_4$	$j_4$	$j_2$	$j_3$

Let us reconsider the DA mechanism under these new priorities. Just as before, it remains a dominant strategy for the informed students to submit their true preferences. However, it turns out for the uninformed student  $j_4$ , her optimal strategy is now to rank  $D$ , her secure school, first. Formally, this will follow from Theorem 1 below which applies to more general markets of arbitrary size, but the intuition is easy to see in the small example: by ranking  $D$  first,  $j_4$  receives  $D$  in every state. If she instead ranked  $C$  above  $D$ ,<sup>25</sup> she will not get  $C$  when it is “good”, but will continue to get it when it is “bad”. Essentially, taking her secure school protects her from the acceptance curse and being taken advantage of by the informed students.

Theorem 1 formalizes this intuition to more general markets of arbitrary size. For the formal results in this section, we continue to work in the common ordinal preferences model introduced earlier, where  $|\Omega| = M!$  and each  $\omega \in \Omega$  is chosen with equal probability and corresponds to an ordinal ranking of schools that is common to all students in state  $\omega$ . This can be thought of as an approximation to highly correlated ordinal preferences, which is a common feature in real-world school choice settings. Many papers have used this assumption when trying to extend classical matching models in various ways (e.g., Abdulkadiroğlu et al. (2011) and Troyan (2012)). This assumption is necessary to get a clear result in the discrete model, as the proofs rely on combinatoric arguments that quickly become intractable without it. However, the intuition applies more broadly, and in the next section, we will use a continuum model to analyze the more general case of heterogeneous (but correlated) preferences.

For the priority structure, we assume that each student has at least one secure school, but otherwise, the priority structure is arbitrary.<sup>26</sup> Define a strategy profile  $\sigma_j^*(\omega)$  as follows: for all informed  $j \in I$ ,  $\sigma_j^*(\omega) = P_j(\omega)$ , and for all uninformed  $j \in U$ ,  $\sigma_j^*(\omega) = \tilde{P}_j$  for all  $\omega$ , where  $\tilde{P}_j$  is any preference ranking that lists one of  $j$ 's secure schools first. As discussed above, following  $\sigma_j^*$  is weakly dominant for the informed students; i.e., after observing their signal, they just report their true preferences. Our theorem shows that taking their secure

<sup>25</sup>Note that  $j_4$  still has the lowest priority at  $A$  and  $B$ , and so once again only the relative rankings of  $C$  and  $D$  are relevant.

<sup>26</sup>All that is really necessary is that the uninformed students have a secure school, but, since since the mechanism designer may not know who is informed or uninformed, it is simpler to just require this for all students.

school is also optimal for the uninformed students.

**Theorem 1.** *Assume that every student has a secure school. Then,  $\sigma^*$  is an equilibrium strategy profile.*

The proof of this theorem in the appendix actually shows something slightly stronger, which is that for each  $j$ ,  $j$ 's outcome under profile  $(\sigma_j^*, \sigma_{-j}^*)$  first-order stochastically dominates her outcome under  $(\sigma'_j, \sigma_{-j}^*)$  for any other strategy  $\sigma'_j$  that  $j$  could choose.<sup>27</sup> This implies that the strategy profile  $\sigma^*$  is an equilibrium of the deferred acceptance mechanism (and is also the reason why the cardinal utility values we assigned in the common ordinal preference model could be changed to any other values that preserve the order).

While Theorem 1 constructs one equilibrium, it is a natural equilibrium. When uninformed students do not have a secure school, determining equilibrium strategies for the uninformed students can be quite complicated, as seen in Example 3. Designing priorities such that every student has a secure school results in equilibrium strategies that are focal and very simple to compute: informed students follow the familiar truth-telling strategy, while uninformed students simply take their secure school. Doing so is optimal because it protects them from the curse of acceptance identified previously.

Furthermore, we think that secure schools may be helpful in this environment not only because there exists an equilibrium in simple strategies, but also because the full matching induced by these strategies is ex-post stable. This is stated in the following theorem.

**Theorem 2.** *Suppose every uninformed student has a secure school. Then, the full matching induced by equilibrium  $\sigma^*$  in the deferred acceptance mechanism is ex-post stable.*

What is it about secure schools that make the full matching ex-post stable? The apparent reason is that the uninformed students do not learn anything from observing their own match, because they are always matched to their secure school. However, this is not the complete argument, because they can still condition on their proposed block being implementable (recall the set  $\mathcal{B}_j(s)$  above). Perhaps a student considers a rematching proposal to some school  $s$  that she knows will be accepted in some state where  $s$  is very good. Then, because the informed students all propose to schools in the same order, when  $s$  is worse, the students who match to  $s$  have weakly lower priority than when  $s$  is better (this is where the common ordinal preference assumption is key). That is, if  $s$  will block with  $j$  when it is a good school, it will continue to block with  $j$  whenever it is worse. So, the best that  $j$  can hope for is to

---

<sup>27</sup>A strategy induces a distribution over the number of states where  $j$  gets the best school, the second-best school, the third-best school etc. according to the common ordinal preference. This is the distribution we are referring to in regards to comments on first-order stochastic dominance.

get  $s$  when it's best, second-best, etc. down to the worst, which gives her the same expected utility as her secure school. In fact, unless she gets  $s$  when it is the best, she strictly prefers matching to her secure school.<sup>28</sup>

## 5 Heterogeneous Preferences

While preferences over schools are likely correlated in the real-world, they are almost certainly not perfectly correlated, which raises the question of how our results extend to this setting. The proof strategies used for the discrete economies in the previous section rely on combinatoric arguments that explode and quickly become intractable when preferences are correlated but heterogeneous. However, the intuition for the curse of acceptance still holds when preferences are partially correlated. We can obtain formal results by moving to a continuum economy.

Continuum economies have recently been receiving significant interest in the matching literature because they simplify analysis considerably while still providing useful insights. For example, Miralles (2008) uses a continuum model to compare the Boston mechanism and DA from an ex-ante welfare perspective in settings without priorities on the school side. Abdulkadiroğlu et al. (2015) use a continuum model to understand properties of their choice-augmented DA algorithm that are too complicated to analyze in the standard discrete model (again assuming no priorities on the school side). Azevedo and Leshno (2016) provide a general framework for analyzing stable matchings in economies with a finite number of agents on one side and a continuum on the other. Hassidim et al. (2018) study the relationship between the equilibria of the Boston mechanism and stable matchings in a continuum model allowing for strict priorities and naive and sophisticated students. Besides being more tractable and providing useful insights, continuum economies are particularly well-suited as an approximation in school choice, where market sizes are often very large. For example, the Boston public school district serves over 60,000 students, while in New York City, this number is over one million students.<sup>29</sup>

All of the aforementioned papers that have used continuum economies have worked in the standard private values domain, whereas we want to generalize the finite economy above and allow for interdependent preferences. We do this as follows. There is a finite set of schools, denoted  $S = \{s_1, \dots, s_M\}$ , but each school now has a unit mass of seats to fill. There is a

---

<sup>28</sup>She can only get  $s$  when it is best if  $s$  is also a secure school or at least one other uninformed student has two secure schools, one of which is  $s$ , and chooses one other than  $s$ . So, in most cases,  $j$  will strictly prefer to stay at her secure school.

<sup>29</sup>Chade et al. (2014), Bodoh-Creed and Hickman (2016), and Che and Koh (2016) also use continuum models to study decentralized college admissions problems.

total mass  $M$  of students. The underlying state space  $\Omega$  is again the set of all permutations of  $S$ , and each state  $\omega$  is still identified with an ordinal ranking of the schools. Nature draws a state  $\omega = (s^{(1)}, \dots, s^{(M)})$  uniformly from  $\Omega$ , which can be interpreted as a ranking of the intrinsic qualities of the schools (i.e.,  $s^{(1)}$  is the highest quality school in state  $\omega$ ,  $s^{(2)}$  is the second highest quality, etc.).<sup>30</sup> However, the preferences of any individual student may depart from this ranking (e.g., due to idiosyncratic components such as location, sibling attendance, or other reasons). Formally, we model each state  $\omega$  as determining a measure  $\lambda_\omega$  over  $\mathcal{P}$ , where, for any ordinal ranking  $P$ ,<sup>31</sup>  $\lambda_\omega(P) > 0$  is the measure of students who have ordinal preferences  $P$  in state  $\omega$  and  $\sum_{P \in \mathcal{P}} \lambda_\omega(P) = M$ . A student  $j$ 's cardinal preference for school  $s$ , given that she has ordinal preferences  $P_j$ , is  $u_j^{P_j}(s) = v(\text{rank}_{P_j}(s))$ , where  $v$  is some strictly decreasing function.

We make the following assumptions on  $\lambda_\omega$ . First,  $\lambda_\omega(P) \neq \lambda_\omega(P')$  for all  $P \neq P'$ . Second, let  $\pi$  and  $\tilde{\pi}$  be two permutations of the set  $\{1, \dots, M\}$ . For any two states  $\omega = (s_{\pi(1)}, \dots, s_{\pi(M)})$  and  $\tilde{\omega} = (s_{\tilde{\pi}(1)}, \dots, s_{\tilde{\pi}(M)})$ , and two preference rankings such that  $P : s_{i_1}, s_{i_2}, \dots, s_{i_M}$  and  $\tilde{P} : s_{\tilde{\pi}(\pi^{-1}(i_1))}, s_{\tilde{\pi}(\pi^{-1}(i_2))}, \dots, s_{\tilde{\pi}(\pi^{-1}(i_M))}$ , we have  $\lambda_\omega(P) = \lambda_{\tilde{\omega}}(\tilde{P})$ . Last, say that a preference ranking  $P$  is **closer to state**  $\omega$  than  $P'$  if there exists two schools  $s, s'$  such that  $\text{rank}_\omega(s) < \text{rank}_\omega(s')$ ,  $\text{rank}_P(s) < \text{rank}_{P'}(s)$ , and for all other  $s'' \neq s, s'$ ,  $\text{rank}_P(s'') = \text{rank}_{P'}(s'')$ ; in other words,  $P$  and  $P'$  differ only in where they rank  $s$  and  $s'$ , and  $P$  ranks the two in the same way as  $\omega$ . Then, we assume that if  $P$  is closer to state  $\omega$  than  $P'$  that  $\lambda_\omega(P) > \lambda_\omega(P')$ .

The first condition rules out knife-edge cases where two preference profiles have exactly the same measure of students, which is not a generic property. The second says that the distribution of preferences depends only on the relative rankings of schools, and not on their labels (in particular, if two preference profiles represent the same permutation of two states, then they have the same measure in the respective states). The last condition, our assumption on how preferences are correlated, captures the idea that higher quality schools (as determined by the state) should be more popular in aggregate: if  $P$  is closer to state  $\omega$  than  $P'$ , then there are more students with preferences  $P$  than preferences  $P'$ .<sup>32</sup>

---

<sup>30</sup>We use superscripts to denote rankings (in a manner reminiscent of the common notation for order statistics), while subscripts denote the fixed “names” of the schools.

<sup>31</sup>For simplicity, we abuse notation and drop the subscript  $j$  on preferences  $P$  in this section when we are not referring to a specific student.

<sup>32</sup>One concrete way to model preferences in a finite market that satisfies these assumptions is to assume an underlying probability distribution  $\tilde{\text{Pr}}$  such that  $\tilde{\text{Pr}}(s^{(1)}) > \tilde{\text{Pr}}(s^{(2)}) > \dots > \tilde{\text{Pr}}(s^{(M)})$ , and have student  $j$  draw her preferences  $P_j$  via repeated draws from  $\tilde{\text{Pr}}$ , removing repetitions, until all  $M$  schools have been drawn. This procedure will induce a probability distribution over all preferences  $\mathcal{P}$ , and the measure  $\lambda$  is the analogue of this probability distribution in the continuum model. This procedure for drawing preferences is commonly used in the literature (see, e.g., Immorlica and Mahdian (2005), Kojima and Pathak (2009) and Kojima et al. (2013)).

Moving to priorities, in most realistic school choice applications, schools rarely have strict priority relations, but instead divide students into a few very broad priority classes, which is often referred to as “weak priorities”. Weak priorities necessitate the introduction of lotteries to break ties between students in the same priority class. To capture this feature as well as understand the role of secure schools while still retaining analytical tractability, we assume that each student has exactly one secure school, and for each school  $s$ , there is a mass 1 of students such that  $s$  is their secure school. Each student  $j$  is endowed with priority numbers  $\ell_j(s) \in [0, 1]$ , one for each school  $s$ , and  $j$ ’s priority score at  $s$  is  $\ell_j(s) + 1$  for each school other than  $j$ ’s secure school and  $\ell_j(s)$  for  $j$ ’s secure school (where we use the convention that smaller numbers are better priority). For any  $x \in [0, 1]$ , the measure of students with  $\ell_j(s) \leq x$  is equal to  $x$  for each  $s$ . All students observe their own priority numbers before submitting their preferences.<sup>33</sup>

Of the total mass  $M$  of students, a mass  $\nu$  are informed and the remaining  $M - \nu$  are uninformed, where, as in the previous section, the former receive a perfectly informative signal of their preferences, while the latter receive no signal. A student’s information type, preferences, and priorities are independent; in other words, in each state  $\omega$ , the measure of informed students who have preferences  $P$  and priority numbers  $\ell_j(s) \leq x$  at school  $s$  is  $\nu\lambda_\omega(P)x$  and the measure of respective uninformed students is  $(M - \nu)\lambda_\omega(P)x$ .

Endowing students with preferences and priorities that add up to the measures given above is a way to get around a technical problem with perhaps the more natural way one would want to define the model in a finite world, which is to have each student  $j$  independently draw an ordinal preference  $P_j$  from some distribution over  $\mathcal{P}$  (which may depend on the state, see footnote 32) and a priority score at each school. The technical problem is that when we move to the continuum, realizations of such random draws may result in the set of students with preference  $P$  and priority scores less than  $x$  being non-measurable. The model as formally defined does not suffer from this problem. Similar approaches are taken by Miralles (2008) and Abdulkadirođlu et al. (2015) to deal with this issue.

As before, each student will submit an ordinal preference relation, which will then be turned into a matching using the DA mechanism. Intuitively, the deferred acceptance mechanism works in the same way as in the discrete case: students apply starting at the top of their (submitted) ordinal preference list, and if the total measure of students who apply to a school is greater than the capacity of the school, the school tentatively admits the measure of students equal to its capacity that have the highest priority. Rejected students apply to their next most preferred school, etc. However, a more useful way to understand deferred acceptance in continuum models is given by Abdulkadirođlu et al. (2015), who show that for

---

<sup>33</sup>The equilibrium results below continue to hold whether students observe their priority numbers or not.

any preferences submitted by the students, the DA output can be characterized by a unique vector of *cutoffs*  $(\bar{\ell}_{s_1}, \dots, \bar{\ell}_{s_M})$  (see also Azevedo and Leshno, 2016). Each student  $j$  is then assigned to her most preferred school for which her priority score is lower than the school's cutoff. Any individual student has no ability to affect the cutoffs, and so effectively acts as a “price-taker”. This simplifies the strategic analysis, since, to analyze the equilibrium outcomes of DA in our model, we only need to understand the structure of these cutoffs. The next proposition provides the key property of the cutoffs on which the subsequent equilibrium results are built. As above, let  $\sigma^*$  be a strategy profile where all informed students report their true preferences and uninformed students report their secure school first in their preference list.

**Proposition 2.** *Fix a state  $\omega = (s^{(1)}, \dots, s^{(M)})$ , and suppose the students play  $\sigma^*$ . Then, the deferred acceptance cutoffs satisfy  $1 < \bar{\ell}_{s^{(1)}} < \bar{\ell}_{s^{(2)}} < \dots < \bar{\ell}_{s^{(M)}} = 2$ .*

This proposition gives a characterization of the equilibrium cutoffs under the proposed strategies. Intuitively, the cutoffs are smaller (and so the school is harder to get) for the schools that are more likely to be popular. Indeed, if any of these inequalities were reversed, then there is some school that is both more popular in aggregate and is easier to get into. The proof in the appendix reaches a contradiction by showing that this would result in some school being over capacity. Recall also that priority scores at any school  $s$  range from 0 to 2, with scores in  $[0, 1]$  reserved for students for which  $s$  is their secure school. Since there is only a unit mass of students for which this is true and all schools must be filled to capacity (because the total measure of students is equal to the total capacity of the schools), all cutoffs will be greater than 1, and the highest cutoff will be 2.

We can now use Proposition 2 to find an equilibrium of the preference submission game. For the informed students, truthful reporting continues to be a weakly dominant strategy. What is left to show is that for an uninformed student, listing her secure school  $\bar{s}$  first gives a higher expected utility than any other strategy. The proof of the following theorem is technical, but of all the many cases, the one that matters is intuitive. Consider any other  $s' \neq \bar{s}$ , and two states,  $\omega$  and  $\omega'$  such that  $rank_{\omega}(\bar{s}) < rank_{\omega}(s')$  and  $rank_{\omega'}(s') < rank_{\omega'}(\bar{s})$  (and the ranks of the other schools are identical). If  $j$  submits her secure school at the top of her list, she will get it in both states. If she instead puts  $s'$  above her secure school, then it is possible for her priority number to only be small enough to get it when  $s'$  is less popular (i.e., in state  $\omega$ ) where by Proposition 2, the cutoff to get into  $s'$  is larger than in state  $\omega'$ , giving her the worse school in both states (i.e.,  $s'$  in state  $\omega$  and  $\bar{s}$  in state  $\omega'$ )

**Theorem 3.** *The strategy profile  $\sigma^*$  is an equilibrium of the deferred acceptance mechanism.*

What about ex-post stability? As in the discrete case, there is no learning by uninformed students. However, we still need to address the issue of conditioning on a potential block being accepted. This turns out to be much simpler in this setting than in the discrete case, because by Proposition 2 we can immediately see that if a school will accept a student in some state  $\omega$ , then the cutoffs are larger in every state where the school is less popular, and so will accept the student in all those states as well.

**Theorem 4.** *The strategy  $\sigma^*$  induces an ex-post stable matching and therefore the deferred acceptance mechanism is ex-post stable.*

In summary, the results that there is a natural and simple equilibrium and that this equilibrium produces an ex-post stable matching carry over to the case where preferences are correlated, but not perfectly so. Of course, this makes sense as the general logic of the curse of acceptance does not rely on perfect correlation, but only requires some interdependence of student preferences.

## 6 The Welfare Effects of School Choice

One of the main motivations for introducing school choice is that it allows parents to provide information on how they rank various schools, and this preference information can then be used by the school district to determine a more efficient assignment than the historical practice of simply assigning students to the school closest to their home (i.e. neighborhood assignment). However, this requires the parents to know their preferences over schools in the first place. As we have pointed out, when this is not the case, less informed parents may be exploited by more informed parents and fall prey to the curse of acceptance. Thus, to the extent that school districts are concerned about the welfare of less informed parents, school choice may actually have significant downsides. We close the paper by showing that the use of secure schools can be seen as a practical compromise: it allows choice for those parents who know their preferences, while still protecting the less informed from being exploited.

We can make this point more formally in the context of the continuum model introduced above. In particular, we consider student welfare under three alternative school assignment procedures that we call No Choice (NC), DA with No Secure Schools (DA-NSS), and DA with Secure Schools (DA-SS). DA-SS is simply the procedure introduced above, in which each student is given a secure school, and assignments are determined by running the DA algorithm on the submitted preferences. DA-NSS is the same, except students are not given any secure schools, and so their priority score at every school is just equal to their priority

number.<sup>34</sup> In the No Choice procedure, students are not asked to submit any information about their preferences, and are simply assigned to a school. To be able to make meaningful comparisons, we assume that under the No Choice procedure, students are assigned to whatever was their secure school under DA-SS.

We are interested in comparing the welfare for each group of students (informed and uninformed) under each of the procedures. For any agent  $j$ , let  $V_j^\alpha(\sigma)$  denote agent  $j$ 's interim expected utility (that is, agent  $j$ 's expected utility after she receives her signal about her own preferences, but before the mechanism is run) under school choice procedure  $\alpha$  when agents follow strategy profile  $\sigma$ , where  $\alpha = SS, NSS$ , or  $NC$ . We say that procedure  $\alpha$  (with equilibrium  $\sigma$ ) **Pareto dominates** procedure  $\alpha'$  (with equilibrium  $\sigma'$ ) for some set of students  $\tilde{J} \subseteq J$  if  $V_j^\alpha(\sigma) \geq V_j^{\alpha'}(\sigma')$  for all  $j \in \tilde{J}$ . Further, if this inequality is strict for a subset  $\tilde{\tilde{J}} \subseteq \tilde{J}$  that has strictly positive measure, we say that  $\alpha$  **strictly Pareto dominates**  $\alpha'$  for  $\tilde{J}$ .

The first point we make is to formalize the idea that school choice can have a significant downside for uninformed students. In particular, if there are no secure schools, then the uninformed students fall prey to the curse of acceptance and will actually be worse, in a Pareto sense. Before we can state this formally, we need to analyze equilibrium under DA-NSS. In fact, there turns out to be a very natural equilibrium of the DA-NSS game. Let  $\hat{\sigma}$  be the strategy profile in which informed students report their true preferences and uninformed students report schools in ascending order of their priority numbers. We then have the following theorem.

**Proposition 3.** *The strategy profile  $\hat{\sigma}$  is an equilibrium of the DA-NSS game.*

The outline of the proof of this proposition follows two similar steps to the proof that  $\sigma^*$  was an equilibrium of DA-SS. First, the cutoff result of Proposition 2 still holds because, just like before, an equal number of uninformed students are assigned to each school. Second, given this cutoff structure, it is optimal for the uninformed students to list schools where they have better priority first, because those are the schools they can get when they are more popular and thus mitigate the curse of acceptance the most. While there may be other equilibria, we think that this one is very natural and intuitive (because the uninformed students report the schools they are most likely to get in to at the top of their preferences) and we will use it to establish our welfare results. Our first main result of this section is that the curse of acceptance means that the uninformed students would prefer no choice. In particular, school choice is doing the exact opposite of its intended effect.

---

<sup>34</sup>We note that once preferences have been submitted, the DA-NSS algorithm is equivalent to the DA with Multiple Tie-Breaking algorithm considered in the continuum model of Abdulkadiroğlu et al. (2015).

**Theorem 5.** Consider  $\tilde{J} = U$ , the set of uninformed students. No Choice strictly Pareto dominates DA with No Secure Schools for  $\tilde{J} = U$ .

However, when there are secure schools, then school choice does have the intended effect. In particular, the uninformed students get their secure school in either case, while informed students can use their information to improve their own assignment.

**Theorem 6.** Consider  $\tilde{J} = I \cup U$ , the set of all students (both informed and uninformed). DA with Secure Schools strictly Pareto dominates No Choice for  $\tilde{J} = I \cup U$ .

To summarize, in general, we have seen that uninformed students may be exploited by informed students and end up worse off than if there were no school choice at all due to the curse of acceptance. However, we have also provided school districts worried about these issues with a potential design response to mitigate their effects. Giving each student a secure school can be seen as a practical compromise: it allows choice for those students who know their preferences, which increases efficiency, while protecting uninformed students from the unintended consequences of the curse of acceptance. Overall, the final result is a Pareto-improving allocation.

## 7 Conclusion

This paper has built on the canonical (private values) school choice model by introducing correlation in student preferences and allowing some students to be more informed about their own preferences than others. This will be the case, for example, if each school has an intrinsic quality component that all parents care about (in addition to other heterogeneous factors that may vary across families), but some have more time and resources to research and learn about the qualities of various schools. We show that many of the standard results in the literature fail to hold: the popular deferred acceptance mechanism is no longer strategy-proof or stable, and uninformed students are made worse off. While we focus on DA because it is the mechanism that has gained by far the most traction in practical applications, our results are not specific to deferred acceptance, and can be attributed to a general *curse of acceptance*: under any (reasonable) mechanism, informed students are more likely to list the good schools highly precisely because they know which ones they are, and so uninformed students expect that any school that accepts them will be of low quality. However, we also show that there is the potential to mitigate these effects by casting the problem as one of priority design. By designing priorities such that all students are given a secure school, positive results are recovered: DA is (ex-post) stable, the equilibrium strategies are simple,

and the uninformed students are protected from the worst consequences of the acceptance curse.

Introducing more realistic assumptions about preferences and information motivates the need to think about the school choice problem from the priority design perspective, rather than just the mechanism selection perspective. While our model is stylized, our goal is to point out important issues regarding preference modeling that have not been studied in the school choice literature, and highlight how they underscore the need to think carefully about the interplay between preference information and the design of priorities. In particular, while we have shown that secure schools can be beneficial from a welfare perspective in an abstract model, our results raise important new questions about bringing these ideas to practice. For instance, one possibility is to make everyone’s secure school their neighborhood school, but if less informed and more disadvantaged students are more likely to live closer to lower quality schools, then this may not be very effective. As we have discussed, secure schools can (and perhaps should) be based on factors other than geography, with the priority design used in Boston serving as one example. Additionally, while we defined a secure school as a school at which a student can ensure a seat with probability 1 if she so desires, the definition need not be binary. In other words, distributing priorities in such a way that students have a relatively high, though not necessarily guaranteeable, priority number at some school will still mitigate, though perhaps not completely eliminate, the curse of acceptance. In short, the optimal way to distribute priorities is a crucial design question that is likely to gain increasing attention in the future. In practice, the ultimate choice will depend on the details of the particular application at hand, but we hope that theoretical models such as ours provide useful guidance. As have shown in this paper, informational asymmetries among parents should be a key factor in this debate, since they are likely to have important consequences for student welfare.

## References

- ABDULKADIROĞLU, A., Y.-K. CHE, P. A. PATHAK, A. E. ROTH, AND O. TER-CIEUX (2017): “Minimizing Justified Envy in School Choice: The Design of New Orleans’ OneApp,” Tech. rep., National Bureau of Economic Research.
- ABDULKADIROĞLU, A., Y.-K. CHE, AND Y. YASUDA (2015): “Expanding “choice” in school choice,” *American Economic Journal: Microeconomics*, 7, 1–42.
- ABDULKADIROĞLU, A. AND T. SÖNMEZ (2003): “School choice: A mechanism design approach,” *American economic review*, 729–747.

- ABDULKADIROĞLU, A., Y.-K. CHE, AND Y. YASUDA (2011): “Resolving Conflicting Preferences in School Choice: The “Boston” Mechanism Reconsidered,” *American Economic Review*, 101, 399–410.
- ABDULKADIROĞLU, A., P. A. PATHAK, AND A. E. ROTH (2005a): “The New York City high school match,” *American Economic Review, Papers and Proceedings*, 95, 364–367.
- ABDULKADIROĞLU, A., P. A. PATHAK, A. E. ROTH, AND T. SÖNMEZ (2005b): “The Boston Public School Match,” *American Economic Review, Papers and Proceedings*, 95, 364–367.
- AFACAN, M. O. (2012): “Group robust stability in matching markets,” *Games and Economic Behavior*, 74, 394–398.
- AZEVEDO, E. M. AND J. D. LESHNO (2016): “A supply and demand framework for two-sided matching markets,” *Journal of Political Economy*, 124, 1235–1268.
- BODOH-CREED, A. AND B. R. HICKMAN (2016): “College assignment as a large contest,” .
- CHADE, H. (2006): “Matching with noise and the acceptance curse,” *Journal of Economic Theory*, 129, 81–113.
- CHADE, H., G. LEWIS, AND L. SMITH (2014): “Student portfolios and the college admissions problem,” *Review of Economic Studies*, 81, 971–1002.
- CHAKRABORTY, A., A. CITANNA, AND M. OSTROVSKY (2010): “Two-sided matching with interdependent values,” *Journal of Economic Theory*, 145, 85–105.
- (2015): “Group stability in matching with interdependent values,” *Review of Economic Design*, 19, 3–24.
- CHE, Y.-K., J. KIM, AND F. KOJIMA (2015): “Efficient assignment with interdependent values,” *Journal of Economic Theory*, 158, 54–86.
- CHE, Y.-K. AND Y. KOH (2016): “Decentralized college admissions,” *Journal of Political Economy*, 124, 1295–1338.
- DUBINS, L. E. AND D. A. FREEDMAN (1981): “Machiavelli and the Gale-Shapley algorithm,” *American Mathematical Monthly*, 88, 485–494.
- DUR, U. M., S. D. KOMINERS, P. A. PATHAK, AND T. SÖNMEZ (2017): “The demise of walk zones in Boston: Priorities vs. precedence in school choice,” *Journal of Political Economy*, forthcoming.
- ERGIN, H. (2002): “Efficient Resource Allocation on the Basis of Priorities,” *Econometrica*, 70, 2489–2498.
- GALE, D. AND L. S. SHAPLEY (1962): “College Admissions and the Stability of Marriage,”

- The American Mathematical Monthly*, 69, 9–15.
- HASSIDIM, A., A. ROMM, AND R. I. SHORRER (2018): “A Continuum Model for Games of School Choice,” working paper.
- IMMORLICA, N. AND M. MAHDIAN (2005): “Marriage, Honesty, and Stability,” in *Proceedings of the 16th ACM-SIAM Symposium on Discrete Algorithms*, 53–62.
- KOJIMA, F. (2011): “Robust stability in matching markets,” *Theoretical Economics*, 6, 257–267.
- KOJIMA, F. AND P. A. PATHAK (2009): “Incentives and Stability in Large Two-Sided Matching Markets,” *American Economic Review*, 99, 608–627.
- KOJIMA, F., P. A. PATHAK, AND A. E. ROTH (2013): “Matching with Couples: Stability and Incentives,” *Quarterly Journal of Economics*, 128, 1585–1632.
- LEE, S.-H. (2009): “Jumping the curse: Early contracting with private information in university admissions,” *International Economic Review*, 50, 1–38.
- MCEACHIN, A., B. STECHER, AND G. EVANS (2015): “Not Everyone Has a Choice,” <https://www.usnews.com/opinion/knowledge-bank/2015/08/31/how-school-choice-could-disadvantage-low-income-students>, (Accessed April 18, 2018).
- MCVITIE, D. G. AND L. B. WILSON (1971): “The stable marriage problem,” *Communications of the ACM*, 14, 486–490.
- MIRALLES, A. (2008): “School Choice: The Case for the Boston Mechanism,” Boston University, unpublished mimeo.
- NATIONAL CENTER FOR EDUCATION STATISTICS (2015): “Digest of Education Statistics,” Tech. rep., United States Department of Education.
- PATHAK, P. A. (2016): “What Really Matters in Designing School Choice Mechanisms,” in *Advances in Economics and Econometrics, 11th World Congress of the Econometric Society*, ed. by L. Samuelson, Cambridge University Press.
- PATHAK, P. A. AND P. SHI (2017): “How Well Do Structural Demand Models Work? Counterfactual Predictions in School Choice,” Tech. rep., National Bureau of Economic Research.
- PATHAK, P. A. AND T. SÖNMEZ (2008): “Leveling the playing field: Sincere and sophisticated players in the Boston mechanism,” *The American Economic Review*, 98, 1636–1652.
- (2013): “School admissions reform in Chicago and England: Comparing mechanisms by their vulnerability to manipulation,” *The American Economic Review*, 103, 80–106.
- SHI, P. (2015): “Guiding school-choice reform through novel applications of operations research,” *Interfaces*, 45, 117–132.

TROYAN, P. (2012): “Comparing School Choice Mechanisms by Interim and Ex-Ante Welfare,” *Games and Economic Behavior*, 75, 936–947.

## A Proofs

We will introduce additional notation as needed below, and then sometimes use that notation again in later proofs. So if notation seems unfamiliar from the main text, refer to previous proofs in this appendix for definitions.

### A.1 Proof of Theorem 1

Recall that a state  $\omega \in \Omega$  can be identified with an ordinal ranking of the schools, and we write a generic state as

$$\omega = (s^{(1)}, s^{(2)}, \dots, s^{(M)})$$

where  $s^{(k)}$  denotes the  $k^{\text{th}}$  best school in state  $\omega$ . Let  $\mu$  denote the full matching induced by the deferred acceptance mechanism.

**Lemma 1.** *Let  $j$  be an uninformed student. Suppose all students  $j' \neq j$  follow the strategy  $\sigma_{j'}^*(\cdot)$ , and let  $j$  choose any arbitrary strategy. Consider a state  $\omega = (s^{(1)}, \dots, s^{(M)})$  such that under this strategy profile, we have  $\mu_j^\omega = s^{(k)}$ . Then, in any other state  $\omega' = (s'^{(1)}, \dots, s'^{(M)})$  where  $s^{(m)} = s'^{(m)}$  for each  $m < k$ , we have  $\mu_j^{\omega'} \notin \{s^{(1)}, \dots, s^{(k-1)}, s'^{(k)}\}$ .*

*Proof.* Given the proposed strategies, all of the uninformed students other than  $j$  apply to a secure school in the first round and are admitted. Thus, it is without loss of generality to consider the submarket that removes these students and their seats. This means that we can consider a market with only one uninformed student, and the rest of the students are informed.

Consider DA for this submarket in state  $\omega$ . First, we show that in all steps  $m < k$ , the informed students who haven't matched in a previous step, apply to  $s^{(m)}$  and the top  $q_{s_m}$  priority of them are admitted (where we abuse notation and let  $q_{s_m}$  denote the capacity of  $s^{(m)}$  in the submarket). Suppose for contradiction that this is not true. By construction of the informed students' preferences, this is only possible if there is some step  $m$  where the uninformed student  $j$  is matched to  $s^{(m)}$  at the end of the step. Since  $j$  is ultimately matched to  $s^{(k)}$  there is some step  $m' > m$  where  $j$  is rejected from  $s^{(m)}$ . But there are only informed students applying in steps  $m+1, \dots, m'$  and they have all been rejected from schools  $s^{(1)}, \dots, s^{(m)}$  so this contradicts that any student will apply to  $s^{(m)}$  in step  $m'$ . This

also means that all the informed students that are tentatively assigned in steps  $m < k$  to  $s^{(m)}$  are matched to  $s^{(m)}$  at the end of the algorithm (as otherwise it could only be  $j$ 's application that rejects them, but then  $j$  would not subsequently get rejected following the same argument just given).

Now, let  $I_k$  denote all informed students who are not assigned to  $s^{(m)}$  for some  $m < k$ . These  $I_k$  students apply to  $s^{(k)}$  in step  $k$  and the top  $q_{s_k} - 1$  priority of them are matched to  $s^{(k)}$  because  $j$  ultimately gets matched to  $s^{(k)}$ , and given the previous paragraph, none of the informed students in  $I \setminus I_k$  ever apply to  $s^{(k)}$ .

Now, consider state  $\omega'$ . Given the first paragraph and that informed students rank the first  $k - 1$  schools identically, steps  $m < k$  follow identically as those steps when the state is  $\omega$ . So the same  $I_k$  informed students are left that could possibly match with  $s^{(k)}$ . As  $j$  matches to  $s^{(k)}$  in state  $\omega$ , we know that  $j$  has lower priority than at most  $q_{s_k} - 1$  which means that if  $j$  applies to  $s^{(k)}$ , then  $j$  will match to  $s^{(k)}$ . Therefore,  $j$  will never match to a school  $j$  ranks lower than  $s^{(k)}$ .

So what remains to show is that, in state  $\omega'$ , if  $j$  ranks  $s^{(m)}$  for some  $m < k$  or  $s'^{(k)}$  better than  $s^{(k)}$ , then  $j$  is rejected. As the same informed students are matched to  $s^{(m)}$  for each  $m < k$  in both states,  $j$  must be rejected from each  $s^{(m)}$  or else they would have matched to it in  $\omega$  as well. For  $s'^{(k)}$ , the top  $q_{s'_k}$  students from  $I_k$  tentatively match to  $s'^{(k)}$ . Suppose for contradiction that  $j$  matched to  $s'^{(k)}$  in state  $\omega'$ . Then  $j$  has higher priority than one of these students. But the students who get matched to  $s'^{(k)}$  in state  $\omega$  is a subset of  $I_k$  so  $j$  is not rejected from  $s'^{(k)}$  in  $\omega$  contradicting that  $j$  is matched to the school they rank worse  $s^{(k)}$ .

□

Fixing the strategies of the other agents at  $\sigma_j^*$ , for any strategy  $\sigma_j$  for the uninformed student  $j$ , let

$$F(k|\sigma_j) = \Pr(j \text{ matches to a school ranked } k^{\text{th}} \text{ or better} | \sigma_j)$$

denote the rank distribution of  $j$ 's assignment when she uses strategy  $\sigma_j$ . We now prove a Theorem from which Theorem 1 is an immediate consequence. Note that  $F(k|\sigma_j^*)$  first-order stochastically dominating  $F(k|\sigma_j)$  immediately implies that  $\sigma_j^*$  is optimal for student  $j$ .

**Theorem.** *For any strategy  $\sigma_j$ ,  $F(k|\sigma_j^*)$  first-order stochastically dominates  $F(k|\sigma_j)$ .*

Before the proof, note that Theorem 1 follows immediately given that first-order stochastic dominance implies that expected utility (for any utility function that is decreasing in rank) is larger from  $\sigma_j^*$  than  $\sigma_j$ .

*Proof.* Suppose not, i.e., there exists some  $\sigma_j$  and some  $k$  such that  $F(k|\sigma_j) > F(k|\sigma_j^*)$ . Note that if  $j$  follows strategy  $\sigma_j^*$ , she gets the same school in every state, call it  $\bar{s}$ . There are  $(M-1)!$  states where  $\bar{s}$  is the best school,  $(M-1)!$  states where it is the second-best school, etc., and so it is easy to see that  $F(k|\sigma_j^*) = (M-1)!k$ . Consider the minimum  $k$  such that  $F(k|\sigma_j) > F(k|\sigma_j^*)$ . For this  $k$ , it must be that  $j$  gets the  $k^{\text{th}}$  or better ranked school in strictly more than  $k(M-1)!$  states, and, because  $k$  is the smallest index for which this is true,  $j$  must get exactly the  $k^{\text{th}}$  ranked school in strictly more than  $(M-1)!$  states.

Partition the state space into groups such that each group contains all of the states for which the best  $k$  schools are the same; in other words, two states  $\omega = (s^{(1)}, \dots, s^{(M)})$  and  $\tilde{\omega} = (\tilde{s}^{(1)}, \dots, \tilde{s}^{(M)})$  belong to the same group  $G$  if and only if  $s^{(\ell)} = \tilde{s}^{(\ell)}$  for all  $\ell = 1, \dots, k$ . There are  $M!/(M-k)!$  groups, and each group contains  $(M-k)!$  states. Note that each group  $G$  can be uniquely identified by listing its top  $k$  schools in order,  $(s^{(1)}, \dots, s^{(k)})$ .

Suppose that in state  $\omega = (s^{(1)}, \dots, s^{(k)}, \dots, s^{(M)})$ , student  $j$  matches to school  $s^{(k)}$ . Let  $G_{\tilde{s}}$  be the group such that  $\tilde{s}^{(1)} = s^{(1)}, \dots, \tilde{s}^{(k-1)} = s^{(k-1)}$ , but  $\tilde{s}^{(k)} = \tilde{s} \neq s^{(k)}$ ; in other words,  $G_{\tilde{s}}$  is the group that has the same first  $(k-1)$  best schools as  $G$ , but replaces  $s^{(k)}$  with  $\tilde{s}$ . There are  $M-k$  possible choices for  $\tilde{s}$ , and hence,  $M-k$  such groups  $G_{\tilde{s}}$ . Let  $\tilde{S} = S \setminus \{s^{(1)}, \dots, s^{(k)}\}$ , and define  $\tilde{G} = \cup_{\tilde{s} \in \tilde{S}} G_{\tilde{s}}$ . Lemma 1 implies that for all  $\omega' \in \tilde{G}$ , student  $j$  ends up with worse than the  $k^{\text{th}}$  ranked school. Note that  $|\tilde{G}| = (M-k) \times (M-k)!$ .

By our hypothesis,  $j$  gets the  $k^{\text{th}}$  best school in strictly more than  $(M-1)!$  states. Every group  $G$  contains  $(M-k)!$  different states, which implies that  $j$  must get the  $k^{\text{th}}$  best school in at least  $(M-1)!/(M-k)!$  different groups. But, by the previous paragraph, for each of these groups  $G$ , there is an associated  $\tilde{G}$  such that  $j$  gets strictly worse than the  $k^{\text{th}}$  best school for all  $\omega' \in \tilde{G}$ . Since  $|\tilde{G}| = (M-k) \times (M-k)!$  and there must be at least  $(M-1)!/(M-k)!$  such  $\tilde{G}$ 's, that means that there are at least

$$(M-k) \times (M-k)! \times \frac{(M-1)!}{(M-k)!} = (M-k) \times (M-1)!$$

states where  $j$  gets worse than the  $k^{\text{th}}$  ranked school. Since there are  $M!$  total states, this leaves at most

$$M! - (M-k) \times (M-1)! = k \times (M-1)!$$

states where  $j$  can get the  $k^{\text{th}}$  ranked or better school. However, this contradicts that  $j$  gets the  $k^{\text{th}}$  ranked or better school in strictly more than  $k \times (M-1)!$  states.  $\square$

## A.2 Proof of Theorem 2

Choose some uninformed student  $j$ . As with the proof of Theorem 1, it is without loss of generality to consider a submarket that has removed all of the other uninformed students  $j' \in U \setminus \{j\}$  together with the seats they take at their secure schools in equilibrium. We consider the set of students is  $I \cup \{j\}$ .

We start with the following lemma.

**Lemma 2.** *Assume that  $j \succ_{s'} j'$  for some  $j'$  assigned to school  $s'$  in state  $\omega = (s^{(1)}, \dots, s^{(k-1)}, s', s^{(k+1)}, \dots, s^{(M)})$ . Then, for any state  $\tilde{\omega} = (\tilde{s}^{(1)}, \dots, \tilde{s}^{(M)})$  such that  $\tilde{s}^{(m)} = s^{(m)}$  for all  $m < k$ , we have  $j \succ_{s'} j''$  for some  $j''$  assigned to school  $s'$  in state  $\tilde{\omega}$ .*

In other words, this lemma says that if student  $j$ 's proposed rematching is accepted by school  $s'$  in some state  $\omega$ , it will *also* be accepted in any state  $\tilde{\omega}$  where  $s'$  is ranked lower; i.e., if  $s'$  accepts  $j$  when it is “good”, it will also accept  $j$  when it is “bad”.

*Proof of lemma.* Consider the DA algorithm in state  $\omega$ . Since all the uninformed students take their secure school, we only need to consider the informed students. Thus, at each step  $m$  of the algorithm, there is a set of unmatched students  $I_m$ , and all of these students apply to school  $s^{(m)}$ . The  $q_{s_m}$ -highest priority students are thus admitted (again, abusing notation to let  $q_{s_m}$  be the capacity left after the uninformed students whose secure school is  $s_m$  are matched to it), and, since no new students apply to  $s^{(m)}$  at any later step of the algorithm, these students are the ones who will ultimately be assigned to school  $s^{(m)}$ . In state  $\omega$ ,  $s^{(k)} = s'$ , and so the  $q_{s'}$ -highest priority students from  $I_k$  are admitted to  $s'$ . By assumption, in this set there is some  $j'$  such that  $j \succ_{s'} j'$ .

Now, consider  $\tilde{\omega}$ . By construction, in state  $\tilde{\omega}$ , school  $s'$  is ranked weakly worse than  $k^{\text{th}}$ :  $s' = \tilde{s}^{(k')}$  for some  $k' \geq k$ . Since the schools ranked  $m < k$  are the same as in state  $\omega$ , the set of students who apply to  $s'$  in state  $\tilde{\omega}$  is a subset of  $I_k$ . This means that the lowest-ranked student admitted to  $s'$  in state  $\tilde{\omega}$ ,  $j''$ , is ranked (weakly) worse than  $j'$  according to  $\succ_{s'}$ , and so  $j \succ_{s'} j' \succ_{s'} j''$ .  $\square$

Continuing with the main proof, let  $\bar{s}$  be the secure school that student  $j$  lists first in equilibrium which  $j$  she receives in every state  $\omega$ . After the matching is implemented, consider  $j$  proposing to form a blocking pair with some other school  $s'$ . Suppose there is a state  $\omega$  where  $j$  has higher priority than some  $j'$  that is assigned to  $s'$  and would prefer  $s'$  to  $\bar{s}$ . In particular, let the ranking of  $s'$  in state  $\omega$  be  $k$ , and the ranking of  $\bar{s}$  be  $\ell > k$ :

$$\omega = (s^{(1)}, \dots, s^{(k-1)}, s', s^{(k+1)}, \dots, s^{(\ell-1)}, \bar{s}, s^{(\ell+1)}, \dots, s^{(M)}).$$

By the lemma,  $j$  also has priority over some  $j''$  assigned to  $s'$  in the state

$$\tilde{\omega} = (s^{(1)}, \dots, s^{(k-1)}, \bar{s}, s^{(k+1)}, \dots, s^{(\ell-1)}, s', s^{(\ell+1)}, \dots, s^{(M)}).$$

So, for every state where  $s'$  accepts  $j$  as the  $k^{\text{th}}$  ranked school and  $\bar{s}$  is the  $\ell^{\text{th}}$  ranked school for  $\ell > k$ , there is a symmetric state where  $s'$  accepts  $j$  as the  $\ell^{\text{th}}$  ranked school, while had  $j$  stuck with  $\bar{s}$ , she would have received the  $k^{\text{th}}$  ranked school. Since each of these states are equally likely, student  $j$  is not better off conditional on being admitted to  $s'$  for these two states. Summing over all state where  $j$  can block with  $s'$  and  $s'$  is better than  $\bar{s}$  and the companion states where the rankings are reversed, the total expected utility of rematching with  $s'$  is less than or equal to staying at  $\bar{s}$ .  $\square$

(Side note: It is tempting to look at this proof and conclude that every school gives the same expected payoff (conditional on a block being accepted), which is obviously not true. The reason is that the above logic does not apply in reverse: that is, if we start with the fact that  $j$ 's block with  $s'$  is accepted in state  $\tilde{\omega}$ , we *cannot* conclude that  $j$ 's block will be accepted in state  $\omega$ . But  $j$  is worse off when making the switch in state  $\tilde{\omega}$ , and if she is not accepted in state  $\omega$  to offset this “loss”, she will be worse off overall.)

### A.3 Proof of Proposition 2

Fix  $\omega$ . This determines the measure of students of each type who report each ordinal preference relation (following the given strategy  $\sigma^*$ ). As shown by Abdulkadiroğlu et al. (2015), there is a unique vector of cutoffs  $(\bar{\ell}_{s_1}, \dots, \bar{\ell}_{s_M})$  that can be used to determine the DA assignment. Consider two schools  $s_A$  and  $s_B$  and assume WLOG that  $\text{rank}_\omega(s_A) < \text{rank}_\omega(s_B)$ . For any ordinal preference ranking  $P \in \mathcal{P}$ , write  $P(r)$  for the  $r^{\text{th}}$  ranked school according to  $P$ . Partition the ordinal preference space  $\mathcal{P}$  into  $\mathcal{P} = \mathcal{P}_A \cup \mathcal{P}_B$ , where  $P \in \mathcal{P}_A$  if and only if  $s_A P s_B$  (and  $P \in \mathcal{P}_B$  if and only if  $s_B P s_A$ ). For every  $P \in \mathcal{P}_A$ , there is a corresponding  $P' \in \mathcal{P}_B$  with  $P(r) = P'(t) = s_A$  and  $P(t) = P'(r) = s_B$  for some  $r < t$ , and  $P(k) = P'(k)$  for all  $k \neq r, t$ . In words,  $P$  and  $P'$  are exactly the same, except that the ranking of schools  $s_A$  and  $s_B$  are swapped. Consider some such  $P \in \mathcal{P}_A$  and corresponding  $P' \in \mathcal{P}_B$ . By our assumption on the preference measure,  $\lambda_\omega(P) > \lambda_\omega(P') > 0$ .

We want to show that  $\bar{\ell}_{s_A} < \bar{\ell}_{s_B}$ . Towards a contradiction, suppose that  $\bar{\ell}_{s_A} \geq \bar{\ell}_{s_B}$ . Each uninformed student is matched to their secure school so there is an equal measure  $\frac{1}{M}(M - \nu)\lambda_\omega(P)$  of uninformed students with preferences  $P$  matched to each of  $s_A$  and  $s_B$  (because a fraction  $1/M$  have each school as their secure school).

For the informed students, it will be helpful to divide them into three distinct classes: (1) students whose secure school is strictly preferred to  $s_A$ ; (2) students whose secure school is  $s_A$ ; and (3) students whose secure school is strictly worse than  $s_A$  (where the rankings are

according to  $P$ ). The measure of students in class (1) is  $\frac{r-1}{M}\nu\lambda_\omega(P)$ , the measure of students in class (2) is  $\frac{1}{M}\nu\lambda_\omega(P)$ , and the measure of students in class (3) is  $\frac{M-r}{M}\nu\lambda_\omega(P)$  (recall that  $\text{rank}_P(s_A) = r$  so there are  $r - 1$  better schools and  $M - r$  worse schools than  $s_A$ ).

Now, no student in class (1) is matched to  $s_A$ , because  $\bar{\ell}_{s_A} \geq 1$  and so they will get into a school they rank at least as well as  $s_A$  (in fact,  $\bar{\ell}_s \geq 1$  for every  $s$  as otherwise it could not be full to capacity, and every school is full to capacity because there is an equal measure of seats and students). Students in class (2) match to  $s_A$  if and only if their priority scores at the  $r - 1$  schools they prefer to  $s_A$  are higher than the cutoffs at (all of) these schools. Thus, the fraction of such students matched to  $s_A$  is  $\prod_{x=1}^{r-1}(2 - \bar{\ell}_{P(x)})$  (we can assume that  $\bar{\ell}_s \leq 2$  for each school as 2 is the largest possible priority score). Students in class (3) are similar, except that they also must have a low enough priority score at  $s_A$ , and so the total fraction of students in class (3) matched to  $s_A$  is  $\prod_{x=1}^{r-1}(2 - \bar{\ell}_{P(x)}) \times (\bar{\ell}_{s_A} - 1)$ . Combining all of this, the total measure of students with preference  $P$  matched to school  $s_A$  in state  $\omega$  is

$$\frac{\lambda_\omega(P)}{M} \times \left[ \nu \times \prod_{x=1}^{r-1}(2 - \bar{\ell}_{P(x)}) \times (1 + (M - r)(\bar{\ell}_{s_A} - 1)) + (M - \nu) \right] \quad (1)$$

We can do an equivalent analysis for the measure of students with preference  $P'$  who get matched to  $s_A$  (recalling that under  $P'$ , school  $s_A$  is ranked  $t^{\text{th}}$ ):

$$\frac{\lambda_\omega(P')}{M} \times \left[ \nu \times \prod_{x=1}^{t-1}(2 - \bar{\ell}_{P'(x)}) \times (1 + (M - t)(\bar{\ell}_{s_A} - 1)) + (M - \nu) \right] \quad (2)$$

Now, recall that  $P(k) = P'(k)$  for all  $k < t$ , with the exception of  $k = r$  where  $P(r) = s_B$ . In particular, we can re-write equation 2 as

$$\frac{\lambda_\omega(P')}{M} \times \left[ \nu \times \prod_{x=1, x \neq r}^{t-1} (2 - \bar{\ell}_{P(x)})(2 - \bar{\ell}_{s_B}) \times (1 + (M - t)(\bar{\ell}_{s_A} - 1)) + (M - \nu) \right]$$

The total measure of students assigned to  $s_A$  that are of preference type either  $P$  or  $P'$

$$\delta(P, s_A)\lambda_\omega(P) + \delta(P', s_A)\lambda_\omega(P')$$

where, to simplify the notation, we define  $\delta(P, s_A)$  and  $\delta(P', s_A)$  as

$$\begin{aligned}\delta(P, s_A) &= \frac{1}{M} \times \left[ \nu \times \prod_{x=1}^{r-1} (2 - \bar{\ell}_{P(x)}) \times (1 + (M - r)(\bar{\ell}_{s_A} - 1)) + (M - \nu) \right] \\ \delta(P', s_A) &= \frac{1}{M} \times \left[ \nu \times \prod_{x=1, x \neq r}^{t-1} (2 - \bar{\ell}_{P(x)})(2 - \bar{\ell}_{s_B}) \times (1 + (M - t)(\bar{\ell}_{s_A} - 1)) + (M - \nu) \right]\end{aligned}$$

We can do the same analysis for school  $s_B$ . By symmetry, the expressions are the same as above, except  $P$  is swapped with  $P'$  and  $s_A$  is swapped with  $s_B$ . The total measure of students matched to  $s_B$  is

$$\delta(P, s_B)\lambda_\omega(P) + \delta(P', s_B)\lambda_\omega(P')$$

where the  $\delta$ 's in this case are defined as

$$\begin{aligned}\delta(P, s_B) &= \frac{1}{M} \times \left[ \nu \times \prod_{x=1, x \neq r}^{t-1} (2 - \bar{\ell}_{P(x)})(2 - \bar{\ell}_{s_A}) \times (1 + (M - t)(\bar{\ell}_{s_B} - 1)) + (M - \nu) \right] \\ \delta(P', s_B) &= \frac{1}{M} \times \left[ \nu \times \prod_{x=1}^{r-1} (2 - \bar{\ell}_{P(x)}) \times (1 + (M - r)(\bar{\ell}_{s_B} - 1)) + (M - \nu) \right]\end{aligned}$$

We are now interested in comparing the  $\delta$ 's. In particular, recall our contradiction hypothesis that  $\bar{\ell}_{s_A} \geq \bar{\ell}_{s_B}$ . This immediately implies that  $\delta(P, s_A) \geq \delta(P', s_B)$  and  $\delta(P', s_A) \geq \delta(P, s_B)$ . Adding these two equations and re-arranging gives

$$\delta(P, s_A) - \delta(P, s_B) \geq \delta(P', s_B) - \delta(P', s_A) \quad (3)$$

Note also that  $\delta(P, s_A) - \delta(P, s_B) \geq 0$  with equality if and only if  $\bar{\ell}_{P(x)} = 2$  for some  $x < r$  (to see this, recall also that  $r < t$ ). Further, recall that  $\lambda_\omega(P) > \lambda_\omega(P')$  by our preference assumption, so we can multiply equation 3 and obtain the inequality

$$[\delta(P, s_A) - \delta(P, s_B)] \lambda_\omega(P) \geq [\delta(P', s_B) - \delta(P', s_A)] \lambda_\omega(P') \quad (4)$$

This re-arranges to

$$\delta(P, s_A)\lambda_\omega(P) + \delta(P', s_A)\lambda_\omega(P') \geq \delta(P', s_B)\lambda_\omega(P') + \delta(P, s_B)\lambda_\omega(P) \quad (5)$$

Note what this says: among those students whose ordinal preference types  $P_j \in \{P, P'\}$ , a weakly greater measure are matched to  $s_A$  than to  $s_B$ . Furthermore, the inequality is an

equality if and only if  $\bar{\ell}_{P(x)} = 2$  for some  $x < r$  in which case no students are matched to either  $s_A$  or  $s_B$ . As  $s_A$  and  $s_B$  must be filled to capacity, the inequality must be strict for at least one pair  $\{P, P'\}$  in the partition. So if we sum over all of the inequalities for every corresponding pair in the partition, we conclude that in state  $\omega$ , the total measure of students matched to  $s_A$  is strictly greater than the total measure of students matched to  $s_B$ . However, this contradicts that the schools have the same measure of informed students in all states. As  $s_A$  and  $s_B$  were arbitrary schools such that  $\text{rank}_\omega(s_A) < \text{rank}_\omega(s_B)$ , we have established that  $\bar{\ell}_{s(1)} < \bar{\ell}_{s(2)} < \dots < \bar{\ell}_{s(M)}$ .

Now, we show that  $\bar{\ell}_{s(1)} > 1$ . Suppose for contradiction that  $\bar{\ell}_{s(1)} = 1$  (it cannot be less than 1, as argued above). Then, no student for which  $s^{(1)}$  is not their secure school will get matched to  $s^{(1)}$ . As there is an equal measure of seats and students all schools must be filled to capacity so every student whose secure school is  $s^{(1)}$  must be matched to  $s^{(1)}$ . Let  $s \neq s^{(1)}$  be any other school. As there are a finite number of schools, information types, and ordinal preferences, for every  $\epsilon > 0$  there is a set of informed students  $\tilde{J} \subset I$  of strictly positive measure such that for all  $j \in \tilde{J}$ : (i)  $j$ 's secure school is  $s^{(1)}$ , (ii)  $s P s^{(1)}$ , and (iii)  $\ell_s(j) \leq \epsilon$ . As these students are not matched to  $s$ , this means that  $\bar{\ell}_s < 1 + \epsilon$ . As  $\epsilon$  can be arbitrarily small, this implies that  $\bar{\ell}_s = 1$ . But, this means that  $\bar{\ell}_{s(1)} = \bar{\ell}_s$ , which contradicts that all cutoffs are distinct, as we just showed in the previous paragraph.

Finally, we show that  $\bar{\ell}_{s(M)} = 2$ . Suppose for contradiction that  $\bar{\ell}_{s(M)} < 2$ . By the above inequalities, we have  $\bar{\ell}_s < 2$  for every school  $s$ . This means that every school rejects students at some step of DA. Let step  $k$  be the first step where the last school rejects students. Once a school rejects a student, then they will be at capacity in every future step and so every school is at capacity at the end of step  $k$ . That is, students are rejected in step  $k$  and every school is tentatively filled to capacity with students at the end of step  $k$ . This is a contradiction, as there is an equal measure of students and seats.

## A.4 Proof of Theorem 3

Consider an uninformed student  $j$ . The cardinal utility for a school is some function  $v(x)$  of its ordinal rank  $x$ , where  $v$  is strictly decreasing (so that  $x = 1$  is the best possible rank). Thus,  $j$ 's (ex-ante) expected utility for school  $s$  in state  $\omega$  is

$$\bar{v}_j^\omega(s) = \sum_{P \in \mathcal{P}} \lambda_\omega(P) v(\text{rank}_P(s)).$$

For a state  $\omega = (s^{(1)}, \dots, s^{(M)})$ , we first show that  $\bar{v}_j^\omega(s^{(1)}) > \bar{v}_j^\omega(s^{(2)}) > \dots > \bar{v}_j^\omega(s^{(M)})$ . Consider any positive integers  $k < l \leq M$  and, similar to the proof of the previous propo-

sition, consider  $P$  and  $P'$  with  $P(r) = P'(t) = s^{(k)}$  and  $P(t) = P'(r) = s^{(l)}$  for some  $r < t$ , and  $P(k) = P'(k)$  for all  $k \neq r, t$ . The contribution to  $\bar{v}_j^\omega(s^{(k)})$  from just  $P$  and  $P'$  is  $\lambda_\omega(P)v(r) + \lambda_\omega(P')v(t)$  and to  $\bar{v}_j^\omega(s^{(l)})$  is  $\lambda_\omega(P)v(t) + \lambda_\omega(P')v(r)$ . As  $\lambda_\omega(P) > \lambda_\omega(P') > 0$  (since  $P$  is closer to state  $\omega$  than  $P'$ ) and  $v(r) > v(t)$  (as  $r < t$ ) it follows that the contribution is greater for  $\bar{v}_j^\omega(s^{(k)})$  than for  $\bar{v}_j^\omega(s^{(l)})$ . Partitioning  $\mathcal{P}$  into  $M!/2$  pairs  $\{P, P'\}$  and summing over all members of the partition, we conclude  $\bar{v}_j^\omega(s^{(k)}) > \bar{v}_j^\omega(s^{(l)})$ . As  $k$  and  $l$  were arbitrary we have established that  $\bar{v}_j^\omega(s^{(1)}) > \bar{v}_j^\omega(s^{(2)}) > \dots > \bar{v}_j^\omega(s^{(M)})$ .

Second, by symmetry, for any other state  $\tilde{\omega} = (\tilde{s}^{(1)}, \dots, \tilde{s}^{(M)})$ , we have  $\bar{v}_j^{\tilde{\omega}}(\tilde{s}^{(k)}) = \bar{v}_j^\omega(s^{(k)})$  for all  $k = 1, \dots, M$ . With slight abuse of notation, define  $\bar{v}_k := \bar{v}_j^\omega(s^{(k)})$  for any  $\omega, k$ . In other words, we have shown that there are  $M$  numbers  $\bar{v}_1 > \dots > \bar{v}_M$  such that, conditional on any state  $\omega = (s^{(1)}, \dots, s^{(M)})$ ,  $j$ 's expected utility for school  $s^{(1)}$  is  $\bar{v}_1$ , for  $s^{(2)}$  is  $\bar{v}_2$ , etc. In other words, her expected utility for a school  $s$  depends only on the  $rank_\omega(s)$  (note that this may be different from  $j$ 's own ordinal ranking of  $s$ , which she only learns ex-post).

Assume that all other players are playing their equilibrium strategy. Since  $j$ 's strategy must be measurable with respect to her information, we can identify each of her possible strategies  $\sigma_j$  with a (mixture over) the space ordinal preference relations  $\mathcal{P}$ . Let  $EU_j(P_j)$  be  $j$ 's expected utility when she reports  $P_j$  (and everyone else plays their equilibrium strategy). More formally, using the above definitions,

$$EU_j(P_j) = \sum_{\omega \in \Omega} \sum_{k=1}^M \Pr(\omega) \times \Pr(j \text{ receives } s^{(k)} | \sigma_{-j}^*, \omega, P_j) \times \bar{v}_k.$$

In stating the next lemma, we slightly abuse notation and let  $\ell_j(s)$  denote  $j$ 's overall priority score at school  $s$  (i.e.  $\ell_j(s)$  is the priority number for  $j$ 's secure school and the priority number plus 1 for all other schools).

**Lemma 3.** *Let  $s_A$  and  $s_B$  be two schools such that  $\ell_j(s_A) > \ell_j(s_B)$ , and consider a preference report  $P_j$  such that  $P_j(r) = s_A$  and  $P_j(r+1) = s_B$ . Let  $P'_j$  be the alternative report such that  $P'_j(r) = s_B$ ,  $P'_j(r+1) = s_A$ , and, for all other  $t \neq r, r+1$ ,  $P'_j(t) = P_j(t)$ . Then,  $EU_j(P_j) \leq EU_j(P'_j)$ .*

Given this lemma, consider any arbitrary strategy  $P_j$  for student  $j$  that does not rank her secure school,  $\bar{s}$ , first, i.e.,  $P_j : s_1, s_2, \dots, s_{r-2}, s_{r-1}, \bar{s}, s_{r+1}, \dots$ . By the lemma,  $EU_j(P_j) \leq EU_j(P'_j)$ , where  $P'_j : s_1, s_2, \dots, s_{r-2}, \bar{s}, s_{r-1}, s_{r+1}, \dots$ . Applying the lemma again,  $EU_j(P'_j) \leq EU_j(P''_j)$ , where  $P''_j : s_1, s_2, \dots, \bar{s}, s_{r-2}, s_{r-1}, s_{r+1}, \dots$ . Continuing in this manner, we eventually find a strategy  $P_j^* : \bar{s}, s_1, s_2, \dots$  such that  $EU_j(P_j^*) \geq EU_j(P_j)$ . Any strategy that ranks  $\bar{s}$  first gives  $j$  school  $\bar{s}$  in every state, with associated expected utility  $\frac{1}{M} \sum_{k=1}^M \bar{v}_k$ . As all

strategies that rank  $\bar{s}$  first give the same expected utility, any such strategy is optimal for player  $j$ .

We now prove the lemma.

*Proof of Lemma 3.*

Proposition 2 shows that for any state  $\omega = (s^{(1)}, \dots, s^{(M)})$ , the lottery cutoffs can be written  $\bar{\ell}_{s^{(1)}} < \dots < \bar{\ell}_{s^{(M)}}$ . In addition, note that by the symmetry assumption on preferences, the cutoffs are independent of the state; that is, given any two states  $\omega = (s^{(1)}, \dots, s^{(M)})$  and  $\tilde{\omega} = (\tilde{s}^{(1)}, \dots, \tilde{s}^{(M)})$  and corresponding vectors of cutoffs, we have  $\bar{\ell}_{s^{(m)}} = \bar{\ell}_{\tilde{s}^{(m)}}$  for all  $m$ . In other words, we can just write  $\bar{\ell}_1 < \dots < \bar{\ell}_M$  for the schools ranked first to last in any state. Let  $\mu_j^\omega(P_j)$  denote  $j$ 's match when  $j$  reports  $P_j$  in state  $\omega$ .

Start by partitioning the state space into  $\Omega = \Omega_A \cup \Omega_B$ , where  $\omega \in \Omega_A$  if and only if  $\text{rank}_\omega(s_A) < \text{rank}_\omega(s_B)$  (and  $\Omega_B = \Omega \setminus \Omega_A$ ). As described in the proof of Proposition 2, for each  $\omega_A \in \Omega_A$ , there is a corresponding  $\omega_B$  that swaps the positions of  $s_A$  and  $s_B$ , and leaves all other schools the same. Consider one such pair  $(\omega_A, \omega_B)$ .

Let  $s = \mu_j^{\omega_A}(P_j)$  and first suppose that  $\text{rank}_{P_j}(s) < r$ . In other words,  $j$  is matched to a school she reported as preferred to  $s_A$  in state  $\omega_A$ . This implies that  $\ell_j(s) \leq \bar{\ell}_{\text{rank}_{\omega_A}(s)}$  and  $\ell_j(s') > \bar{\ell}_{\text{rank}_{\omega_A}(s')}$  for all  $s'$  with  $\text{rank}_{P_j}(s') < \text{rank}_{P_j}(s)$ . But, note that in moving to state  $\omega_B$ , the rankings of  $s$  and all such  $s'$  do not change ( $s_B$  is not one of the  $s'$  as  $\text{rank}_{P_j}(s_B) = r+1$ ) so the cutoffs  $\bar{\ell}_{\text{rank}_{\omega_A}(s)}$  and  $\bar{\ell}_{\text{rank}_{\omega_A}(s')}$  do not change either. Furthermore, by construction,  $\text{rank}_{P_j}(s) = \text{rank}_{P'_j}(s)$  and  $\text{rank}_{P_j}(s') = \text{rank}_{P'_j}(s')$  for all such  $s'$  and so  $j$  is always matched to  $s$ . In summary, we conclude that  $j$ 's match is the same in all of these scenarios:  $\mu_j^{\omega_A}(P_j) = \mu_j^{\omega_B}(P_j) = \mu_j^{\omega_A}(P'_j) = \mu_j^{\omega_B}(P'_j) = s$ .

Second, suppose that  $\text{rank}_{P_j}(s) \geq r$ . For ease of notation, define  $\text{rank}_{\omega_A}(s_A) = \text{rank}_{\omega_B}(s_B) = k$  and  $\text{rank}_{\omega_A}(s_B) = \text{rank}_{\omega_B}(s_A) = k'$ , where  $k < k'$ . There are several subcases, depending on the relative magnitudes of  $\ell_j(s_A), \ell_j(s_B), \bar{\ell}_k$ , and  $\bar{\ell}_{k'}$ . Recall that  $\ell_j(s_B) < \ell_j(s_A)$  (by assumption) and  $\bar{\ell}_k < \bar{\ell}_{k'}$  (by Proposition 2), which will eliminate many possibilities.

Subcase (i):  $\ell_j(s_B) < \ell_j(s_A) \leq \bar{\ell}_k < \bar{\ell}_{k'}$ . Note that  $j$  has a priority score good enough to be admitted to both  $s_A$  and  $s_B$  in both states  $\omega_A$  and  $\omega_B$ . Thus, she will be admitted to whichever school she ranks higher in her preferences, regardless of the state. That is,  $\mu_j^{\omega_A}(P_j) = \mu_j^{\omega_B}(P_j) = s_A$  and  $\mu_j^{\omega_A}(P'_j) = \mu_j^{\omega_B}(P'_j) = s_B$ .<sup>35</sup>

Subcase (ii):  $\ell_j(s_B) \leq \bar{\ell}_k < \ell_j(s_A) \leq \bar{\ell}_{k'}$ . In this case, if  $j$  submits  $P'_j$ , then she will be admitted to  $s_B$  in both states. However, if she submits  $P_j$ , then she will only be admitted to  $s_A$  in state  $\omega_B$ , since her priority score is not good enough in state  $\omega_A$ . Thus,  $\mu_j^{\omega_B}(P_j) = s_A$  and  $\mu_j^{\omega_A}(P_j) = \mu_j^{\omega_A}(P'_j) = \mu_j^{\omega_B}(P'_j) = s_B$ .

---

<sup>35</sup>Recall that she can do no "better" (according to her reported preferences), since we are in the case  $\text{rank}_{P_j}(s) \geq r$ .

Subcase (iii):  $\ell_j(s_B) \leq \bar{\ell}_k < \bar{\ell}_{k'} < \ell_j(s_A)$ : In this case,  $j$ 's priority score is not good enough to be admitted to  $s_A$  in either state, but is good enough for  $s_B$  in both states. That is,  $\mu_j^{\omega_A}(P_j) = \mu_j^{\omega_B}(P_j) = \mu_j^{\omega_A}(P'_j) = \mu_j^{\omega_B}(P'_j) = s_B$ .

Subcase (iv):  $\bar{\ell}_k < \ell_j(s_B) < \ell_j(s_A) < \bar{\ell}_{k'}$ : In this case,  $j$  is matched to the lower-ranked school in both states regardless of which she submits. That is,  $\mu_j^{\omega_A}(P_j) = \mu_j^{\omega_A}(P'_j) = s_B$  and  $\mu_j^{\omega_B}(P_j) = \mu_j^{\omega_B}(P'_j) = s_A$ .

Subcase (v):  $\bar{\ell}_k < \ell_j(s_B) \leq \bar{\ell}_{k'} < \ell_j(s_A)$ : In this case,  $j$ 's priority score is not good enough to be admitted to  $s_A$  in either state, but is good enough to be admitted to  $s_B$  in state  $\omega_A$ , which happens under both  $P_j$  and  $P'_j$ . In state  $\omega_B$ ,  $j$  does not have a good enough priority number for  $s_A$  or  $s_B$ , and so she gets some school  $s$  that is ranked (strictly) worse than  $(r+1)^{th}$ . Recall that  $P_j(t) = P'_j(t)$  for all  $t > r+1$ , and so this will be the same school  $s$  under both reports in state  $\omega_B$ . To summarize, in this case we have  $\mu_j^{\omega_A}(P_j) = \mu_j^{\omega_A}(P'_j) = s_B$  and  $\mu_j^{\omega_B}(P_j) = \mu_j^{\omega_B}(P'_j) = s$  for some  $s$  such that  $rank_{P_j}(s) = rank_{P'_j}(s) = t > r+1$ .

Subcase (vi):  $\bar{\ell}_k \leq \bar{\ell}_{k'} < \ell_j(s_B) < \ell_j(s_A)$ : In this case,  $j$  does not have a good enough priority score for either  $s_A$  or  $s_B$  in either state  $\omega_A$  or  $\omega_B$ . By similar reasoning to subcase (v), we have  $\mu_j^{\omega_A}(P_j) = \mu_j^{\omega_A}(P'_j) = \mu_j^{\omega_B}(P_j) = \mu_j^{\omega_B}(P'_j) = s$  for some  $s$  such that  $rank_{P_j}(s) = rank_{P'_j}(s) = t > r+1$ .

Looking back through all of the subcases,  $j$ 's assignment is independent of her choice between reporting  $P_j$  and  $P'_j$  (for a fixed state) in all cases except subcases (i) and (ii). In subcase (i), if she reports  $P_j$ , she gets  $s_A$  in both states. Since both states are equally likely, her expected utility conditional on the true state being  $\omega \in \{\omega_A, \omega_B\}$  is  $\frac{1}{2}(\bar{v}_k + \bar{v}_{k'})$ . If she reports  $P'_j$ , she gets  $s_B$  in both states, and again her expected utility conditional on the true state being  $\omega \in \{\omega_A, \omega_B\}$  is  $\frac{1}{2}(\bar{v}_k + \bar{v}_{k'})$ . Thus, in this subcase again,  $j$  is indifferent between  $P_j$  and  $P'_j$ . Last, consider subcase (ii). In this case, if she reports  $P_j$ , she receives the  $(k')^{th}$ -ranked school (the worse school of  $s_A$  and  $s_B$ ) in both states  $\omega_A$  and  $\omega_B$ , for a expected utility conditional on  $\omega \in \{\omega_A, \omega_B\}$  of  $\bar{v}_{k'}$ . If she reports  $P'_j$ , she receives  $s_B$  in both states, for a conditional expected utility of  $\frac{1}{2}(\bar{v}_k + \bar{v}_{k'}) > \bar{v}_{k'}$ . In this case, she strictly prefers to report  $P'_j$ .

In summary,  $j$  always weakly prefers  $P'_j$  to  $P_j$ , and she strictly prefers it if her priority scores fall in subcase (ii) conditioned on the state being either  $\omega_A$  or  $\omega_B$ . Formally,  $EU_j(P_j|\omega \in \{\omega_A, \omega_B\}) \leq EU_j(P'_j|\omega \in \{\omega_A, \omega_B\})$ . As every pair of states is equally likely ex-ante, summing over all such pairs gives  $EU_j(P_j) \leq EU_j(P'_j)$ .  $\square$

## A.5 Proof of Theorem 4

By the properties of DA, informed students have no justified claims at a school they prefer. Thus, consider an uninformed student  $j$ . Let her secure school where she has high priority be  $\bar{s}$  (and note that she is matched to  $\bar{s}$  in the equilibrium). Consider  $j$  potentially proposing a block with some other school  $s'$ . We use many of the ideas and notation from the proof of Proposition 2 and Theorem 3. In particular, we again partition the states into two groups  $\Omega_{\bar{s}}$  and  $\Omega_{s'}$ :  $\Omega_{\bar{s}}$  consisting of those where the rank of  $\bar{s}$  is lower and  $\Omega_{s'}$  consisting of those where the rank of  $s'$  is lower. Take two states  $\omega \in \Omega_{\bar{s}}$  and  $\hat{\omega} \in \Omega_{s'}$  that differ only in that the relative positions of  $\bar{s}$  and  $s'$  are swapped. Let  $\text{rank}_{\omega}(s') = \text{rank}_{\hat{\omega}}(\bar{s}) = k$  and  $\text{rank}_{\omega}(\bar{s}) = \text{rank}_{\hat{\omega}}(s') = k'$ , where  $k' > k$ . Recall that  $\bar{\ell}_k < \bar{\ell}_{k'}$  (Proposition 2) and  $\bar{v}_k > \bar{v}_{k'}$  (proof of Theorem 3). There are three possibilities for the outcome states  $\omega$  and  $\hat{\omega}$  if  $j$  proposes a block with  $s'$ .

Case (i):  $\ell_j(s') \leq \bar{\ell}_k$ . In this case,  $i$  will be rematched with  $s'$  in both states  $\omega$  and  $\hat{\omega}$ . So  $i$ 's payoff conditional on the true state being in  $\{\omega, \hat{\omega}\}$  is  $\frac{1}{2}(\bar{v}_k + \bar{v}_{k'})$ , whether she stays with  $\bar{s}$  or proposes a block with  $s'$ .

Case (ii):  $\bar{\ell}_k < \ell_j(s') \leq \bar{\ell}_{k'}$ . In this case,  $i$  is rematched with  $s'$  in state  $\hat{\omega}$ , but not in state  $\omega$ . So  $i$ 's payoff conditional on the true state being in  $\{\omega, \hat{\omega}\}$  is  $\bar{v}_{k'} < \frac{1}{2}(\bar{v}_k + \bar{v}_{k'})$ . She is worse off from proposing a block.

Case (iii):  $\bar{\ell}_{k'} \leq \ell_j(s')$ . In this case,  $i$  will not rematch to  $s'$  in either state  $\omega$  or  $\hat{\omega}$ , and hence she is indifferent between proposing a block or not.

Combining these three cases, we see that conditional on the true state lying in  $\{\omega, \hat{\omega}\}$ ,  $i$  prefers to stay at school  $\bar{s}$  (she is indifferent in cases (i) and (iii), and is strictly better off in case (ii)). As  $j$  is matched to  $\bar{s}$  in every state,  $j$  does not update their beliefs over the true state upon observing their own match. That is, she must consider the expected utility of the block conditioned on the state being in  $\Omega$ . Summing over each of the  $M!/2$  pairs of states from  $\Omega_{\bar{s}}$  and  $\Omega_{s'}$  does just this. As she (weakly) prefers not proposing the block for each pair, she (weakly) prefers not proposing the block conditioned on the state being in  $\Omega$  and thus the matching is ex-post stable.  $\square$

## A.6 Proof of Proposition 3

The first thing to show is that if all students choose  $\hat{\sigma}$  then Proposition 2 holds as well. Fix a state  $\omega$ . Let  $\gamma_{\omega}(P)$  denote the measure of students who submit preference profile  $P$ . As the uninformed students are equally likely to have any priority numbers, the measure of uninformed students who submit  $P$  is  $\frac{M-\nu}{M!}$  for every profile  $P$ . So  $\gamma_{\omega}(P) = \nu\lambda_{\omega}(P) + \frac{M-\nu}{M!}$ . In particular, using the same  $P$  and  $P'$  defined in the proof of Proposition 2, we have

$\lambda_\omega(P) > \lambda_\omega(P')$  and therefore  $\gamma_\omega(P) > \gamma_\omega(P')$ .

The proof follows very similarly to the proof of Proposition 2 and so only the differences will be noted here. One conceptual difference though is that we count up the measure of students who submit preference profiles  $P$  and  $P'$  who match to  $s_A$  and  $s_B$  rather than the measure of students whose true preferences are  $P$  and  $P'$  who match to  $s_A$  and  $s_B$ . The measures are actually a little easier to calculate without secure schools. The measure of students assigned to  $s_A$  that submit preference type  $P$  or  $P'$  is

$$\delta(P, s_A)\gamma_\omega(P) + \delta(P', s_A)\gamma_\omega(P')$$

where  $\delta(P, s_A)$  and  $\delta(P', s_A)$  are

$$\begin{aligned}\delta(P, s_A) &= \prod_{x=1}^{r-1} (1 - \bar{\ell}_{P(x)}) \times \bar{\ell}_{s_A} \\ \delta(P', s_A) &= \prod_{x=1, x \neq r}^{t-1} (1 - \bar{\ell}_{P(x)}) \times (1 - \bar{\ell}_{s_B}) \times \bar{\ell}_{s_A}\end{aligned}$$

The total measure of students matched to  $s_B$  is

$$\delta(P, s_B)\gamma_\omega(P) + \delta(P', s_B)\gamma_\omega(P')$$

where the  $\delta$ 's in this case are defined as

$$\begin{aligned}\delta(P, s_B) &= \prod_{x=1, x \neq r}^{t-1} (1 - \bar{\ell}_{P(x)}) \times (1 - \bar{\ell}_{s_A}) \times \bar{\ell}_{s_B} \\ \delta(P', s_B) &= \prod_{x=1}^{r-1} (1 - \bar{\ell}_{P(x)}) \times \bar{\ell}_{s_B}\end{aligned}$$

We still have  $\delta(P, s_A) \geq \delta(P', s_B)$  and  $\delta(P', s_A) \geq \delta(P, s_B)$  and so the rest of the proof follows similarly. The one final difference is that we also may have a weak inequality if  $\bar{\ell}_{P(x)} = 0$  for all  $x = r, \dots, t-1$  but this is not a problem because without secure schools it is obvious that  $\bar{\ell}_s > 0$  for all  $s$  (otherwise, some strictly positive mass of students remains unassigned).

Lemma 3 depended only on the cutoffs increasing and so it applies now as well. In particular, suppose an uninformed student  $j$  chose a profile  $P_j$  where  $\text{rank}_{P_j}(s) < \text{rank}_{P_j}(s')$  for some  $s$  and  $s'$  where  $\ell_j(s) > \ell_j(s')$ . We cannot directly apply Lemma 3 because  $P_j$  may not rank them consecutively. However, ranked between  $s$  and  $s'$  there must be consecutively

ranked schools where  $j$ 's priority is worse at the better ranked school. In particular, there must be some schools  $\hat{s}$  and  $\hat{s}'$  with  $rank_{P_j}(\hat{s}) = rank_{P_j}(\hat{s}') - 1$  and  $\ell_j(\hat{s}) > \ell_j(\hat{s}')$ . For these schools, the lemma does apply to show that  $P'_j : s_1, s_2, \dots, \hat{s}', \hat{s}, \dots$  and  $EU_j(P_j) \leq EU_j(P'_j)$ . This process can be iterated as long there exists  $s$  and  $s'$  where  $rank_{P_j}(s) < rank_{P_j}(s')$  and  $\ell_j(s) > \ell_j(s')$ . Eventually this will lead to the strategy profile  $\hat{\sigma}_j$  which ranks schools in increasing order of priority numbers.  $\square$

## A.7 Proof of Theorems 5 and 6

Theorems 5 and 6 are closely related, and so we prove them jointly. Mathematically, we show the following:

- (i) For all  $j \in U$ ,  $V_j^{SS}(\sigma^*) \geq V_j^{NC} \geq V_j^{NSS}(\hat{\sigma})$ .
- (ii) For all  $j \in I$ ,  $V_j^{SS}(\sigma^*) \geq V_j^{NC}$ .
- (iii) There exists a subset  $\tilde{J} \subseteq U$  with strictly positive measure such that  $V_j^{NC} > V_j^{NSS}(\hat{\sigma})$  for all  $j \in \tilde{J}$ .
- (iv) There exists a subset  $\tilde{J} \subseteq I$  with strictly positive measure such that  $V_j^{SS}(\sigma^*) > V_j^{NC}$  for all  $j \in \tilde{J}$ .

To start, consider an uninformed student  $j \in U$ . Under both DA-SS and NC,  $j$  is matched to their secure school in every state, and so  $V_j^{SS}(\sigma^*) = V_j^{NC} = \bar{v} := \frac{1}{M} \sum_{k=1}^M \bar{v}_k$ . We will show that  $V_j^{NSS}(\hat{\sigma}) \leq \bar{v}$  for all  $j \in U$ , which will complete the proof of (i).

First, partition the state space  $\Omega$  as follows. Take an arbitrary state  $\omega$ , and suppose that in equilibrium, in state  $\omega$ ,  $j$  is matched to school  $s$ . Let  $r = rank_{\omega}(s)$ . Consider every state  $\omega'$  which maintains the relative rankings of all schools other than  $s$  but moves  $s$  to a worse ranking (and include  $\omega$  too for ease of notation later). Formally, define the set

$$\Omega_1 = \{\omega' : rank_{\omega'}(s) \geq r \text{ and } rank_{\omega'}(s') < rank_{\omega'}(s'') \text{ iff } rank_{\omega'}(s') < rank_{\omega'}(s'') \text{ for all } s', s'' \neq s\}$$

As this just moves  $s$  down the rankings, each school other than  $s$  (weakly) moves up in rankings and so, by the extension of Proposition 2 in the previous proof,  $\bar{\ell}_{s'}$  is smaller in state  $\omega'$  than in  $\omega$  for all  $s' \neq s$ . On the other hand,  $s$  gets worse and so  $\bar{\ell}_s$  gets larger. Therefore, as  $j$  is matched to  $s$  in  $\omega$ ,  $j$  is also matched to  $s$  in  $\omega'$ . There is precisely one state in  $\Omega_1$  where  $rank_{\omega'}(s) = k$  for each  $k = r, \dots, M$ . Thus, conditional on the true state lying

in  $\Omega_1$ ,  $j$ 's expected utility is

$$\frac{1}{M-r+1} \sum_{k=r}^M \bar{v}_k \leq \frac{1}{M} \sum_{k=1}^M \bar{v}_k = \bar{v}.$$

Now, consider any state  $\omega \in \Omega \setminus \Omega_1$ , and form  $\Omega_2$  in the same way (find the school  $j$  is matched to in  $\omega$  and then move it down the rankings). By construction, once again,  $j$ 's expected utility conditional on the state being in  $\Omega_2$  is weakly less than  $\bar{v}$ .

Continue this procedure until we run out of states. Note that it might be that  $\Omega_l \cap \Omega_k \neq \emptyset$  for some  $l, k$ . However, if this is the case, then either  $\Omega_k \subset \Omega_l$  or  $\Omega_l \subset \Omega_k$ , and so we simply choose the larger set to form the partition.<sup>36</sup> Now, conditional on each element of the partition,  $j$ 's expected utility is at most  $\bar{v}$ ; as all states are equally likely ex-ante,  $V_j^{NSS}(\hat{\sigma}) \leq \bar{v}$ . This completes the proof of (i).

To show (ii), consider an informed student  $j \in I$  with ordinal preferences  $P_j$ , and let her secure school be  $\bar{s}$ . Under No Choice,  $j$  is assigned to  $\bar{s}$ . Under DA-SS, since  $j$  knows her preferences and it is optimal for informed students to report truthfully, it is clear that her final match will be (weakly) preferred to  $\bar{s}$ . Thus, we conclude that  $V_j^{SS}(\sigma^*) \geq V_j^{NC}$ .

So far, we have shown DA-SS weakly Pareto dominates NC for all students, and NC weakly Pareto dominates DA-NSS for the uninformed students. What is left to show is that these Pareto dominance relations are actually strict, which are statements (iii) and (iv) above.

We first show (iv), that a strictly positive mass of students prefer DA-SS to NC. Fix a state  $\omega = (s^{(1)}, \dots, s^{(M)})$ . Recall that under DA-SS, we have  $\bar{\ell}_{s^{(1)}} \in (1, 2)$  (Proposition 2). Now, for any small  $\epsilon > 0$ , there is a set of informed students of mass  $\epsilon^M \times \nu$  that have priority  $\ell_j(s) \leq 1 + \epsilon$  for all  $s \in S$ .<sup>37</sup> By choosing  $\epsilon$  such  $1 + \epsilon \leq \bar{\ell}_{s^{(1)}}$ , we have a set of students  $I'$  who all receive their first choice school, no matter what it is. Since the distribution of secure schools is independent of preferences, we can find a subset  $I'' \subseteq I'$  of strictly positive measure such that for all  $j \in I''$ ,  $j$ 's first-choice school is different from her secure school.<sup>38</sup> All of the students in  $I''$  strictly prefer DA-SS to NC.

Last, we use a similar argument to show (iii), that a strictly positive mass of uninformed students strictly prefer NC to DA-NSS. To do so, first recall that in every state, the second-

<sup>36</sup>Suppose  $\Omega_l$  was created by starting with some  $\omega$  where  $j$  matches to  $s$  and  $rank_\omega(s) = r$ . If  $\Omega_k \cap \Omega_l \neq \emptyset$ , then  $\Omega_k$  must have been created by starting with some  $\omega'$  where  $j$  matches to  $s$  and the rankings of all other schools are preserved. So if  $rank_{\omega'}(s) = r' < r$  then  $\Omega_l \subset \Omega_k$ , and if  $rank_{\omega'}(s) = r' > r$  then  $\Omega_k \subset \Omega_l$ .

<sup>37</sup>The 1 comes from the fact that we add 1 to the priority of any student for whom  $s$  is not their secure school.

<sup>38</sup>There are many ways to construct this set, but, for example, consider two schools  $s, s'$ , and some ordinal preference  $P$  that ranks  $s$  first. Then, there is a subset  $I''$  of measure  $\epsilon^M \times \nu \times \frac{1}{M} \times \lambda_\omega(P)$  whose secure school is  $s'$ , but rank school  $s$  first and have high enough priority to be admitted to it.

to-highest cutoff satisfies  $\bar{\ell}_{s(M-1)} < 1$ .<sup>39</sup> Consider some small  $\epsilon > 0$  such that  $1 - \epsilon > \bar{\ell}_{s(M-1)}$ . There is a set of uninformed students  $\tilde{U} \subset U$  of mass  $\epsilon^M \times (M - \nu)$  such that for all  $j \in \tilde{U}$ ,  $\ell_j(s) \geq 1 - \epsilon > \bar{\ell}_{s(M-1)}$  for all schools  $s \in S$ . In other words, these students will receive the worst school in every state state, and hence  $V_j^{NSS}(\hat{\sigma}) = 0$  for all  $j \in \tilde{U}$ . Since  $V_j^{NC} = \bar{v} > 0$  for these students, they strictly prefer NC to DA-NSS.

## B An ex-post stable mechanism for Example 2

Example 2 shows that the DA mechanism is not ex-post stable. However, there is an ex-post stable full matching in that example. This leaves open the question of whether ex-post stable full matchings always exist and if/so, whether we can find an ex-post stable mechanism. In this appendix, we show that the answer to both of these questions is yes. But, this comes with a strong caveat in that the mechanism must ignore stated preferences. We formally describe this below, and then discuss why we think ignoring stated preferences is not practical for real-world implementation.

We start by showing that an ex-post stable full matching exists and then find a mechanism that is ex-post stable. To find an ex-post stable full matching, implement any matching that is stable in the classical sense (as defined at the beginning of Section 2.3) in each state. This will be ex-post stable because classical stability means that if in state  $\omega$ ,  $|\mu_s| < q_s$  or  $j \succ_s j'$  for some  $j' \in \mu_s$  then  $u_j^\omega(s) \leq u_{j'}^\omega(\mu_{j'}^\omega)$ . Therefore  $u_j^\omega(s) \leq u_j^\omega(\mu_j^\omega)$  for every  $\omega \in C_j(\mu_j^\omega, s)$  and so the expected value is also less from  $s$ . An ex-post stable full matching can be implemented with what we call the *state-learning DA mechanism* (as long as  $|I| \geq 3$ ). This mechanism is defined as follows:

**State-Learning:** Fix an arbitrary  $\bar{\omega}$ . If the reported preferences are all  $\bar{\omega}$  or  $\omega$  for some  $\omega \in \Omega$ , or there is a single report that is not  $\bar{\omega}$  or  $\omega$  and at least two reports that are  $\omega$ , then determine that the state is  $\omega$  (including the case where  $\omega = \bar{\omega}$ ). Otherwise, determine that the state is  $\bar{\omega}$ .

**DA:** Run DA with all students submitting the determined state  $\omega$ .

The idea of the mechanism is that all the uninformed students submit  $\bar{\omega}$  while all the informed students submit the true state  $\omega$ . This is incentive compatible and the state is identified (as long as  $|I| \geq 3$ ) because any unilateral deviation is ignored. Then, the mechanism runs DA to guarantee a matching that is stable in the classical sense in every state (alternatively, it could run any stable mechanism after determining the state).

---

<sup>39</sup>Under DA-NSS, there are no secure schools, and so the cutoffs range from  $[0, 1]$ , rather than  $[0, 2]$ .

However, the state-learning mechanism (and as we shall argue other ex-post stable mechanisms) is odd, because it may not satisfy the following property. For a matching  $\mu$ , call a student and a school  $(j, s)$  a *blocking pair with respect to  $P$*  if: (i)  $sP_j\mu_s$  and (ii) either  $|\mu_s| < q_s$  or  $j \succ_s j'$  for some  $j' \in \mu_s$ . A matching  $\mu$  is then called *stable with respect to  $P$*  if there are no blocking pairs with respect to  $P$ . A mechanism is called *stable with respect to  $P$*  if, for every state  $\omega$ , the equilibrium reported ordinal preference profile  $P$  induces a matching  $\mu$  that is stable with respect to  $P$ .

Going back to Example 2, the unique ex-post stable full matching has  $\mu_{j_4} = C$  and  $\mu_{j_2} = D$  in  $\omega = ABCD$  and  $\mu_{j_4} = D$  in  $\omega = BADC$ . Suppose first that  $j_4$  reports  $DP_{j_4}C$ . Then, any mechanism that is stable with respect to  $P$  cannot assign  $j_4$  to  $C$  and  $j_2$  to  $D$ , because then  $j_4$  and  $D$  would form a blocking pair. Likewise, if  $j_4$  reports  $CP_{j_4}D$ , then any mechanism that is stable with respect to  $P$  cannot assign  $j_4$  to  $D$  and  $j_3$  to  $C$ , because if  $j_4$  was informed then  $j_4$  and  $C$  would form a blocking pair. That is, no mechanism that is stable with respect to  $P$  can implement the ex-post stable full matching.

So why do we care whether the mechanism is stable with respect to  $P$ ? In our model, there is a common ordinal preference so once the mechanism knows what it is, it can just implement a matching that is stable relative to that preference ordering. But it also means that the mechanism can learn everyone's preferences from just learning one person's preference and, importantly, this means that the mechanism can learn the preferences of the uninformed students. In reality, this is not the case (in fact, if it were, one might argue that we would not need school choice anyways in the first place, since we already know everyone's preferences). We do this here to make a tractable point about the curse of acceptance, and in fact, later generalize to allow for imperfect correlation. Given that the mechanism will not know everyone's preferences and will not know who is informed and who is uninformed, it seems that for real-world implementation it should "take people at their word". In the context of our example, the mechanism would not know that  $j_4$  was uninformed and so, for example, assigning her to  $C$  when she stated a preference for  $D$  and she has high enough priority to get  $D$  seems impractical (for example, if the mechanism designer was wrong and she actually is an informed student, she would have justified envy).