

First-Choice Maximal and First-Choice Stable School Choice Mechanisms*

Umut Dur[†] Timo Mennle[‡] Sven Seuken[‡]

May 18, 2018

Abstract

We investigate the class of school choice mechanisms that are *first-choice maximal (FCM)* (i.e., they match a maximal number of students to their reported first choices) and *first-choice stable (FCS)* (i.e., no students form blocking pairs with their reported first choices). FCM is a ubiquitous desideratum in school choice, and we show that FCS is the only rank-based relaxation of stability that is compatible with FCM. The class of FCM and FCS mechanisms includes variants of the well-known Boston mechanism as well as certain Asymmetric Chinese Parallel mechanisms. Regarding incentives, we show that while no mechanism in this class is strategyproof, the Pareto efficient ones are least susceptible to manipulation. Regarding student welfare, we show that the Nash equilibrium outcomes of these mechanisms correspond precisely to the set of stable matchings. By contrast, when some students are sincere, we show that more students may be matched to their true first choices in equilibrium than under any stable matching. Finally, we show how our results can be used to obtain a new characterization of the Boston mechanism (i.e., the most widely used FCM and FCS mechanism). On a technical level, this paper provides new insights about an influential class of school choice mechanisms. For practical market design, our results yield a potential rationale for the popularity of FCM and FCS mechanisms in practice.

Keywords: School Choice, Matching, Boston Mechanism, First Choices, Stability, Vulnerability to Manipulation, Nash Equilibrium, Augmented Priorities

*This paper is published as: Dur, Umut, Timo Mennle, and Sven Seuken. “First-Choice Maximal and First-Choice Stable School Choice Mechanisms.” In *Proceedings of the 19th ACM Conference on Economics and Computation (EC)*, 2018. In addition to the content of the published paper, this version also contains the appendix with all proofs. We would like to thank (in alphabetical order) Antonio Miralles, Axel Ockenfels, Alexander Westkamp, as well as participants of the ZEW Market Design Research Seminar, the Design and Behavior Research Seminar (Uni Cologne), the 12th MiP Workshop, and MATCH-UP 2017 for helpful comments. Any errors remain our own. Part of this research was supported by the Swiss National Science Foundation under grant #156836.

[†]North Carolina State University, Department of Economics, 2801 Founders Drive, Raleigh, NC, 27695; e-mail: udur@ncsu.edu; web: <https://sites.google.com/site/umutdur/>

[‡]University of Zurich, Binzmuhlestrasse 14, CH-8050 Zurich; e-mail: {mennle,seuken}@ifi.uzh.ch; web: <https://www.ifi.uzh.ch/ce/people/>

1. Introduction

School choice programs give students an opportunity to express their preferences over which public schools they would like to attend. Ideally, one would like to match all students to their respective true first choices. However, this ideal may not be achievable because schools have limited capacities and some schools may be more popular than others. Therefore, administrators need to design school choice mechanisms that reconcile students' conflicting interests with capacity constraints. Generating high student welfare is one of the key objectives in this task (Abdulkadiroğlu and Sönmez, 2003).

One way to measure student welfare is to consider the *number of students who are matched to their first choices*. This measure is particularly tangible because maximizing it is an obvious compromise between capacity constraints and the desire to ideally match all students to their top choices. It comes as no surprise that the share of students who are matched to their first choices receives attention in the media with headlines such as

*“One in six secondary pupils in England doesn't get first school choice”*¹

and

*“45% of New York City 8th-graders got into top high school choice [...]”*²

Furthermore, administrators report the share of students who are matched to their first choices as part of their public communication. For example, Denver Public Schools prominently feature this measure on the website that informs parents about school choice.³

In line with the popularity of this measure, many school choice mechanisms in practice attempt to maximize it, the most prominent example being the *Boston mechanism (BM)* (Abdulkadiroğlu and Sönmez, 2003). Arguably, BM owes much of its popularity to the intuitive way in which it attempts to maximize first choices. The common focus on first choices motivates our definition of *first-choice maximality (FCM)*, which requires that a mechanism matches a maximal number of students to their reported first choices.

A second important desideratum in school choice is stability. A matching is called *stable* if it is *non-wasteful* (i.e., no student would rather be matched to a school with unfilled

¹Sally Weale (June 14, 2016). *The Guardian*. Retrieved October 10, 2016: www.theguardian.com

²Sophia Rosenbaum and Erica Pearson (March 11, 2014). *New York Daily News*. Retrieved October 10, 2016: nydailynews.com

³Denver Public Schools Website. Retrieved October 10, 2016: schoolchoice.dpsk12.org/schoolchoice/

seats), *individually rational* (i.e., no matched student would rather be unmatched), and if it *eliminates justified envy* (i.e., if a student prefers a different school to her match, then any student matched to that other school must have higher priority⁴) (Balinski and Sönmez, 1999). Equivalently, stability can be defined as the absence of blocking pairs.⁵ A student in a blocking pair may feel as though she has been treated unfairly by the mechanism. On top of that, she may even have a basis for pursuing legal actions against the school district.⁶ Administrators would naturally try to avoid both, perceived unfairness and the risk of legal actions. Stability is therefore a common criterion in school choice.

Our first insight in this paper is that FCM and stability are incompatible. This raises the question whether there exists a relaxed notion of stability that can serve as a useful second best but is not in conflict with FCM. We answer this question in the affirmative: A matching is called *first-choice stable (FCS)* if no student forms a blocking pair with her reported first choice (but students may form blocking pairs with other choices). We show that among all rank-based relaxations of stability, FCS is the only one that is compatible with FCM.⁷

This insight gives rise to a natural class of school choice mechanisms: We start with the ubiquitous desideratum to maximize the number of students who are matched to their first choices (i.e., FCM). Given this, it would be unreasonable to match these first choices in a way that violates priorities; thus, we also require FCS. On the other hand, all more demanding notions of stability are in conflict with FCM. This motivates our analysis of the class of mechanisms that satisfy both FCM and FCS. For the sake of brevity, we refer to these mechanisms as *first-choice (FC) mechanisms*.

The class of FC mechanisms includes all commonly employed variants of the Boston mechanism (e.g., with varying tie-breakers, varying limits on the length of preference lists, and where filled schools are or are not skipped in the application process). Such

⁴Priorities are common in school choice. For example, a school may grant priority to students whose siblings attend the same school, who live nearby, or based on grades (Abdulkadiroğlu et al., 2006).

⁵A student-school-pair (i, s) is *blocking* if i would prefer s to her current match and s has unfilled seats, s is i 's outside option, or some student who is matched to s has lower priority at s than i .

⁶A Wisconsin student's lawsuit succeeded on the grounds of *justified envy* (Abdulkadiroğlu and Sönmez, 2003). The admission to medical degree programs in Germany is organized via a centralized mechanism (Westkamp, 2013). Rejected applicants sometimes obtain a seat by suing universities for leaving seats unfilled. These lawsuits succeed on the grounds of *wastefulness*.

⁷A *rank-based relaxation of stability* requires that there are no blocking pairs for which the respective school has a certain rank in the respective student's preference order.

mechanisms are used in many school districts, including Minneapolis, Seattle, Lee County (Kojima and Ünver, 2014), San Diego, Amsterdam (until 2014) (de Haan et al., 2015), Wake County (until 2015) (Dur, Hammond and Morrill, 2016), and in Nordrhein Westfalen and Freiburg (Germany) (Basteck, Huesmann and Nax, 2015). Asymmetric Chinese Parallel mechanisms with $e_0 = 1$ also belong to this class and are used for college admissions in the Chinese Beijing, Gansu, and Shandong provinces (Chen and Kesten, 2016a). A plausible motivation for using FC mechanisms is the desire to achieve FCM and FCS. As market designers, we are of course aware of the fact that strategic misreporting by students may impede this objective. Nonetheless, administrators may find FC mechanisms appealing, for example, if they believe that students will report their preferences truthfully despite contrary incentives.⁸ Moreover, even if students strategize, administrators may prefer FC mechanisms for cosmetic reasons (e.g., if they are driven by other considerations such as favorable media coverage).

For market designers, the question arises whether and to what extent *FC mechanisms actually achieve the intended desiderata* to match a maximal number of students to their *true* first choices and to avoid blocking pairs of students with their respective *true* first choices. This research question is the focus of our present paper. To answer it, we proceed in two steps:

- Step 1. We identify the incentives for students under FC mechanisms (by comparing these mechanisms by their vulnerability to manipulation).
- Step 2. We investigate how strategic reporting by students affects outcomes (by studying the Nash equilibria of the induced preference revelation games).

Regarding **incentives** (Step 1), our first insight is that FCM alone is already incompatible with strategyproofness, even without the additional restriction of FCS. Despite this incompatibility, some FC mechanisms may invite more strategic misreporting than others. Towards understanding these differences, we employ the concept of comparing mechanisms by their *vulnerability to manipulation*, introduced by Pathak and Sönmez

⁸For example, the Teach for America (TfA) program uses a non-strategyproof mechanism to match teachers-to-be to teaching positions. However, TfA administrators feel that the use of this mechanism is justified because participants would find it hard to acquire the skills and information necessary to successfully manipulate the mechanism (Featherstone, 2015).

(2013).⁹ We show that all Pareto efficient FC mechanisms are manipulable at exactly the same problems and are therefore equivalent in this sense. Moreover, we show that for any Pareto *inefficient* FC mechanism, there exists a Pareto efficient FC mechanism that Pareto dominates the original mechanism but is also at most as manipulable.

Our results have two significant consequences: First, the two most widely studied FC mechanisms are the classic *Boston mechanism (BM)* (Abdulkadiroğlu and Sönmez, 2003) and the *adaptive Boston mechanism (ABM)* (Alcalde, 1996; Miralles, 2008; Dur, 2015; Harless, 2015; Mennle and Seuken, 2017b). One would intuitively suspect ABM to have better incentive properties than BM. However, since both mechanisms are Pareto efficient, they are manipulable at the same problems. Thus, surprisingly, the intuitive difference in their incentive properties cannot be formalized via the comparison by vulnerability to manipulation.¹⁰ The second consequence of our results pertains to FC mechanisms used in practice that are not Pareto efficient, such as Asymmetric Chinese Parallel mechanisms (Chen and Kesten, 2016a). Our results imply that these mechanisms are more manipulable and less efficient than necessary, even if administrators are restricted to only using FC mechanisms. Thus, the motivation for using Pareto *inefficient* FC mechanisms must rely on other considerations besides incentives. Otherwise, unambiguous improvements to these mechanisms would be possible, even within the class of FC mechanisms.

Regarding the **impact of strategic behavior on outcomes** (Step 2), we show that the set of Nash equilibrium outcomes of any FC mechanism corresponds precisely to the set of matchings that are stable with respect to the true preferences. This means that the equilibrium outcomes of any FC mechanism are first-choice stable with respect to the true preferences, but they may not match a maximal number of students to their true first choices.¹¹ Our result generalizes the main result of Ergin and Sönmez (2006), who showed this correspondence for BM. For market designers, the most important consequence of our result is that in markets where all students are strategic, there is no reason for using an FC mechanism instead of the *student-proposing deferred-acceptance*

⁹A mechanism φ is *at most as manipulable as* another mechanism ψ if the set of problems where some student can benefit from misreporting under φ is a subset of the respective set under ψ .

¹⁰ABM differs from BM in that students automatically skip exhausted schools in the application process under ABM, while under BM they apply to exhausted schools (thereby wasting some rounds). More nuanced assumptions are required to establish a formal understanding of the different incentive properties of BM and ABM, e.g., when school may have zero capacity or find some students unacceptable (Dur, 2015) or when priorities are random (Mennle and Seuken, 2017b).

¹¹This follows because stability implies FCS but is incompatible with FCM.

(*DA mechanism* (Abdulkadiroğlu and Sönmez, 2003): DA produces the *student-optimal* stable matching, which is the unique stable matching that Pareto dominates all other stable matchings, and DA is strategyproof. In contrast, FC mechanisms merely produce *some* stable matchings, and they do this only subject to a weaker solution concept.

Our observations so far pertain to the case when *all* students strategize. In practice, however, students may exhibit varying levels of strategic sophistication. For example, it may be cognitively challenging to determine a beneficial misreport, or acquiring the necessary information may be costly. Experimental results suggest that under BM, a significant share of the participants report their preferences truthfully despite incentives to misreport (Chen and Sönmez, 2006; Chen and Kesten, 2016*b*), and a lack of information further increases this share (Pais and Pinter, 2008).¹² We therefore follow Pathak and Sönmez (2008), who considered *mixed* problems with two types of students: *Sincere* students report their preferences truthfully, independent of incentives, while *sophisticated* students recognize the strategic aspect of the matching process. For these mixed problems, we identify the Nash equilibrium outcomes of FC mechanisms. Specifically, we show that, from the perspective of the sophisticated students, the set of Nash equilibrium outcomes corresponds to the set of matchings that are stable with respect to the true preferences and certain *augmented priorities*.¹³ Our result partially generalizes the result of Pathak and Sönmez (2008), who showed this correspondence for BM.¹⁴ Our result also implies the existence of equilibrium outcomes that are preferred by all sophisticated students to any other equilibrium outcomes.

With these results at our disposal, we can conclusively answer our main research question: To what extent do **FC mechanisms actually achieve FCM and FCS**? Towards FCS, we show that all equilibrium outcomes of any FC mechanism are FCS with respect to the true preferences. Thus, FC mechanisms actually implement FCS, independent of which students strategize and independent of which equilibrium is chosen.

¹²In laboratory experiments conducted by Chen and Sönmez (2006), this share was 13%, and it increased to 28% when priorities were random. Chen and Kesten (2016*b*) observed shares between 23% and 46% in similar experiments, and Pais and Pinter (2008) found that withholding information significantly increased these shares from 47% to 87%.

¹³*Augmented priorities* arise as an adjustment of the original priorities. They account for the positions in which students rank schools and whether or not students are sophisticated (see Section 6.1).

¹⁴The correspondence proven by Pathak and Sönmez (2008) holds for all students. Our restriction to the perspective of sophisticated students is necessary because FC mechanisms are identified by how they match students to their reported first choices but are unrestricted in how they match students to other choices (see Lemma 1). In equilibrium, this freedom only affects sincere students.

Towards FCM, we isolate the effect of any individual student’s decision (say i') whether to strategize or to report truthfully by comparing the respective sophisticated-student-optimal equilibrium outcomes. We show that under any FC mechanism, i' prefers being sophisticated to being sincere, all other sophisticated students prefer i' being sincere, and for sincere students the difference is ambiguous with one notable exception: Any student who is matched to her true first choice when i' is sophisticated is still matched to her true first choice when i' is sincere. Thus, the number of true first choices matched in strategic-student-optimal equilibrium outcomes is lowest if all students strategize, increases in mixed problems with more sincere students, and is maximal when all students are sincere. In this sense, FC mechanisms have the potential to match more true first choices in equilibrium than strategyproof alternatives like DA. This prediction is consistent with experimental findings that BM and ACPM both match significantly more students to their true first choices than DA (Result 4 of Chen and Kesten (2016b)).

Finally, we show how our results regarding the properties of FC mechanisms can be used to obtain a new characterization of BM (i.e., the most widely used FC mechanism). To this end, we first relax FCM to *best-choice-maximality*, which requires the mechanism to maximize the number of students matched to their reported *best choice that is attainable under some matching*. Similarly, we relax FCS to *best-choice-stability*, which requires the mechanism to respect each student’s priority at her *best attainable choice*. Lastly, we follow Kojima and Ünver (2014), and require a mechanism to be *consistent*, i.e., if we remove a subset of students with their assignments and update the remaining seats, then the remaining students are assigned to the same schools as before. Using our results obtained for the class of FC mechanisms we then show that BM is the unique best-choice-stable, best-choice-maximal, and consistent mechanism.

For market designers, our main results imply that the use of FC mechanisms in practice may be justified if FCM and FCS are primary desiderata and if a sufficiently large share of the students can be expected to report their preferences truthfully. Our characterization result of BM shows that, within the class of FC mechanisms, BM stands out a little more than the others, which may be a potential rationale for its popularity in practice.

2. Preliminaries

2.1. Formal Model

A *school choice problem* is a tuple $(I, S, q, P, >)$ with a finite set of *students* $I = \{i_1, \dots, i_n\}$ and a finite set of *schools* $S = \{s_1, \dots, s_m\}$. $q = (q_s)_{s \in S}$ is the vector of *school capacities* (i.e., q_s is the number of seats available at school s), $P = (P_i)_{i \in I}$ is the *preference profile* in which each P_i is the strict *preference order* of student i over the schools in S and the *outside option*, denoted by \emptyset . $> = (>_s)_{s \in S}$ is the *priority profile* in which each $>_s$ is the *priority order* of school s over students in I . $s P_i s'$ means that student i strictly prefers school s to school s' , and $i >_s i'$ means that student i has priority over student i' at school s . We assume that there is at least one seat at each school (i.e., $q_s \geq 1$ for all $s \in S$) and that all students can be matched to their outside option (i.e., $q_\emptyset = n$). For a preference order P_i , the corresponding *weak preference order* is denoted by R_i (i.e., $s R_i s'$ if either $s P_i s'$ or $s = s'$). Throughout the paper, we fix I , S , and q , and we use $(P, >)$ to denote a specific *problem*.

A *matching* is a function $\mu : I \rightarrow S \cup \{\emptyset\}$. For a given matching μ , $\mu(i)$ is the school to which student i is matched, and $\mu^{-1}(s)$ is the set of students who are matched to schools s . We focus on *feasible* matchings (i.e., $|\mu^{-1}(s)| \leq q_s$ for every $s \in S \cup \{\emptyset\}$), and we simplify notation by writing μ_i and μ_s for $\mu(i)$ and $\mu^{-1}(s)$, respectively.

For a problem $(P, >)$ and matchings μ and ν , we say that μ *weakly Pareto dominates* ν if $\mu_i R_i \nu_i$ for all students $i \in I$, and μ *Pareto dominates* ν if μ weakly Pareto dominates ν and $\mu_{i'} P_{i'} \nu_{i'}$ for at least one student $i' \in I$. The matching μ is *Pareto efficient* if it is not Pareto dominated by any other matching, μ is *individually rational* if $\mu_i R_i \emptyset$ for all students $i \in I$, μ is *wasteful* if there exists some student $i \in I$ and some school $s \in S$ such that $|\mu_s| < q_s$ but $s P_i \mu_i$, and μ is *non-wasteful* if it is not wasteful. A student $i \in I$ has *justified envy (under μ)* if there exists another student $i' \in I$ and a school $s \in S$ such that $s P_i \mu_i$, $\mu_{i'} = s$, and $i >_s i'$, and μ is *fair* if no student has justified envy. Finally, μ is *stable* if it is individually rational, non-wasteful, and fair. Observe that for any unstable matching μ , there exists at least one pair $(i, s) \in I \times (S \cup \{\emptyset\})$ such that $s P_i \mu_i$ and $|\mu_s| < q_s$ (if μ is wasteful), or $s = \emptyset$ (if μ is not individually rational), or there exists a student $i' \in I$ with $\mu_{i'} = s$ and $i >_s i'$ (if i has justified envy). Any such student-school pair is called a *blocking pair*. Obviously, the matching μ is stable if and

only if there exist no blocking pairs. Throughout the paper, we employ this equivalent definition of stability to simplify definitions and proofs.

A *mechanism* φ is a mapping that receives a problem (P, \succ) as input and selects a matching, denoted by $\varphi(P, \succ)$. We denote by $\varphi_i(P, \succ)$ the school to which student i is matched under $\varphi(P, \succ)$. A mechanism φ is called *Pareto efficient/individually rational/non-wasteful/fair/stable* if it selects matchings with the respective property for all problems. The mechanism φ *Pareto dominates* another mechanism ψ if the matching $\varphi(P, \succ)$ weakly Pareto dominates the matching $\psi(P, \succ)$ for all problems (P, \succ) and this dominance is not weak for at least one problem. Observe that these properties are formulated in terms of how the mechanisms handle reported preferences. However, students may lie about their preferences so that the input to the mechanism may differ from the true problem. Regarding incentives, a mechanism φ is called *strategyproof* if, for all problems (P, \succ) , all students $i \in I$, and all preference orders P'_i , we have $\varphi_i(P, \succ) R_i \varphi_i((P'_i, P_{-i}), \succ)$ where $P_{-i} = (P_j)_{j \neq i}$ are the preferences of all students except i .

Finally, we introduce some auxiliary notation: Given a problem (P, \succ) , let $\text{choice}_{P_i}(k)$ be the k^{th} choice of student i according to P_i . For a matching μ , let $I(\mu, k, P)$ be the set of students who are matched to their k^{th} choice under the matching μ ; formally, $I(\mu, k, P) = \{i \in I : \text{choice}_{P_i}(k) = \mu_i\}$. Thus, $I(\varphi(P, \succ), k, P)$ is the set of students who are matched to their k^{th} choice (according to P) when the mechanism φ is applied to the problem (P, \succ) .

2.2. School Choice Mechanisms

In this section, we describe common school choice mechanisms. Table 1 provides an overview of their properties as well as examples of their use in practice.

The *Boston mechanism (BM)* (Abdulkadiroğlu and Sönmez, 2003) works in rounds. In the first round, all students apply to their respective first choices, and each school permanently accepts applications from students in order of priority until all applications are accepted or until all seats are filled. Students whose applications are not accepted enter the second round where they apply to their respective second choices. Again, schools accept applications into unfilled seats by priority and reject all remaining applications once all seats are filled. This process continues (i.e., students who were rejected in round

Mechanism	SP	PE	ST	FCM	FCS	Examples of use in practice
BM		✓		✓	✓	Minneapolis, Seattle, Lee County, San Diego
ABM		✓		✓	✓	Amsterdam (until 2014), Nordrhein Westfalen
DA	✓		✓		✓	New York, Boston, Mexico City
ACPM					✓	Various Chinese provinces
ACPM, $e_0 = 1$				✓	✓	Beijing, Gansu, and Shangdong provinces
TTC	✓	✓				New Orleans Recovery School District (2012)

Table 1: Mechanisms, their properties, and examples of their use in practice; strategyproof (*SP*), Pareto efficient (*PE*), stable (*ST*).¹⁵

$k - 1$ apply to their k^{th} choices in round k) until no school receives new applications.

The *adaptive Boston mechanism (ABM)* (Alcalde, 1996; Miralles, 2008; Dur, 2015; Harless, 2015; Mennle and Seuken, 2017b) is similar to BM, except that students who are rejected in round $k - 1$ apply to their respective *most-preferred school that still has at least one unfilled seat* in round k . Students thus automatically skip schools in the application process when applications to these schools are bound to be rejected, independent of priority.

Under the *Student-Proposing Deferred Acceptance (DA) mechanism* (Gale and Shapley, 1962; Abdulkadiroğlu and Sönmez, 2003), students also apply to schools in rounds, and priorities determine which applications are accepted. However, acceptances are *tentative* rather than permanent. This means that students can be rejected by a school where they were tentatively accepted in a previous round if other students apply in later rounds who have higher priority. Newly rejected students apply to their most preferred school that has not yet rejected them. The application process ends when no school receives any new applications.

Asymmetric Chinese Parallel mechanisms (ACPM) (Chen and Kesten, 2016a) are mechanisms that combine elements of BM and DA. They are parametrized by a vector of integers (e_0, e_1, \dots) with $e_k \geq 1$. Initially, all matches are tentative as under DA. However, unlike under DA, students who are rejected by their e_0^{th} choices pause in

¹⁵Sources: (Abdulkadiroğlu, Pathak and Roth, 2005; Abdulkadiroğlu et al., 2006; Kojima and Ünver, 2014; de Haan et al., 2015; Basteck, Huesmann and Nax, 2015; Chen and Pereyra, 2015; Chen and Kesten, 2016a), and Andrew Vanacore (April 16, 2012). *The Times-Picayune*. Retrieved March 22, 2017: www.nola.com

the application process. When all students are either matched tentatively or have been rejected by their e_0^{th} choices, all tentative matches are finalized. In the next phase, the rejected students continue to apply but pause again when they have been rejected by their $(e_0 + e_1)^{\text{th}}$ choices. The tentative matches are again finalized if all students are either tentatively accepted at some school or pausing. This process continues until all students are matched or have been rejected by all schools on their preference list. Observe that ACPM specifies a class of mechanisms. This class subsumes BM (for $e_k = 1$ for all k) and DA (for $e_0 \geq |S|$).

The *Top Trade Cycles (TTC) mechanism* (Abdulkadiroğlu and Sönmez, 2003) works by forming a directed graph: Each student points to her most-preferred school with unfilled seats, each school points to the student who has highest priority at that school, and the outside option points to all students who are pointing to it. In each step, a cycle of this graph is selected and implemented (i.e., each student in the cycle is permanently matched to the school to which she is pointing, and the respective seats and students are removed from the mechanism). Students and schools then adjust where they are pointing and the process repeats. The process ends when all students have been removed.

3. Setting the Stage: First-Choice Maximality and First-Choice Stability

As we have argued in the introduction, the number of students who are matched to their (reported) first choices receives a lot of attention. Following this observation, we formally define first-choice maximality.

Definition 1. Given a problem $(P, >)$, a matching μ is *first-choice maximal* if there exists no other matching ν such that $|I(\mu, 1, P)| < |I(\nu, 1, P)|$. A mechanism φ is *first-choice maximal (FCM)* if, for all problems $(P, >)$, the matching $\varphi(P, >)$ is first-choice maximal.¹⁶

Next, we define rank-based relaxations of stability.

¹⁶FCM is a strictly weaker requirement than the axiom that a mechanism *favors higher ranks* (Kojima and Ünver, 2014): Specifically, a mechanism is FCM if and only if it favors the *first* rank.

Definition 2. Given a problem (P, \succ) and an integer $k \in \{1, \dots, m\}$, a matching μ is k^{th} -choice stable if there exists no blocking pair $(i, s) \in I \times (S \cup \{\emptyset\})$ where s is the k^{th} choice of i , and μ is first-choice stable if this holds for $k = 1$. A mechanism is k^{th} -choice stable (first-choice stable (FCS)) if, for all problems (P, \succ) , the matching $\varphi(P, \succ)$ is k^{th} -choice stable (first-choice stable).

We observe that FCM and FCS are compatible: It is easy to see that BM satisfies both properties. On the other hand, FCM by itself is already a severe restriction: As the next example shows, it is incompatible with strategyproofness, with stability, or even with k^{th} -choice stability for $k \geq 2$.

Example 1. There are three students $I = \{1, 2, 3\}$, two schools $S = \{a, b\}$ with a single seat each. The preferences and priorities are

$$\begin{aligned} P_i \text{ for } i \in I: & \quad a P_i b P_i \emptyset, \\ \succ_s \text{ for } s \in S: & \quad 1 \succ_s 2 \succ_s 3. \end{aligned}$$

Let φ be an FCM mechanism. By feasibility, φ must leave one student unmatched. Without loss of generality, suppose that this is student 3 (i.e., $\varphi_3(P, \succ) = \emptyset$). If student 3 reports $b P'_3 a P'_3 \emptyset$ instead of reporting P_3 truthfully, then φ must match student 3 to school b . Since $\varphi_3((P'_3, P_{-3}), \succ) = b P_3 \emptyset = \varphi_3(P, \succ)$, P'_3 is a beneficial misreport for student 3. Thus, φ cannot be strategyproof.

Next, observe that at the problem $((P'_3, P_{-3}), \succ)$, there exists a unique stable matching μ where $\mu_1 = a$, $\mu_2 = b$, and $\mu_3 = \emptyset$. However, any FCM mechanism φ must match student 3 to school b , so φ cannot be stable. A straightforward extension illustrates that φ also violates k^{th} -choice stability for any $k \geq 2$: Consider a problem with k schools $S = \{s_1, \dots, s_k\}$ with a single seat each, and $k + 1$ students $I = \{1, 1', 2, \dots, k\}$. The preferences and priorities are

$$\begin{aligned} P_i \text{ for } i \in \{1, 1'\}: & \quad s_1 P_i s_2 P_i \dots P_i s_k P_i \emptyset, \\ P_i \text{ for } i \in \{2, \dots, k\}: & \quad s_i P_i \dots, \\ \succ_s \text{ for } s \in S: & \quad 1 \succ_s 1' \succ_s 2 \succ_s \dots \succ_s k. \end{aligned}$$

FCM implies that φ matches each student $i \in \{2, \dots, k\}$ to s_i . Thus, either student 1 or

student $1'$ is unmatched. Then that student forms a blocking pair with her k^{th} choice.

It follows from Example 1 that among all rank-based relaxations of stability, FCS is the only one compatible with FCM. For the remainder of this paper, we therefore focus on the class of mechanisms that are both FCM and FCS.

Definition 3. A mechanism is a *first-choice (FC) mechanism* if it is FCM and FCS.

For example, BM and ABM are both FCM and FCS and known to be Pareto efficient, but they violate strategyproofness and stability. ACPMs with $e_0 = 1$ are also FCM and FCS. Their first phase is the same as the first round of BM, and the matches made in this phase are not undone in subsequent phases. These mechanisms thus satisfy FCM and FCS (by Lemma 1 in Section 4). Moreover, as Chen and Kesten (2016a) pointed out, ACPMs are not Pareto efficient unless $e_k = 1$ for all k , in which case they are equivalent to BM, and they are not stable unless $e_0 > m$, in which case they are equivalent to DA. DA is stable and therefore FCS. However, FCM is incompatible with stability, so DA cannot be FCM. It is straightforward to see that TTC is neither FCM nor FCS, and it is known to be strategyproof and Pareto efficient but unstable.¹⁷

4. Incentives under FC Mechanisms

As Example 1 illustrates, no FC mechanism is strategyproof. However, some FC mechanisms may be *more manipulable* than others. Towards understanding these differences, we employ the concept of *comparing mechanisms by their vulnerability to manipulation* introduced by Pathak and Sönmez (2013), which we restate formally before presenting our results.

Definition 4. A mechanism φ is *manipulable at* $(P, >)$ if there exist a student $i \in I$ and a preference order P'_i such that $\varphi_i((P'_i, P_{-i}), >) P_i \varphi_i(P, >)$. For two mechanisms φ and ψ , we say that φ is *at least as manipulable as* ψ if, for all problems $(P, >)$, manipulability of ψ at $(P, >)$ implies manipulability of φ at $(P, >)$. φ is *more manipulable than* ψ if in addition there exists a problem where φ is manipulable but ψ is not.

¹⁷TTC violates FCM at the problem $((P'_3, P_{-3}), >)$ from Example 1, and it violates FCS at the same problem if the priorities of school a are changed to $3 >_a 2 >_a 1$.

In words, the comparison by vulnerability to manipulation classifies mechanisms by the sets of problems at which they are manipulable. If φ is at least as manipulable as ψ , then the set of problems at which φ is manipulable is a superset of the set of problems where ψ is manipulable, and φ is more manipulable if it is a strict superset.

Our next results, Theorem 1 and Proposition 1, reveal how FC mechanisms compare in terms of their vulnerability to manipulation. First, we show that all Pareto efficient FC mechanisms are manipulable at *exactly the same problems*. Second, we show that any Pareto *inefficient* FC mechanism is manipulable at a (possibly strict) *superset* of these problems. In this sense, the Pareto efficient FC mechanisms form a minimally manipulable subset within the class of FC mechanisms.

Theorem 1. *Let φ and ψ be two Pareto efficient FC mechanisms. Then φ is at least as manipulable as ψ and vice versa.*

Proof. We state two lemmas, which we also use in other proofs. Recall that $\text{choice}_{P_i}(k)$ denotes the k^{th} choice school according to P_i and that $I(\mu, k, P)$ denotes the set of students who are matched to their k^{th} choice according to P under μ .

Lemma 1. *φ is an FC mechanism if and only if $I(\varphi(P, >), 1, P) = I(\text{BM}(P, >), 1, P)$ for all problems $(P, >)$.*

Lemma 2. *Given an FC mechanism φ , a problem $(P, >)$, and a subset of students $A \subseteq I$, let (P'_A, P_{-A}) be a preference profile such that $\text{choice}_{P'_i}(1) = \varphi_i(P, >)$ for all $i \in A$. Then, $\varphi_i((P'_A, P_{-A}), >) = \varphi_i(P, >)$ for all $i \in A$.*

The proofs of these Lemmas are given in Appendixes A and B. Lemma 1 characterizes the set of FC mechanisms as those mechanisms that match exactly the same students to their reported first choices as BM; but otherwise they are free in how they match the remaining students. Lemma 2 shows that when some students change their reported preferences by ranking first the school to which they are matched under an FC mechanism, then this mechanism continues to match these students to the same schools.¹⁸

¹⁸This corresponds to a relaxed notion of Maskin monotonicity for FC mechanisms: The preference profile (P'_A, P_{-A}) is a monotonic transformation of P at $\varphi(P, >)$, and under any FC mechanism the matching of the students in A may not change. Observe that this property is independent of the *rank respecting invariance* property, a different relaxation of Maskin monotonicity that Kojima and Ünver (2014) used for their axiomatic characterization of BM.

Next, let φ be manipulable at $(P, >)$. Then $\varphi_i((P'_i, P_{-i}), >) P_i \varphi_i(P, >)$ for some student $i \in I$ and some preference order P'_i . Without loss of generality, let $s = \varphi_i((P'_i, P_{-i}), >)$ be the most preferred school according to P_i at which i can obtain a seat by misreporting (and possibly $s = \emptyset$). By Lemma 2, we can choose P'_i such that s is ranked first. Observe that s cannot be the first choice under P_i (otherwise, P'_i would not be a strictly beneficial deviation for i). There are two cases:

Case 1: $\varphi(P, >) = \psi(P, >)$. Then $I(\varphi((P'_i, P_{-i}), >), 1, P) = I(\psi((P'_i, P_{-i}), >), 1, P)$ by Lemma 1. Since $i \in I(\varphi((P'_i, P_{-i}), >), 1, P)$ and s is i 's first choice under P'_i , we have that $i \in I(\psi((P'_i, P_{-i}), >), 1, P)$ and $\psi_i((P'_i, P_{-i}), >) = s$. Hence, $\psi_i((P'_i, P_{-i}), >) P_i \psi_i(P, >)$, that is, i can manipulate ψ at $(P, >)$.

Case 2: $\varphi(P, >) \neq \psi(P, >)$. Since both $\varphi(P, >)$ and $\psi(P, >)$ are FCS and Pareto efficient, there exists a student $i' \in I$ such that $\varphi_{i'}(P, >) \neq \psi_{i'}(P, >)$ and $\varphi_{i'}(P, >) P_{i'} \psi_{i'}(P, >)$. Let $s' = \varphi_{i'}(P, >)$ and let $P'_{i'}$ be a preference order in which s' is ranked first. Then $\varphi_{i'}((P'_{i'}, P_{-i'}), >) = s'$ by Lemma 1, and $I(\varphi((P'_{i'}, P_{-i'}), >), 1, (P'_{i'}, P_{-i'})) = I(\psi((P'_{i'}, P_{-i'}), >), 1, (P'_{i'}, P_{-i'}))$ by Lemma 2. Hence, $\psi_{i'}((P'_{i'}, P_{-i'}), >) = s' P_{i'} \psi_{i'}(P, >)$, that is, i' can manipulate ψ at $(P, >)$.

In conclusion, manipulability of φ at $(P, >)$ implies manipulability of ψ at $(P, >)$. Symmetrically, it follows that if ψ is manipulable at some problem, then so is φ . \square

While Theorem 1 pertains to Pareto efficient FC mechanisms, the next proposition closes the remaining gap for Pareto *inefficient* FC mechanisms.

Proposition 1. *Let φ be an FC mechanism that violates Pareto efficiency. Then there exists a mechanism ψ with the following properties:*

1. ψ is a Pareto efficient FC mechanism,
2. ψ Pareto dominates φ ,
3. φ is at least as manipulable as ψ .

The formal proof is given in Appendix C. Here, we explain the intuition for the proof: If φ is not Pareto efficient, then there exists at least one problem $(P, >)$ where the matching $\varphi(P, >)$ is Pareto dominated by some other matching μ . We can define a new mechanism $\bar{\varphi}$ to be exactly the same as φ except at the problem $(P, >)$, where we set

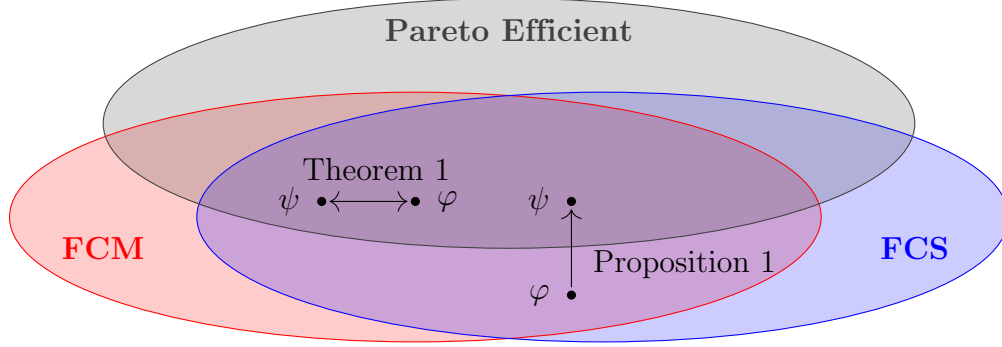


Figure 1: Illustration of Theorem 1 and Proposition 1.

$\bar{\varphi}(P, >) = \mu$. It is straightforward to show that $\bar{\varphi}$ is FCM and FCS and that it Pareto dominates φ . Careful inspection reveals that $\bar{\varphi}$ is manipulable only at problems where φ is manipulable (but possibly at fewer problems). Iterated application of this argument yields a Pareto efficient mechanism ψ that satisfies all three properties.

The Venn diagram in Figure 1 illustrates the results of Theorem 1 and Proposition 1. All Pareto efficient FC mechanisms (i.e., mechanisms in the intersection of all three areas) are equivalent when comparing them by their vulnerability to manipulation. The horizontal arrow on the left symbolizes this equivalence. Moreover, any Pareto *inefficient* FC mechanism (i.e., a mechanism in the intersection of the blue and red areas at the bottom) is at least as manipulable as some (and therefore any) Pareto efficient FC mechanism. The vertical arrow on the right symbolizes this relationship.

Theorem 1 and Proposition 1 have two implications for market design: First, recall that under BM, students apply to their k^{th} choices in the k^{th} round, even if these schools have no more unfilled seats. Students can therefore strategize by omitting full schools in their ranking. In contrast, under ABM, students automatically skip such schools and apply to their most preferred school with one or more unfilled seats. While students may still strategize in other ways, the above-mentioned manipulation becomes unnecessary. Thus, intuitively, we would expect ABM to have better incentive properties than BM. Surprisingly, this difference does not surface because both mechanism are equivalent in

this sense by Theorem 1.¹⁹

Second, ACPMs with $e_0 = 1$ are Pareto *inefficient* (unless $e_k = 1$ for all k , in which case they are equivalent to BM), but they are used in practice (e.g., for college admission in the Chinese Beijing, Gansu, and Shandong provinces). Proposition 1 shows that the choice of such mechanisms cannot be justified by the desiderata FCM, FCS, and “good” incentives alone, because administrators could choose a different FC mechanism that simultaneously yields unambiguous improvements in terms of student welfare and robustness to manipulation.

This concludes Step 1, our discussion of incentives under FC mechanisms.

5. Equilibrium under FC Mechanisms

Recall that no FC mechanism is strategyproof. We must therefore be concerned about the impact of strategic reporting by students. Following prior work, we identify the equilibrium outcomes of the preference revelation game induced by the mechanisms and use *Nash equilibrium under complete information* as the main solution concept (Ergin and Sönmez, 2006; Pathak and Sönmez, 2008; Haeringer and Klijn, 2009; Jaramillo, Kayi and Klijn, 2016).

Definition 5. Given a mechanism φ and a problem $(P, >)$, a preference profile P^* is a *Nash equilibrium (of φ at $(P, >)$)* if, for all students $i \in I$ and all preference orders $P'_i \neq P^*_i$, we have that $\varphi_i(P^*, >) R_i \varphi_i((P'_i, P^*_{-i}), >)$.

In words, no student can be matched to a school that she strictly prefers by unilaterally deviating from the equilibrium profile.²⁰

¹⁹The indistinguishability of BM and ABM is even more severe: It cannot be recovered via the *as-strongly-manipulable-as* relation, and for random priorities, even the weak distinction by vulnerability to manipulation is inconclusive (see Appendices G and H). More nuanced approaches are needed to obtain meaningful distinctions (see, e.g., (Dur, 2015; Mennle and Seuken, 2017b)).

²⁰The *preference revelation game induced by φ at $(P, >)$* is a simultaneous move game (N, O, τ) , where the students are the agents (i.e., $N = I$), the outcomes are the matchings that are possible under φ (i.e., $O = \{\varphi(\hat{P}, >) \mid \hat{P} \text{ preference profile}\}$), and for each agent i , the weak preference order τ_i over outcomes is induced by the respective student’s weak preference order over schools (i.e., for outcomes $x, y \in O$, $x \tau_i y$ if and only if $x_i R_i y_i$).

Observe that we simplify the formal exposition of Nash equilibrium in two ways: First, we use the term *Nash equilibrium of φ at $(P, >)$* to mean a Nash equilibrium of the induced preference revelation game. Second, we consider *preference profiles* instead of strategy profiles. This is without loss of generality because we consider only direct revelation mechanisms.

For the special case of BM, Ergin and Sönmez (2006) showed that the Nash equilibrium outcomes of BM correspond precisely to the matchings that are stable with respect to the true preferences and priorities.

Fact 1 (Theorem 1 of Ergin and Sönmez (2006)). *Given a problem (P, \succ) , a matching μ is stable if and only if there exists a Nash equilibrium P^* of BM at (P, \succ) with $\mu = BM(P^*, \succ)$.*

Our next theorem shows that the same result holds for all FC mechanisms.

Theorem 2. *Given a problem (P, \succ) and an FC mechanism φ , a matching μ is stable at (P, \succ) if and only if there exists a Nash equilibrium P^* of φ at (P, \succ) with $\mu = \varphi(P^*, \succ)$.*

Proof. Necessity. Let P^* be a Nash equilibrium. Assume towards contradiction that the matching $\varphi(P^*, \succ)$ is not stable wrt. (P, \succ) . Then there exist $i \in I$ and $s \in S \cup \{\emptyset\}$ such that $s P_i \varphi_i(P^*, \succ)$ and either $|\varphi_s(P^*, \succ)| < q_s$, or there exists $i' \in I$ such that $\varphi_{i'}(P^*, \succ) = s$ and $i \succ_s i'$. Let P'_i be a preference order with $\text{choice}_{P'_i}(1) = s$. When BM is applied to the preference profile (P'_i, P^*_{-i}) , i is matched to s in the first round. By Lemma 1 we get $\varphi_i((P'_i, P^*_{-i}), \succ) = \text{BM}_i((P'_i, P^*_{-i}), \succ) = s$. Therefore, P_i^* is not a best response to P^*_{-i} for i , a contradiction to the assumption that P^* is a Nash equilibrium.

Sufficiency. Let μ be stable wrt. (P, \succ) . Consider a preference profile P^* where each student i ranks μ_i first. Since μ is feasible, φ produces the matching μ in the first step when applied to (P^*, \succ) by Lemma 1. Moreover, P^* is a Nash equilibrium of φ : Assume towards contradiction that some student i can improve her match by deviating from P^* (i.e., get matched to a school s). By Lemma 2, she can do so by ranking s first. Then i either displaces another student with lower priority at s , or s has unfilled seats under μ . In both cases, the pair (i, s) blocks μ , a contradiction to stability of μ . \square

Remark 1. Theorem 2 shows that stability (with respect to the true preferences and priorities) is the characterizing feature of the Nash equilibrium outcomes of FC mechanisms. This generalizes Fact 1 from BM to all FC mechanisms. Ergin and Sönmez (2006) also showed that all monotonic rank-priority mechanisms have this property. Relatedly, within the class of rank-priority mechanisms, Jaramillo, Kayi and Klijn (2016) characterized those that Nash implement the set of stable matchings. Our Theorem 2 is independent of both of these results because the set of FC mechanisms is neither a subset

nor a superset of the rank-priority mechanisms (see Appendix D). An interesting subject for future research would be a characterization of all direct-revelation mechanisms that Nash implement the set of stable matchings that unifies these results.

DA is known to implement the *student-optimal* stable matchings in weakly-dominant strategies.²¹ In contrast, by Theorem 2, FC mechanisms implement *all* stable matchings in Nash equilibrium, not just the student-optimal ones. Thus, DA produces (weakly) Pareto dominant matchings, and it does so subject to a more robust solution concept. This provides a partial answer to our main research question whether FC mechanisms actually achieve the desiderata FCM and FCS with respect to the true preferences: They achieve FCS because they lead to stable matchings (in equilibrium when all students strategize), but they fail to achieve FCM (which is incompatible with stability).

6. Equilibrium under FC Mechanisms When Some Students are Sincere

In practice, students exhibit varying levels of strategic sophistication. Some students may report their preferences truthfully, e.g., because they lack the information that is necessary to determine beneficial misreports. Following Pathak and Sönmez (2008), we consider *mixed problems* with two groups of students: *Sincere students* simply report their preferences truthfully, independent of incentives, while *sophisticated students* recognize the strategic nature of the preference revelation game and play best responses. In this section, we identify the Nash equilibrium outcomes of FC mechanisms (Section 6.1) and study the implications of strategic reporting on student welfare (Section 6.2).

6.1. Identification of Equilibrium Outcomes

We first extend our definition of Nash equilibrium to mixed problems with both sophisticated and sincere students.

Definition 6. Given a problem (P, \succ) , a set of sophisticated students $A \subseteq I$, and a mechanism φ , a preference profile P^* is an *A-Nash equilibrium* (of φ at (P, \succ)) if $P_i^* = P_i$

²¹A stable matching is *student-optimal* if all students prefer it to any other stable matching. These matchings are unique for the problems we consider (Roth, 1982).

for all sincere students $i \in I \setminus A$ and $\varphi_i(P^*, >) R_i \varphi_i((P'_i, P_{-i}^*), >)$ for all sophisticated students $i \in A$ and all preference orders P'_i .

In words, in an A -Nash equilibrium, all sincere students report truthfully and no sophisticated student can benefit by unilaterally deviating from the equilibrium profile. Pathak and Sönmez (2008) showed that the A -Nash equilibrium outcomes of BM correspond to the matchings that are stable with respect to the true preferences and an *augmented priority profile*. These augmented priorities capture three intuitive aspects of BM: First, students who have been accepted keep their seats, even if they have lower priority than another student who applies in a later round of BM. The fact that a student ranks a school higher thus overrules the fact that her priority at that school may be lower. Augmented priorities capture this by giving higher priority to students who rank a school higher. Second, while sincere students report their preferences truthfully, sophisticated students can misreport their preferences. In particular, they can rank any school first. Augmented priorities reflect this advantage by treating sophisticated students as if they ranked every school first. Third, BM breaks ties according to the original priority profile, and augmented priorities do the same. The next definition formalizes these aspects (where I_k^s denotes the set of students who rank s in k^{th} position according to P).

Definition 7. Given a problem $(P, >)$ and a set of sophisticated students $A \subseteq I$, for each school $s \in S$, the *augmented priority order* $\widehat{>}_s$ is constructed as follows:

- $i \widehat{>}_s j$ if $i \in A \cup I_1^s$ and $j \in I_2^s \setminus A$
- $i \widehat{>}_s j$ if $i \in I_k^s \setminus A$ and $j \in I_{k+1}^s \setminus A$ for any $k \geq 2$
- $i \widehat{>}_s j$ if $i >_s j$ and either $i, j \in A \cup I_1^s$ or $i, j \in I_k^s \setminus A$ for any $k \geq 2$
- All undefined priorities are implied by transitivity

A matching μ is *augmented stable* if it is stable with respect to the problem $(P, \widehat{>}_s)$.

With the notions of A -Nash equilibrium and augmented stability, we can now formally restate the main result of Pathak and Sönmez (2008).

Fact 2 (Proposition 1 of Pathak and Sönmez (2008)). *Given a problem $(P, >)$ and a set of sophisticated students $A \subseteq I$, a matching μ is augmented stable if and only if there exists an A -Nash equilibrium P^* of BM at $(P, >)$ with $\mu = BM(P^*, >)$.*

In words, Fact 2 means that the set of A -Nash equilibrium outcomes of BM corresponds precisely to the set of augmented-stable matchings. Our next result extends this characterization to the entire class of FC mechanisms, albeit with one limitation: We identify the equilibrium outcomes of FC mechanisms only up to equivalence from the perspective of the sophisticated students. Formally, we say that two matchings μ and ν are A -equivalent if $\mu_i = \nu_i$ for all sophisticated students $i \in A$, denoted $\mu =_A \nu$.

Theorem 3. *Given a problem (P, \succ) , a set of sophisticated students $A \subseteq I$, and an FC mechanism φ , a matching μ is A -equivalent to some augmented-stable matching if and only if there exists an A -Nash equilibrium P^* of φ at (P, \succ) with $\mu =_A \varphi(P^*, \succ)$.*

A formal proof is given in Appendix E. In words, Theorem 3 shows that augmented stability describes the equilibrium outcomes of FC mechanisms in mixed problems from the perspective of the sophisticated students. To see why the limitation to sophisticated students is needed, recall that FC mechanisms are restricted in how they handle first choices but they are free in how they to handle other choices (Lemma 1). In equilibrium, this freedom only affects the matching of sincere students because sophisticated students can always get matched to their equilibrium school by ranking it first (Lemma 2).

If all students are sincere, then the equilibrium outcomes of FC mechanisms trivially satisfy FCM and FCS with respect to the true preferences. With sophisticated students, this may no longer be true. Example 2 illustrates that FCM can be violated even if there is just one sophisticated student.

Example 2. There are three students $I = \{1, 2, 3\}$ and two schools $S = \{a, b\}$ with one seat each, and only student 2 is sophisticated (i.e., $A = \{2\}$). The preferences and priorities are

$$\begin{aligned}
 P_i \text{ for } i \in \{1, 2\} : & \quad a P_i b P_i \emptyset, \\
 P_3 : & \quad b P_3 a P_3 \emptyset, \\
 \succ_a : & \quad 3 \succ_a 1 \succ_a 2, \\
 \succ_b : & \quad 1 \succ_b 2 \succ_b 3.
 \end{aligned}$$

Notice that at most two students can be matched to their true first choices. But with $A = \{2\}$, the unique augmented-stable matching is μ with $\mu_1 = a$, $\mu_2 = b$, and $\mu_3 = \emptyset$.

By Thm. 3, student 2 is matched to school b in any $\{2\}$ -Nash equilibrium outcome of any FC mechanism. Thus, at most one student is matched to her true first choice which violates FCM with respect to the true preference.

While FC mechanisms can fail to match a maximal number of students to their true first choices in equilibrium, the next corollary shows that all equilibrium outcomes are FCS with respect to the true preferences.

Corollary 1. *Given a problem $(P, >)$, a set of sophisticated students $A \subseteq I$, and an FC mechanism φ , let P^* be an A -Nash equilibrium of φ at $(P, >)$. Then the matching $\varphi(P^*, >)$ is FCS with respect to $(P, >)$.*

Proof. Assume towards contradiction that $\varphi(P^*, >)$ is not FCS with respect to the true problem $(P, >)$. Then there exists a pair $(i, s) \in I \times (S \cup \{\emptyset\})$ that blocks $\varphi(P^*, >)$ with $s = \text{choice}_{P_i}(1)$. If $i \in A$, then i could get matched to s by ranking it first, a contradiction to the assumption that P^* is an A -Nash equilibrium. If $i \notin A$, then i already ranks s first. However, first choices are matched respecting priorities under any FC mechanism (by Lemma 1). Thus, all students matched to s have priority over i at s , again a contradiction. \square

In summary, we have found that the equilibrium outcomes of FC mechanisms in mixed problems always satisfy FCS with respect to the true preferences, but they may violate FCM with respect to the true preferences.

6.2. Student Welfare in Equilibrium

So far, we have analyzed student welfare under FC mechanisms in two extreme cases: When *all students are sophisticated*, then FC mechanisms match at most as many students to their true first choices in equilibrium as DA;²² when *all students are sincere*, then FC mechanisms are trivially FCM. But what happens in intermediate cases, when some but not all students are sincere?

Towards this question, we consider specific equilibrium outcomes that are unanimously preferred by all sophisticated students. To make this formal, fix an FC mechanism φ , a set of sophisticated students $A \subseteq I$, and a problem $(P, >)$. By the theory of stable matchings

²²Assuming truthful reporting, which is a weakly dominant strategy under DA.

(Roth and Sotomayor, 1990), there exists a *student-optimal augmented-stable matching*, say $\hat{\mu}$. By Theorem 3, there exists an A -Nash equilibrium P^* (of φ at $(P, >)$) such that the outcome $\varphi(P^*, >)$ is A -equivalent to $\hat{\mu}$, and all sophisticated students (weakly) prefer the matching $\varphi(P^*, >)$ to all other A -Nash equilibrium matchings. Because of this unanimous preference by the sophisticated students for $\varphi(P^*, >)$, we now focus on these *A -optimal Nash equilibrium matchings*.

For our analysis of student welfare in intermediate cases (i.e., with some sincere students), we identify how an individual student's behavior impacts student welfare by comparing the respective A -optimal Nash equilibrium matchings. Before we formalize the comparison, we discuss an example to build intuition.

Example 3. There are six students $I = \{1, \dots, 6\}$ and four schools $S = \{a, b, c, d\}$ with one seat each. The preferences and priorities are

$$\begin{aligned}
 P_i \text{ for } i \in \{1, 2, 3\} : & \quad a P_i b P_i c P_i \emptyset, \\
 P_4 : & \quad a P_4 b P_4 c P_4 d P_4 \emptyset, \\
 P_5 : & \quad a P_5 b P_5 d P_5 \emptyset, \\
 P_6 : & \quad b P_6 c P_6 \emptyset, \\
 >_s \text{ for } s \in S : & \quad 1 >_s \dots >_s 6.
 \end{aligned}$$

Suppose that the FC mechanism ABM is used to match students to schools. For the sets of sophisticated students $A = \{6\}$ and $A' = \{3, 6\}$, the equilibrium outcomes in each case are unique and given in the following table.

Sophisticated students	Unique equilibrium matchings under ABM					
$A = \{6\}$	$\mu_1 = a,$	$\mu_2 = c,$	$\mu_3 = \emptyset,$	$\mu_4 = \emptyset,$	$\mu_5 = d,$	$\mu_6 = b$
$A' = \{3, 6\}$	$\nu_1 = a,$	$\nu_2 = \emptyset,$	$\nu_3 = b,$	$\nu_4 = d,$	$\nu_5 = \emptyset,$	$\nu_6 = c$

This example illustrates four aspects of the relationship between μ and ν .

1. The sophisticated student 6 prefers the outcome when student 3 is sincere because she strictly prefers $\mu_6 = b$ to $\nu_6 = c$.
2. Student 3, who is either sophisticated or sincere, prefers the outcome when she is sophisticated because she strictly prefers $\nu_3 = b$ to $\mu_3 = \emptyset$.

3. Whether student 3 is sophisticated or sincere has ambiguous effects for the sincere students: Students 2 and 5 prefer μ , student 4 prefers ν , and student 1 is indifferent between both matchings.
4. When more students are sincere, more student are matched to their true first choices: If students 3 and 6 are both sophisticated, then only student 1 is matched to her true first choice. However, if only student 6 is sophisticated, then students 1 and 6 are both matched to their true first choices.

The following theorem establishes that our observations about the relationship between μ and ν in Example 3 hold in general for all FC mechanisms and all mixed problems.

Theorem 4. *Given a problem $(P, >)$, an FC mechanism φ , sophisticated students $A \subseteq I$, and a sincere student $i' \notin A$, let μ be an A -optimal Nash equilibrium matching and let ν be an A' -optimal Nash equilibrium matching where $A' = A \cup \{i'\}$. Then:*

1. *For all $i \in A$: $\mu_i R_i \nu_i$, i.e., all sophisticated students prefer μ to ν .*
2. *For i' : $\nu_{i'} R_{i'} \mu_{i'}$, i.e., i' prefers ν to μ .*
3. *For $i \notin A \cup \{i'\}$: $\nu_i P_i \mu_i$, $\mu_i P_i \nu_i$, or $\nu_i = \mu_i$ are all possible, i.e., for all sincere students except i' the impact of the behavior of i' is ambiguous.*
4. *For all $i \in I$: If $\nu_i = \text{choice}_{P_i}(1)$, then $\mu_i = \nu_i$, i.e., any student who is matched to her true first choice under ν is also matched to her true first choice under μ .*

A formal proof is given in Appendix F. Our result generalizes the corresponding Proposition 4 of Pathak and Sönmez (2008) in two ways: On the one hand, we extend its scope from BM to all FC mechanisms. On the other hand, beyond the impact for sophisticated students and i' , we also identify the impact for sincere students and for those students who are matched to their true first choices.

Remark 2. It is intuitive that strategizing improves the outcome for i' and harms the other sophisticated students. Thus, Statements 1 and 2 in Theorem 4 may appear trivial. However, this intuition is deceptive: By misreporting her preferences, i' changes the outcome for herself and others. But more importantly, she also changes the game for the other sophisticated students. They in turn respond by changing their own preference reports, which could deteriorate the outcome for i' in general. Showing that

FC mechanisms do not induce such dynamics is precisely the contribution of Theorem 4, and the proof critically relies on the two properties FCM and FCS.

For market designers, Theorem 4 answers the question whether FC mechanisms can bring us closer to achieving first-choice maximality with respect to the true preferences. In the respective optimal equilibrium outcomes, the number of students who are matched to their true first choices is lowest if all students are sophisticated, increases when more students are sincere, and peaks when all students are sincere. This provides a potential justification for the use of FC mechanisms in practice: If administrators wish to match many students to their true first choices, if they care about a minimal fairness guarantee in terms of FCS, and if they expect some share of the students to report their preferences truthfully, then FC mechanisms may be an attractive design alternative.

Remark 3. This insight comes with a caveat: Nash equilibrium is a weaker solution concept than dominant-strategy equilibrium. Moreover, robust predictions of equilibrium play are not the only reason for the appeal of strategyproofness. By simplifying participation for students, strategyproofness also alleviates the cognitive cost that sophisticated students incur when they strategize (Azevedo and Budish, 2015). On top of that, strategyproofness yields another form of fairness because it levels the playing field between students with different levels of strategic sophistication (Pathak and Sönmez, 2008). Our finding that the use of non-strategyproof FC mechanisms may be justified by the fact that they may match more students to their true first choices is agnostic to this broader role of strategyproofness as a desideratum for market design.

7. A Characterization of the Boston Mechanism

In Theorems 1 and 2, we have shown that any FC mechanism has the same level of vulnerability to manipulation and the same set of Nash equilibrium outcomes. Moreover, in Lemma 1 we have shown that any FC mechanism assigns the same set of students to their first choices in any problem. These results imply that no FC mechanism performs better than any other FC mechanism regarding vulnerability to manipulation and regarding the structure of the Nash equilibrium outcomes. In this section, we go one step further and provide a characterization of the most widely used FC mechanism:

BM. Here, we relax our restriction on the school quotas and we allow $q_s = 0$ for some school $s \in S$.

We first provide the new relaxed notions that we will use in our characterization. Given a problem $(I, S, q, P, >)$, choice k is the *best possible choice* for a student if there does not exist a matching μ in which the student is assigned to her \bar{k} 'th choice where $\bar{k} < k$. When $q_s > 0$ for all $s \in S$, the best possible choice is the first choice. However, when $q_s = 0$ for some school s and school s is the top choice of all students, then the best possible choice is the first choice.

Next, we provide more generalized versions of first-choice-maximality and first-choice-stability. We say a mechanism is *best-choice-maximal* if it maximizes the number of students assigned to the best possible choice in any problem $(I, S, q, P, >)$. Similarly, a mechanism is *best-choice-stable* if it is individually rational, non-wasteful, and it respects each student's priority at her best possible choice in any problem $(I, S, q, P, >)$.

If we restrict our domain to the problems in which $q_s > 0$ for each school $s \in S$, then any FC mechanism is best-choice-maximal and best-choice-stable. When $q_s = 0$ for some school s , then not all FC mechanisms are best-choice-maximal and best-choice-stable. However, one can easily show, using its definition that BM is both best-choice-maximal and best-choice-stable.

In addition to these two desiderata, we follow Kojima and Ünver (2014) and introduce a third desideratum, *consistency*, to obtain our characterization result. We say that a mechanism is *consistent* if, for any problem $(I, S, q, P, >)$, when we remove a subset of students with their assignments and update the remaining seats of the schools, then the remaining students are assigned to the same schools as in the initial problem. Kojima and Ünver (2014) have shown that BM is consistent.

We are now ready to provide our characterization.

Theorem 5. *BM is the unique best-choice-stable, best-choice-maximal, and consistent mechanism.*

Proof. Best-choice-stability and best-choice-maximality of BM follow from its definition. Moreover, BM is consistent (Kojima and Ünver, 2014).

On the contrary, suppose Ψ is best-choice-stable, best-choice-maximal, and consistent, but $\Psi(I, S, P, >, q) \neq BM(I, S, q, P, >)$ for some problem $(I, S, q, P, >)$. The steps of Lemma 1 imply that any best-choice-maximal and best-choice-stable mechanism assigns

the same set of students to their best possible choices for any problem. In particular, if choice $k > 1$ is the best possible choice under problem $(I, S, q, P, >)$, then we can update the preferences of all students in I by truncating the first $k - 1$ choices and apply Lemma 1 to show that any best-choice-maximal and best-choice-stable mechanism assigns the same set of students to their k 'th choices in problem $(I, S, q, P, >)$. If we remove these students with their assignments and update the number of remaining seats in each school and get the set of students I' and quota profile q' , then by consistency of Ψ and BM, both of these mechanisms assign the remaining students to the same schools under both $(I, S, q, P, >)$ and $(I', S, q', P, >)$. Then, by using best-choice-maximality and best-choice-stability of both mechanisms, we can show that in problem $(I', S, q', P, >)$, BM and Ψ assign the same set of students to their best possible choice. As we follow the same procedure, we can show that $\Psi(I, S, q, P, >) = BM(I, S, q, P, >)$, a contradiction. \square

8. Conclusion

In school choice, FC mechanisms arise naturally, and many school choice mechanisms used in practice belong to this class. Understanding this class is therefore important for researchers and practitioners alike. The class of FC mechanisms is fairly large because FC mechanisms are only restricted in how they handle (reported) first choices. Nonetheless, we were able to analyze the incentives and equilibria of all mechanisms in this class (including BM, ABM, and ACPMs with $e_0 = 1$) purely based on the defining properties FCM and FCS. This suggests FCM and FCS as useful concepts for the analysis of school choice markets.

Our findings contribute to an ongoing debate about the respective merits and shortcomings of BM and DA. On the one hand, DA has the obvious advantage of being strategyproof (Abdulkadiroğlu and Sönmez, 2003), and its outcomes weakly Pareto dominate the Nash equilibrium outcomes of BM in settings with strict priorities and when all students strategize optimally (Ergin and Sönmez, 2006). However, the comparison becomes less clear when either of the assumptions are relaxed. When priorities are weak, the equilibrium outcomes of BM can dominate those of DA from an *ex-ante* perspective (Miralles, 2008; Erdil and Ergin, 2008; Abdulkadiroğlu, Che and Yasuda, 2011). Furthermore, if all students are sincere, BM *rank dominates* DA whenever they are comparable by rank dominance (Harless, 2015; Mennle and Seuken, 2017b), and BM

satisfies the welfare property of *favoring higher ranks* (Kojima and Ünver, 2014). Our present paper adds to these insights: If first choices matter and if a share of the students is expected to be sincere, then any FC mechanism (including BM) may yield more appealing matchings in equilibrium than DA. Our characterization result further shows that BM takes an elevated position among the FC mechanisms. Nevertheless, we refrain from recommending any specific mechanism in general; instead, our findings highlight the implicit trade-offs one must make when choosing between any of the mechanisms we have studied.

The prevalence of non-strategyproof mechanisms in practice has inspired new ways of thinking about incentives. For example, the concept for comparing mechanisms by their vulnerability to manipulation, put forward by (Pathak and Sönmez, 2013), was instrumental in identifying a trend towards better incentive properties in the USA, UK, and Ghana. Interestingly, our results yield a criticism of this concept because it does not identify the intuitive differences between BM and ABM. To distinguish these two mechanisms by their incentive properties, more nuanced approaches are necessary (e.g., (Dur, 2015) for the case when schools find some students unacceptable or (Mennle and Seuken, 2017b) when priorities are random).

Finally, our results give rise to promising directions for future research: Whether or not FC mechanisms can outperform DA hinges on the share of sophisticated students and their ability to coordinate on equilibrium. Some recent work has already considered field data to study student behavior and outcomes in school choice markets (Calsamiglia and Güell, 2014; de Haan et al., 2015; Dur, Hammond and Morrill, 2016). However, further research focusing specifically on sincerity and coordination would constitute an important contribution to the debate about school choice mechanisms, and we are aware of ongoing efforts in this direction.

References

- Abdulkadiroğlu, Atila, Alvin Roth, Parag Pathak, and Tayfun Sönmez.** 2006. “Changing the Boston School Choice Mechanism: Strategy-proofness as Equal Access.” Working Paper.
- Abdulkadiroğlu, Atila, and Tayfun Sönmez.** 2003. “School Choice: A Mechanism Design Approach.” *American Economic Review*, 93(3): 729--747.
- Abdulkadiroğlu, Atila, Parag Pathak, and Alvin Roth.** 2005. “The New York City High School Match.” *American Economic Review*, 95(2): 364--367.
- Abdulkadiroğlu, Atila, Yeon-Koo Che, and Yosuke Yasuda.** 2011. “Resolving Conflicting Preferences in School Choice: The Boston Mechanism Reconsidered.” *American Economic Review*, 101(1): 399--410.
- Alcalde, Jose.** 1996. “Implementation of Stable Solutions to Marriage Problems.” *Journal of Economic Theory*, 69(1): 240--254.
- Azevedo, Eduardo, and Eric Budish.** 2015. “Strategy-proofness in the Large.” Working Paper.
- Balinski, Michel, and Tayfun Sönmez.** 1999. “A Tale of Two Mechanisms: Student Placement.” *Journal of Economic Theory*, 84(1): 73--94.
- Basteck, Christian, Katharina Huesmann, and Heinrich Nax.** 2015. “Matching Practices for Secondary Schools - Germany.” matching-in-practice.eu, accessed May 16, 2015.
- Calsamiglia, Caterina, and Maia Güell.** 2014. “The Illusion of School Choice: Empirical Evidence from Barcelona.” Working Paper.
- Chen, Li, and Juan Sebastián Pereyra.** 2015. “Self-selection in School Choice.” Working Paper.
- Chen, Yan, and Onur Kesten.** 2016*a*. “Chinese College Admissions and School Choice Reforms: A Theoretical Analysis.” Working Paper.

- Chen, Yan, and Onur Kesten.** 2016b. “Chinese College Admissions and School Choice Reforms: An Experimental Study.” Working Paper.
- Chen, Yan, and Tayfun Sönmez.** 2006. “School Choice: an Experimental Study.” *Journal of Economic Theory*, 127(1): 202--231.
- de Haan, Monique, Pieter Gautier, Hessel Oosterbeek, and Bas van der Klaauw.** 2015. “The Performance of School Assignment Mechanisms in Practice.” Working Paper.
- Dur, Umut.** 2015. “The Modified Boston Mechanism.” Working Paper.
- Dur, Umut, Robert Hammond, and Thayer Morrill.** 2016. “The Secure Boston Mechanism: Theory and Experiments.” Working Paper.
- Erdil, Aytek, and Haluk Ergin.** 2008. “What’s the Matter with Tie-Breaking? Improving Efficiency in School Choice.” *American Economic Review*, 98(3): 669--89.
- Ergin, Haluk, and Tayfun Sönmez.** 2006. “Games of School Choice Under the Boston Mechanism.” *Journal of Public Economics*, 90(1-2): 215--237.
- Featherstone, Clayton.** 2015. “Rank Efficiency : Investigating a Widespread Ordinal Welfare Criterion.” Working Paper.
- Gale, David, and Lloyd Shapley.** 1962. “College Admissions and the Stability of Marriage.” *The American Mathematical Monthly*, 69(1): 9--15.
- Haeringer, Guillaume, and Flip Klijn.** 2009. “Constrained School Choice.” *Journal of Economic Theory*, 144(5): 1921--1947.
- Harless, Patrick.** 2015. “A School Choice Compromise: Between Immediate and Deferred Acceptance.” Working Paper.
- Jaramillo, Paula, Çağatay Kayi, and Flip Klijn.** 2016. “School Choice: Nash Implementation of Stable Matchings Through Rank-Priority Mechanisms.” Working Paper.
- Kojima, Fuhito, and Utku Ünver.** 2014. “The Boston School-Choice Mechanism: an Axiomatic Approach.” *Economic Theory*, 55(3): 515--544.

- Mennle, Timo, and Sven Seuken.** 2017*a*. “Partial Strategyproofness: Relaxing Strategyproofness for the Random Assignment Problem.” Working Paper.
- Mennle, Timo, and Sven Seuken.** 2017*b*. “Trade-offs in School Choice: Comparing Deferred Acceptance, the Naïve and the Adaptive Boston Mechanism.” Working Paper.
- Miralles, Antonio.** 2008. “School Choice: the Case for the Boston Mechanism.” Working Paper.
- Pais, Joana, and Agnes Pinter.** 2008. “School Choice and Information: an Experimental Study on Matching Mechanisms.” *Games and Economic Behavior*, 64: 303--328.
- Pathak, Parag, and Tayfun Sönmez.** 2008. “Leveling the Playing Field: Sincere and Sophisticated Players in the Boston Mechanism.” *American Economic Review*, 98(4): 1636--52.
- Pathak, Parag, and Tayfun Sönmez.** 2013. “School Admissions Reform in Chicago and England: Comparing Mechanisms by Their Vulnerability to Manipulation.” *American Economic Review*, 103(1): 80--106.
- Roth, Alvin.** 1982. “The Economics of Matching: Stability and Incentives.” *Mathematical Operations Research*, 7: 617--628.
- Roth, Alvin, and Marilda Sotomayor.** 1990. *Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis*. Cambridge University Press.
- Westkamp, Alexander.** 2013. “An Analysis of the German University Admissions System.” *Economic Theory*, 53(3): 561--589.

Appendix

A. Proof of Lemma 1

Sufficiency. Since BM and φ are FCM, $|I(\varphi(P, \succ), 1, P)| = |I(\text{BM}(P, \succ), 1, P)|$. Thus, it suffices to show that for all students $i \in I$, if $i \notin I(\text{BM}(P, \succ), 1, P)$, then $i \notin I(\varphi(P, \succ), 1, P)$. Assume towards contradiction that there exists some $i \in I$ such that $i \notin I(\text{BM}(P, \succ), 1, P)$ but $i \in I(\varphi(P, \succ), 1, P)$. Let $s = \text{choice}_{P_i}(1)$. Consider the first round of BM. If i is not matched to s , there exist at least q_s students whose first choice is s and who has higher priority than i at s . If i is matched to s under $\varphi(P, \succ)$, then at least one of these students is not matched to s under $\varphi(P, \succ)$, a contradiction to first-choice stability of φ .

Necessity. $I(\varphi(P, \succ), 1, P) = I(\text{BM}(P, \succ), 1, P)$ holds for all problems (P, \succ) . Since BM is FCM, so is φ . Now, assume towards contradiction that φ is not FCM. Thus, there exist a problem (P, \succ) and a pair $(i, s) \in I \times (S \cup \{\emptyset\})$ that blocks the matching $\varphi(P, \succ)$ and where s is i 's first choice. By assumption, the same set of students (excluding i) is matched to s under $\text{BM}(P, \succ)$. But then the pair (i, s) blocks the matching $\text{BM}(P, \succ)$ as well, a contradiction to the fact that BM is FCS.

B. Proof of Lemma 2

First, assume that $\text{choice}_{P_i}(1) \neq \varphi_i(P, \succ)$ for all $i \in A$. By Lemma 1, $I(\varphi(P, \succ), 1, P) = I(\text{BM}(P, \succ), 1, P)$. Thus, all schools to which some student in A is matched under $\varphi(P, \succ)$ are not exhausted in the first round of BM. Therefore, $\text{BM}_i((P'_A, P_{-A}), \succ) = \varphi_i(P, \succ)$ for all $i \in A$. However, $I(\text{BM}((P'_A, P_{-A}), \succ), 1, (P'_A, P_{-A})) = I(\varphi((P'_A, P_{-A}), \succ), 1, (P'_A, P_{-A}))$ by Lemma 1, so that $\varphi_i(P, \succ) = \text{BM}_i((P'_A, P_{-A}), \succ) = \varphi_i((P'_A, P_{-A}), \succ)$ for all $i \in A$.

Second, suppose that A also contains students who are matched to their first choices under $\varphi(P, \succ)$, and let $B \subseteq A$ be the set of these students. In this case, apply the above argument to $A \setminus B$. Next, observe that for any $i \in B$, if the preference order of i is changed to some P'_i with $\text{choice}_{P'_i}(1) = \text{choice}_{P_i}(1)$, then the matching of first choices under BM does not change. In particular, i is still matched to $\varphi_i(P, \succ)$. Thus, $\varphi_i(P, \succ) = \text{BM}_i((P'_A, P_{-A}), \succ) = \varphi_i((P'_A, P_{-A}), \succ)$ for all $i \in A$.

C. Proof of Proposition 1

Suppose that $\varphi(P, >)$ is strictly Pareto dominated by some other matching μ at $(P, >)$. Define a mechanism $\bar{\varphi}$ to be the same as φ , except that $\bar{\varphi}(P, >) = \mu$. First, observe that $\bar{\varphi}$ must be FCM because it Pareto improves over an FCM mechanism and therefore cannot match strictly fewer first choices. However, since φ is FCM, $\bar{\varphi}$ also cannot match strictly more first choices. Thus, $\bar{\varphi}$ matches exactly the same first choices as φ , and therefore, it must be FCS.

Next, we verify that φ is as manipulable as $\bar{\varphi}$. Since φ and $\bar{\varphi}$ select the same matching for all problems except $(P, >)$, we only need to consider two cases:

Case A. $(P, >)$ is the true problem and some student $i \in I$ is considering some misreport $P'_i \neq P_i$,

Case B. $((P'_i, P_{-i}), >)$ with $P'_i \neq P_i$ is the true problem and student $i \in I$ is considering the particular misreport P_i .

In Case A, suppose that i can manipulate $\bar{\varphi}$ by reporting P'_i instead of reporting P_i truthfully in the problem $(P, >)$, i.e., $\bar{\varphi}_i((P'_i, P_{-i}), >) P_i \bar{\varphi}_i(P, >)$. Since μ Pareto dominates $\varphi(P, >)$, we get

$$\varphi((P'_i, P_{-i}), >) = \bar{\varphi}_i((P'_i, P_{-i}), >) P_i \bar{\varphi}_i(P, >) = \mu R_i \varphi_i(P, >), \quad (1)$$

or equivalently, i can also manipulate φ by reporting P'_i in the problem $(P, >)$.

In Case B, suppose that i can manipulate $\bar{\varphi}$ by reporting P_i instead of reporting P'_i truthfully in the problem $((P'_i, P_{-i}), >)$, i.e., $\mu_i = \bar{\varphi}_i(P, >) P'_i \bar{\varphi}_i((P'_i, P_{-i}), >)$. Let $P''_i \neq P_i$ be a preference order in which μ_i is ranked in first position. i can obtain μ_i by reporting P_i , and since $\bar{\varphi}$ is an FC mechanism, i can also obtain μ_i by ranking it first (in particular by reporting P''_i) by Lemma 2. $\mu_i = \bar{\varphi}_i((P''_i, P_{-i}), >) = \varphi_i((P''_i, P_{-i}), >)$ holds by construction, which implies that i can obtain μ_i by reporting P''_i under φ in the problem $(P, >)$. Thus, φ must be manipulable in $(P, >)$.

Finally, we can construct a mechanism ψ that is FCM, FCS, and Pareto efficient by iteratively applying the above construction. Since I , S , and q are held fixed, there are only finitely many possible problems and matchings, so that the construction ends after finitely many steps.

D. FC Mechanisms and Rank-Priority Mechanisms

BM is both an FC mechanism and a monotonic rank-priority mechanism. Thus, our Theorem 2 as well as Theorem 4 of Ergin and Sönmez (2006) and Theorem 1 of Jaramillo, Kayi and Klijn (2016) apply to BM. To show that our result is independent of the two other results, we construct two mechanisms: First, an FC mechanism that is not a rank-priority mechanism, and second, a monotonic rank-priority mechanism that is not an FC mechanism.

Example 4. ABM is an FC mechanism that is not a rank-priority mechanism. To see this, consider four students $I = \{1, 2, 3, 4\}$ and three schools $S = \{a, b, c\}$. The preferences and priorities are

$$\begin{aligned} P_i \text{ for } i \in \{1, 2, 3\} : & \quad a P_i b P_i c P_i \emptyset, \\ P_4 : & \quad a P_4 c P_4 \emptyset, \\ >_s \text{ for } s \in S : & \quad 1 >_s 2 >_s 3 >_s 4. \end{aligned}$$

Observe that $\text{ABM}_3(P, >) = \emptyset$ and $\text{ABM}_4(P, >) = c$. In the notation of Ergin and Sönmez (2006), this implies $\pi(2, 4) < \pi(3, 3)$. Next, consider the preference profile $P' = (P_1, P'_2, P_3)$ where student 2 ranks school b first under P'_2 . Then $\text{ABM}_3(P', >) = c$ and $\text{ABM}_4(P', >) = \emptyset$, which implies $\pi(2, 4) > \pi(3, 3)$. It follows that no rank-priority order is consistent with the way in which ABM matches students to schools.

Example 5. The *school-proposing Boston mechanism (SPBM)* is a monotonic rank-priority mechanism that is not an FC mechanism. This mechanism works like the Boston mechanism, but schools and students swap roles.

In the first round, each school offers a seat to the student who has highest priority at that school. Each student then accepts the best offer she receives, unless none of the offers come from a school that she finds acceptable; in this case, she remains unmatched. In the k^{th} round, each school with unfilled seats makes an offer to the student who is in the k^{th} position of the priority order of that school. Students who have not matched in any previous round accept the offer from the school they prefer most, unless all offers are unacceptable. The process ends when all students have been matched or all seats are filled.

SPBM violates FCM: To see this, consider two students $I = \{1, 2\}$ and two schools $S = \{a, b\}$ with one seat each. The preferences and priorities are

$$\begin{aligned} P_1 &: a P_i b \emptyset, \\ P_2 &: b P_4 a \emptyset, \\ >_a &: 2 >_s 1, \\ >_b &: 1 >_s 2. \end{aligned}$$

The resulting matching is $\text{SPBM}_1(P, >) = b$ and $\text{SPBM}_2(P, >) = a$. However, only the matching μ with $\mu_1 = a$ and $\mu_2 = b$ is FCM. There exists a monotonic rank-priority order (i.e., π with $\pi(k, l) \leq \pi(k', l')$ whenever $(k, l) \leq (k', l')$) that is consistent with SPBM, namely

$$\pi(1, 1) < \dots < \pi(n, 1) < \pi(1, 2) < \dots < \pi(n, 2) < \pi(1, 3) < \dots < \pi(n, m). \quad (2)$$

E. Proof of Theorem 3

For the fixed problem $(P, >)$ and sophisticated students $A \subseteq I$, let ν be an augmented-stable matching.

Necessity. Let (P_A^*, P_{-A}) be an A -Nash equilibrium, for all $i \in A$, let $\varphi_i((P_A^*, P_{-A}), >)$ be the school to which student i is matched in this equilibrium, and let P'_i be any preference order with $\text{choice}_{P'_i}(1) = \varphi_i((P_A^*, P_{-A}), >)$. Our proof uses the following two claims:

Claim 1. (P'_A, P_{-A}) is an A -Nash equilibrium, and the outcomes $\varphi((P'_A, P_{-A}), >)$ and $\varphi((P_A^*, P_{-A}), >)$ are A -equivalent.

Proof of Claim 1. Observe that $\varphi_i((P'_A, P_{-A}), >)$ and $\varphi_i((P_A^*, P_{-A}), >)$ are A -equivalent by Lemma 2. Assume towards contradiction that (P'_A, P_{-A}) is not an A -Nash equilibrium. Then, there exists a student $i \in A$ and a preference order P''_i such that

$$s = \varphi_i((P''_i, P'_{A \setminus \{i\}}, P_{-A}), >) P_i \varphi_i((P'_A, P_{-A}), >) = \varphi_i((P_A^*, P_{-A}), >). \quad (3)$$

By Lemma 2, we can assume $\text{choice}_{P''_i}(1) = s$. Since (P_A^*, P_{-A}) is an A -Nash equilibrium,

i cannot find a beneficial deviation (e.g., to obtain s). Thus, there exist at least q_s students who rank s first under (P_A^*, P_{-A}) , have higher priority than i at s , and are therefore matched to a s under $\varphi((P_A^*, P_{-A}), >)$. By construction, these students still rank s first under $(P_i'', P_{A \setminus \{i\}}', P_{-A})$ and thus exhaust s in the first step of φ . This contradicts the assumption that i could obtain s by deviating from P_i' to P_i'' . \square

Claim 2. $DA((P_A', P_{-A}), \widehat{\succ})$ is an augmented-stable matching with respect to the true preference profile P and A -equivalent to the outcome $\varphi((P_A^*, P_{-A}), >)$.

Proof of Claim 2. Let $\nu = DA((P_A', P_{-A}), \widehat{\succ})$. To see A -equivalence between ν and $\varphi((P_A^*, P_{-A}), >)$, observe that at the problem $((P_A', P_{-A}), >)$ the first round of DA and φ (at $((P_A', P_{-A}), \widehat{\succ})$) match exactly the same students to their reported first choices. Students who enter further rounds under DA are necessarily sincere (i.e., not in A) and only apply to schools which they have not ranked first in any subsequent rounds. By construction of $\widehat{\succ}$, they have lower priority at any school where they apply than any student who was tentatively accepted in the first round. Thus, all students from A are matched to the school they ranked first under P_A' . With Claim 1, this implies that the matchings $DA((P_A', P_{-A}), \widehat{\succ})$, $\varphi((P_A', P_{-A}), >)$, and $\varphi((P_A^*, P_{-A}), >)$ are A -equivalent.

To see augmented stability of ν , assume towards contradiction that there exists some student $i \in I$ and some $s \in S \cup \{\emptyset\}$ such that $s \neq \nu_i$ and either s has an unfilled seat under ν (or $s = \emptyset$) or there exists another student $i' \in I$ with $s = \nu_{i'}$ and $i \succ_s i'$. If $i \notin A$ is a sincere student, then she cannot be part of a blocking pair. If $i \in A$ and s has an unfilled seat, then there are strictly less than q_s students who rank s first under (P_A', P_{-A}) . Thus, s is not exhausted in the first step when φ is applied to $((P_A', P_{-A}), >)$. Consequently, i can obtain s (instead of ν_i) by ranking s first in the A -Nash equilibrium (P_A', P_{-A}) of φ , a contradiction. If a student i' with lower priority under $\widehat{\succ}_s$ than i holds a seat at s , i can claim this seat in the same way, again a contradiction. \square

In Claim 2 (using Claim 1), we have constructed the matching ν that is augmented stable with respect to the true preference profile P and A -equivalent to $\varphi((P_A^*, P_{-A}), >)$. This concludes the proof of necessity.

Sufficiency. Fix an arbitrary augmented-stable matching μ , and let (P_A^*, P_{-A}) be a preference profile such that $\text{choice}_{P_i^*}(1) = \mu_i$ for all $i \in A$. Our proof uses the following two claims:

Claim 3. *The matchings μ and $\varphi((P_A^*, P_{-A}), >)$ are A -equivalent.*

Proof of Claim 3. Assume towards contradiction that there exists a student $i \in A$ such that $\mu_i \neq \varphi_i((P_A^*, P_{-A}), >)$. Since i ranks μ_i first under P_i^* , μ_i must be exhausted in the first step of φ by students who rank μ_i as their first choices and have higher priority than i at μ_i under $>$. Any of these q_s students who is strategic (i.e., $i' \in A$) must be matched to μ_i under the matching μ , otherwise, i' would not rank μ_i first in $P_{i'}^*$; and if $i' \notin A$, then μ_i must be the true first choice of i' , so that $i' \hat{\succ}_{\mu_i} i$ by construction. Since i is matched to μ_i under μ , one of these q_s students, i' say, is not matched to μ_i but some other school $\mu_{i'} \neq \mu_i$. This implies $i' \notin A$ (otherwise, $\mu_{i'} = \mu_i$ by the argument above). Thus, μ_i is the true first choice of i' and $i' \hat{\succ}_{\mu_i} i$, so i' and μ_i form a blocking pair, a contradiction. \square

Claim 4. *The preference profile (P_A^*, P_{-A}) is an A -Nash equilibrium of φ at $(P, >)$.*

Proof of Claim 4. Assume towards contradiction that there exists some student $i \in A$ and a preference order P'_i such that $s = \varphi_i((P'_i, P_{A \setminus \{i\}}^*, P_{-A}), >) \neq \varphi_i((P_A^*, P_{-A}), >) = \mu_i$. By Lemma 2 we can assume that i ranks s first under P'_i . Thus, under the deviation, i takes s in the first step, displacing another student or claiming an empty seat. If i claims an empty seat, then μ is not augmented stable, a contradiction. If i displaces another student, i' say, then $i \hat{\succ}_s i'$. This implies that i and s block the matching μ with respect to the augmented priorities $\hat{\succ}$, again a contradiction. \square

We have constructed the preference profile (P_A^*, P_{-A}) that is an A -Nash equilibrium such that the matching $\varphi((P_A^*, P_{-A}), >)$ is A -equivalent to μ . This concludes the proof of sufficiency.

F. Proof of Theorem 4

For simplicity we introduce the following notation: If all sophisticated students $i \in A$ simultaneously prefer some matching μ to another matching ν , then we say that μ A -dominates ν .

Proof of Statement 1. We construct a preference profile P^* that is an A -Nash equilibrium and the matching $\varphi(P^*, >)$ weakly A -dominates ν . The statement then

follows from the fact that the matching $\varphi(P^*, >)$ is weakly A -dominated by any A -optimal Nash equilibrium matching. The preference profile P^* is created as follows: Let $P^0 = (P_{A'}^0, P_{-A'})$ be the preference profile where each student $i \in A'$ ranks the school ν_i first and all other students report truthfully. By Claim 1 in the proof of Theorem 3, P^0 is also an A' -Nash equilibrium and the matching $\varphi(P^0, >)$ is A' -equivalent to ν . Let $P^1 = (P_A^0, P_{i'}, P_{-A'}) = (P_A^0, P_{-A})$ be the same preference profile except that i' reports $P_{i'}$ truthfully instead of reporting $P_{i'}^0$.

Consider the simple case where $s' = \nu_{i'}$ is the true first choice of i' . Then P^1 is an A -Nash equilibrium and $\varphi(P^1, >)$ is A -equivalent to ν . We can simply set $P^* = P^1$.

Next, consider the case where s' is not the true first choice of i' . Then, if we apply φ to the problem $(P^1, >)$, i' is rejected by her first choice in the first step (otherwise, i' could have obtained her first choice by ranking it first, which contradicts the fact that P^0 is an A' -Nash equilibrium). Let $\mu^1 = \varphi(P^1, >)$. Observe that $\mu_i^1 = \nu_i$ for all $i \in A$. We now construct preference profiles $P^k, k \in \{2, \dots, K\}$ in steps, where $P^* = P^K$ is constructed in the last step.

Step 0. Set $k = 1$, $s^1 = s'$, and $\mu^1 = \varphi(P^1, >)$.

Step 1. Let A^k be the (possibly empty) set of sophisticated students who prefer s^k to their match under μ^k (i.e., $i \in A^k$ if $i \in A$ and $s^k P_i \mu_i^k$) and let I^k be the (possibly empty) set of sincere students who rank s^k as their first choice but are not matched to it under μ^k (i.e., $i \in I^k$ if $i \notin A$, $\text{choice}_{P_i}(1) = s^k$, and $s^k \neq \mu_i^k$).

Step 2. If $A^k \cup I^k = \emptyset$, set $P^* = P^k$; end the process.

Step 3. Else, let $i^k \in (A^k \cup I^k)$ be the student with the highest priority at s^k of those students (i.e., $i^k >_{s^k} i$ for all $i \in (A^k \cup I^k), i \neq i^k$).

Step 4. If $i^k \in I^k$, then set $P^* = P^k$ and end the process.

Step 5. If $i^k \in A^k$, then define the new preference profile P^{k+1} by setting

$$P_i^{k+1} = \begin{cases} P_i^k, & \text{if } i \in A \setminus \{i^k\}, \\ P_i^{k+1}, & \text{if } i = i^k, \\ P_i, & \text{if } i \in I \setminus A, \end{cases} \quad (4)$$

where $P_{i^k}^{k+1}$ is a preference order under which i^k ranks s^k as her first choice.

Step 6. Set $\mu^{k+1} = \varphi(P^{k+1}, >)$, then set $k \rightsquigarrow k + 1$ and return to Step 1.

The following Claim 5 completes the proof of statement 1 in Theorem 4.

Claim 5. *The construction of the preference profile P^* as described above ends after finitely many steps, the matching $\varphi(P^*, >)$ A -dominates the matching ν , and P^* is an A -Nash equilibrium of φ at $(P, >)$.*

Proof of Claim 5. First, observe that for any $k \geq 1$, each sophisticated student $i \in A$ weakly prefers μ_i^{k+1} to μ_i^k , and the student $i^k \in A$ strictly prefers μ_i^{k+1} to μ_i^k . This rules out cycles. The process therefore ends after finitely many iterations and $\varphi(P^*, >)$ A -dominates ν .

Finally, we need to show that P^* is an A -Nash equilibrium. Assume towards contradiction that some student $i \in A$ has a beneficial deviation $P'_i \neq P_i^*$. Let $s = \varphi_i((P_A^*, P_{-A}), >)$ and $s' = \varphi_i((P_{A \setminus \{i\}}^*, P'_i, P_{-A}), >)$. By Lemma 2, we can assume that i ranks s' as her first choice under P'_i .

Suppose that i can obtain s' because s' is not exhausted in the first step when φ is applied to P^* . Since $i \in A$, i weakly prefers s to ν_i , which implies $s' P_i \nu_i$. Thus, s' must have been exhausted in the first step when φ is applied to P^1 ; otherwise, P'_i would have been a strictly beneficial deviation from the A -Nash equilibrium P^1 for i . Therefore, s' must have become available in some step k of the transition from P^1 to P^* . At this point, we know that $i \in A^k$ was a candidate to claim the seat at s' , which implies that the seat must have been taken by some student (possibly not by i but a student with higher priority at s' than i), a contradiction.

Conversely, suppose that i can obtain s' because she has higher priority at s' than some other student \tilde{i} who is matched to s' under $\varphi(P^*, >)$. \tilde{i} can not be matched to s' under $\varphi(P^1, >)$, otherwise, i could have benefited by deviating from the A -Nash equilibrium P^1 . Thus, \tilde{i} must be matched to s' in some step of the transition. But if i has priority over \tilde{i} at s' , then i would have been chosen to receive s' in this step, a contradiction. \square

Proof of Statement 2. Let (P_A^*, P_{-A}) be an A -optimal Nash equilibrium of φ at $(P, >)$ that leads to the matching μ . First, suppose that i' receives her first choice under μ (i.e., $\mu_{i'} = \text{choice}_{P_{i'}}(1)$). Then truthful reporting is a best response for i' to the preference reports (P_A^*, P_{-A}) from the other students. Since (P_A^*, P_{-A}) is an A -Nash equilibrium and truthful reporting is a best response for i' , it must also be an A' -Nash equilibrium. Therefore, the matching ν A' -dominates the matching μ , which implies that i' receives her first choice under ν as well.

Next, suppose that i' does not receive her first choice under μ . Then i' is rejected by her first choice in the first step when φ is applied to the problem $((P_A^*, P_{-A}), \succ)$. The best school that i' can possibly obtain is therefore the school she prefers most out of all the schools that are not exhausted in the first step of φ applied to $((P_A^*, P_{-A}), \succ)$, say s . Now, let $\hat{\succ}^A$ and $\hat{\succ}^{A'}$ denote the augmented priority profiles that arise from \succ when the sets of sophisticated students are given by A and $A' = A \cup \{i'\}$, respectively, and let $\hat{\succ}^{A, i' \sim s}$ be the same priority profile as $\hat{\succ}^A$ except that i' has priority at s over all sincere students who do not rank s as their first choice. Next, we prove the two Claims 6 and 7.

Claim 6. $DA_i(P, \hat{\succ}^A) = \varphi_i((P_A^*, P_{-A}), \succ)$ for all $i \in A$ and those $i \notin A$ for whom $\varphi_i((P_A^*, P_{-A}), \succ) = \text{choice}_{P_i}(1)$.

Proof of Claim 6. $DA(P, \hat{\succ}^A)$ is the student-optimal augmented-stable matching with respect to augmented priorities $\hat{\succ}^A$. In particular, it is preferred by all sophisticated students to any other augmented-stable matching. Thus, $DA(P, \hat{\succ}^A)$ is A -equivalent to the A -optimal Nash equilibrium outcome $\varphi((P_A^*, P_{-A}), \succ)$ by Theorem 3.

Now, assume towards contradiction that there exists some sincere student $i \notin A$ who receives her first choice \tilde{s} under $\varphi((P_A^*, P_{-A}), \succ)$ but is rejected by \tilde{s} under $DA(P, \hat{\succ}^A)$. By A -equivalence of the matchings $DA(P, \hat{\succ}^A)$ and $\varphi((P_A^*, P_{-A}), \succ)$, the sophisticated students who are matched to \tilde{s} are exactly the same in both matchings. Thus, there exists a sincere student \tilde{i} who gets \tilde{s} under $DA(P, \hat{\succ}^A)$ but does not get \tilde{s} under $\varphi((P_A^*, P_{-A}), \succ)$. Since i ranks \tilde{s} first, \tilde{i} must also rank \tilde{s} first (otherwise, i would have $\hat{\succ}^A$ -priority over \tilde{i} at \tilde{s} by construction) and \tilde{i} must have \succ -priority over i at \tilde{s} . But then both i and \tilde{i} compete for \tilde{s} in the first step of φ , and since i gets \tilde{s} , \tilde{i} must get \tilde{s} as well, a contradiction. \square

Claim 7. $DA_{i'}(P, \hat{\succ}^{A, i' \sim s}) = s$.

Proof of Claim 7. Let \tilde{S} be the set of schools that are filled in the first step of φ when it is to the problem $((P_A^*, P_{-A}), \succ)$. By Claim 6, \tilde{S} coincides with the set of schools that are filled exclusively by sophisticated students and those sincere students who rank them as their first choices under the matching $DA(P, \hat{\succ}^A)$. Recall that s is the school that i' prefers most of all the schools that are not in \tilde{S} and observe that i' does not get any of these schools, independent of her priority at s . Thus, s is the most preferred school that i' could possibly obtain in any matching that is stable with respect to the priority profile $\hat{\succ}^{A, i' \sim s}$. $DA(P, \hat{\succ}^{A, i' \sim s})$ is such a matching. Now, consider the application process

when DA is applied to the preference profile P and the priority profile $\hat{\succ}^{A, i' \sim s}$: i' definitely applies to s at some point because she will have been rejected by all more preferred schools. Under $\hat{\succ}^{A, i' \sim s}$, only strategic students and those sincere students who rank s as their first choices have higher priority than i' at s . Thus, i' can only be rejected by s if s is exhausted by such students at that point (and therefore at any later point as well). But $s \notin \tilde{S}$, so i' is not rejected by s . \square

To complete the proof of Statement 2 of Theorem 4, we observe that

$$\nu_{i'} = \text{DA}_{i'}(P, \hat{\succ}^{A'}) R_{i'} \text{DA}_{i'}(P, \hat{\succ}^{A, i' \sim s}) = s R_{i'} \mu_{i'}, \quad (5)$$

where the first equality holds because the matching $\text{DA}(P, \hat{\succ}^{A'})$ is A' -equivalent to any A' -optimal Nash equilibrium matching (by Theorem 3), the first preference relation holds because DA respects improvements (Balinski and Sönmez, 1999), the second equality holds because of Claim 7, and the last preference relation holds by definition of s .

Proof of Statement 3. This follows from Example 3.

Proof of Statement 4. Assume towards contradiction that Statement 4 is false, i.e., there exists a student i who receives her true first choice, say s , under ν but not under μ . Then $i \notin A$, since all sophisticated students in A prefer μ to ν by Statement 1. On the other hand, there must exist some student \tilde{i} who receives s under μ but not under ν and this student must have higher priority than i at s . Since \tilde{i} is not sophisticated, she ranks s first and therefore applies to s in the first step of φ . By first-choice stability of φ and the fact that i receives s in ν , \tilde{i} must also receive s in ν , a contradiction.

G. Failure of Strong Comparison

Definition 8 (As Strongly Manipulable As-relation). ψ is *as strongly manipulable as* φ if whenever φ is manipulable by some student $i \in I$ at some problem $(P, >)$, then ψ is also manipulable by the same student i at the problem $(P, >)$.

Proposition 2. *BM is not as strongly manipulable as ABM.*

Proof. Consider a problem $(P, >)$ with five students $I = \{1, \dots, 5\}$ and five schools

$S = \{a, b, c, d, e\}$ with a single seat each. Let the preference profile be given by

$$\begin{aligned}
 a & P_1 \dots, \\
 b & P_2 \dots, \\
 d & P_3 \dots, \\
 a & P_4 b P_4 d P_4 c P_4 e, \\
 a & P_5 b P_5 c P_5 d P_5 e,
 \end{aligned}$$

and let the priorities be the same at all schools such that $1 \succ_s \dots \succ_s 5$ for all $s \in S$.

Under BM and truthful reporting we have $\text{BM}_5(P, \succ) = c$. Since a and b are taken by students 1 and 2 in the first round and both students have priority over student 5, student 5 cannot improve her assignment by misreporting. Under the adaptive Boston mechanism and truthful reporting we have $\text{ABM}_5(P, \succ) = e$. However, if student 5 ranks school c in first position (i.e., she reports $c P'_5 \dots$), then $\text{ABM}_5((P'_5, P_{-5}), \succ) = c$. This represents a strict improvement for student 5. Thus, for the problem (P, \succ) , ABM is manipulable by student 5 but BM is not. \square

Corollary 2. *BM and ABM are incomparable by the as strongly manipulable as-relation.*

Proof. Proposition 2 already shows that BM is not as strongly manipulable as ABM. If the comparison was possible, then ABM would have to be more strongly manipulable than BM. However, a simple example shows that this is not the case: Consider a problem (P, \succ) with four students $I = \{1, \dots, 4\}$ and four schools $S = \{a, b, c, d\}$ with a single seat each. Let the preference profile be given by

$$a P_1 \dots, \tag{6}$$

$$b P_2 \dots, \tag{7}$$

$$a P_3 b P_3 c P_3 d, \tag{8}$$

$$a P_4 c P_4 \dots, \tag{9}$$

and let the priorities be the same at all schools such that $1 \succ_s \dots \succ_s 4$ for all $s \in S$. Under BM and truthful reporting we have $\text{BM}_3(P, \succ) = d$, but if student 3 reports $a P'_3 c P'_3 \dots$ she will be assigned to c . Under the adaptive Boston mechanism and

truthful reporting we have $\text{ABM}_3(P, \succ) = c$. Since a and b are taken by students 1 and 2 in the first round, student 3 cannot improve her assignment by misreporting. \square

H. Failure of Weak Comparison for Random Priorities

H.1. Modeling Random Priorities in School Choice Mechanisms

In many school choice settings, priorities are not strict but coarse. This means that when two students apply for the same seat at some school, then ties between these students must be broken. Normally, this is done using a random tie-breaker. A *random assignment* of the students to the schools is represented by an $n \times m$ -matrix $x = (x_{i,s})_{i \in I, s \in S}$ where $x_{i,s} \in [0, 1]$ denotes the probability that student i is assigned to school s . From the perspective of the students, reporting a different preference order leads to a different random assignment. Therefore, we need to extend their preferences, which we do by endowing them with vNM utility functions. Given a preference order P_i , a utility function $u_i : S \rightarrow \mathbb{R}^+$ is *consistent with P_i* if $u_i(s) > u_i(s')$ whenever $s P_i s'$, and we denote by U_{P_i} the set of all vNM utility functions that are consistent with P_i . We assume that students wish to maximize their expected utility.

To model the uncertainty from random tie-breaking, we assume that any mechanism first collects the preference orders of the students and then chooses a priority profile randomly from a distribution \mathbb{P} over priority profiles. We denote the resulting mechanism by $\varphi^{\mathbb{P}}$, where $\varphi_i^{\mathbb{P}}(P)$ denotes the random assignment vector to student i if the reported preference profile is P . This is the i^{th} row of the random assignment matrix. In particular, we denote by \mathbb{U} the uniform distribution over all single priority profiles, that is the priority profiles $\succ = (\succ_1, \dots, \succ_1)$ where the priority order \succ_s is the same at all schools.

To study mechanisms in the presence of random priorities we consider *random problems* u , which simply consist of a profile of utility functions. Note that this problem contains no priority profile because determining this profile is now part of the mechanism. We consider ordinal mechanisms, which depend only on the ordinal preference profile induced by the utility profile u . This means that for two different utility profiles u, u' that are both consistent with the same preference profile P , any mechanism $\varphi^{\mathbb{P}}$ has to select the same random assignment. Thus, it is without loss of generality that we consider mechanisms as functions of preference profiles (rather than utility profiles).

H.2. Comparing Random Mechanisms by Their Vulnerability to Manipulation

Since we have changed the structure of the problem to accommodate random mechanisms, we also need to re-define the concepts for the comparison of mechanisms by their vulnerability to manipulation. Let \mathbb{P} be a priority distribution and let $\varphi^{\mathbb{P}}$ and $\psi^{\mathbb{P}}$ be two random mechanisms.

Definition 9. $\psi^{\mathbb{P}}$ is *manipulable by student i at problem u* if there exists a preference order $P'_i \neq P_i$ such that $\mathbb{E}_{\psi^{\mathbb{P}}(P'_i, P_{-i})}[u_i] > \mathbb{E}_{\psi^{\mathbb{P}}(P)}[u_i]$. $\psi^{\mathbb{P}}$ is *manipulable at u* if it is manipulable by some student $i \in I$ at the problem u . $\psi^{\mathbb{P}}$ is *manipulable* if it is manipulable at some problem.

Definition 10. $\psi^{\mathbb{P}}$ is *as manipulable as $\varphi^{\mathbb{P}}$* if whenever $\varphi^{\mathbb{P}}$ is manipulable at some problem u , then $\psi^{\mathbb{P}}$ is also manipulable at the same problem u . $\psi^{\mathbb{P}}$ is *as strongly manipulable as $\varphi^{\mathbb{P}}$* if whenever $\varphi^{\mathbb{P}}$ is manipulable by some student i at some problem u , then $\psi^{\mathbb{P}}$ is also manipulable by i at u .

H.3. Failure of the Comparison of BM and ABM

Proposition 3. $BM^{\mathbb{U}}$ is not as manipulable as $ABM^{\mathbb{U}}$, and $ABM^{\mathbb{U}}$ is not as manipulable as $BM^{\mathbb{U}}$.

Proof. We construct a problem $u^{(1)}$ for which $BM^{\mathbb{U}}$ is manipulable but $ABM^{\mathbb{U}}$ is not, and we construct a second problem $u^{(2)}$ for which $ABM^{\mathbb{U}}$ is manipulable but $BM^{\mathbb{U}}$ is not.

- (1) Consider a problem $u^{(1)}$ with four students $I = \{1, \dots, 4\}$ and four schools $S = \{a, b, c, d\}$ with a single seat each. Let the preference profile be given by

$$\begin{aligned} a & P_1^{(1)} & b & P_1^{(1)} & c & P_1^{(1)} & d, \\ a & P_2^{(1)} & c & P_2^{(1)} & b & P_2^{(1)} & d, \\ a & P_3^{(1)} & c & P_3^{(1)} & b & P_3^{(1)} & d, \\ b & P_4^{(1)} & \dots & & & & \end{aligned}$$

Student 1's assignment vector under BM^{U} is $\text{BM}_1^{\text{U}}(P^{(1)}) = (1/3, 0, 0, 2/3)$ for a, b, c, d , respectively. If student 1 swaps b and c in her report (i.e., she reports $a P'_1 c P'_1 b P'_1 d$), her assignment vector changes to $\text{BM}_1^{\text{U}}(P'_1, P_{-1}^{(1)}) = (1/3, 0, 1/3, 1/3)$. Since $\text{BM}_1^{\text{U}}(P'_1, P_{-1}^{(1)})$ first order-stochastically dominates $\text{BM}_1^{\text{U}}(P^{(1)})$ at $P_1^{(1)}$, this misreport is an unambiguous improvement for student 1 (independent of her vNM utility function $u_1^{(1)}$).

For settings with four students and four schools with a single seat each, ABM^{U} is $1/3$ -partially strategyproof (Mennle and Seuken, 2017a). Thus, if all students have utilities $(9, 3, 1, 0)$ for their first, second, third, and last choices, respectively, truthful reporting is a dominant strategy for all of them. With the utility profile $u^{(1)}$ defined in this way BM^{U} is manipulable (by student 1) at the problem $u^{(1)}$ but ABM^{U} is not manipulable (by any student).

- (2) Consider a problem $u^{(2)}$ with six students $I = \{1, \dots, 6\}$ and six schools $S = \{a, b, c, d, e, f\}$ with a single seat each. Let the preference profile be given by

$$a P_1^{(2)} e P_1^{(2)} c P_1^{(2)} d P_1^{(2)} f P_1^{(2)} b, \quad (10)$$

$$a P_2^{(2)} e P_2^{(2)} c P_2^{(2)} d P_2^{(2)} f P_2^{(2)} b, \quad (11)$$

$$a P_3^{(2)} e P_3^{(2)} d P_3^{(2)} c P_3^{(2)} f P_3^{(2)} b, \quad (12)$$

$$a P_4^{(2)} e P_4^{(2)} d P_4^{(2)} c P_4^{(2)} f P_4^{(2)} b, \quad (13)$$

$$b P_5^{(2)} c P_5^{(2)} \dots, \quad (14)$$

$$b P_6^{(2)} d P_6^{(2)} \dots \quad (15)$$

Suppose that all students have utilities $(120, 30, 19, 2, 1, 0)$ for their first through sixth choices, respectively.

First, we study the incentives to manipulate under BM^{U} . Note that student 1 cannot improve her expected utility by ranking another school than a first. To see this, observe the following: under truthful reporting she obtains a with probability $1/4$ but if she ranks a different school first she will at best obtain her second choice e with certainty. Since b is exhausted in the first round, she will not obtain b with any positive probability unless she ranks it first. Therefore, it is a weakly better response to rank b last. Since f is the worst school which she obtains with positive probability, it is a weakly better response to leave f in fifth position. Otherwise,

she will only reduce her chances at more preferred schools. Thus, without loss of generality, any beneficial misreport only involves the order of the schools e, c, d . It is a simple exercise to compute the changes in expected utility for student 1 under any such misreport:

Report P'_1	Change in expected utility
$a P'_1 e P'_1 d P'_1 c P'_1 f P'_1 b$	-2.1
$a P'_1 c P'_1 e P'_1 d P'_1 f P'_1 b$	-0.4
$a P'_1 c P'_1 d P'_1 e P'_1 f P'_1 b$	-0.3
$a P'_1 d P'_1 e P'_1 c P'_1 f P'_1 b$	-9.5
$a P'_1 d P'_1 c P'_1 e P'_1 f P'_1 b$	-8.1

This shows that truthful reporting is a best response for student 1. The same is true for student 2, 3, and 4 by symmetry. Student 5 receives her first and second choices with probabilities $1/2$ each. Thus, the only way she can improve her expected utility is by increasing her probability for her first choice but this is obviously impossible. The same is true for student 6 by symmetry. In combination this means that for the problem $u^{(2)}$ the mechanism BM^{U} is not manipulable.

Second, we study the incentives to manipulate under ABM^{U} for the same problem: under truthful reporting, student 1's assignment vector is $\text{ABM}_1^{\text{U}}(P^{(2)}) = (1/4, 1/4, 1/8, 1/8, 1/4, 0)$ for the schools a, e, c, d, f, b , respectively. If student 1 reported P'_1 with $a P'_1 c P'_1 d P'_1 e P'_1 f P'_1 b$ instead, her assignment vector would be $\text{ABM}_1^{\text{U}}(P'_1, P_{-1}^{(2)}) = (1/4, 0, 71/120, 3/40, 1/12, 0)$. This means an increase in expected utility from 40.375 to 41.475, a strict improvement.

This concludes the proof. □