First Choice-Maximizing School Choice Mechanisms*

Umut Dur† Timo Mennle‡ Sven Seuken‡

November 1, 2016

Abstract

We investigate *first choice-maximality (FCM)* (i.e., a maximal share of students is matched to their reported first choices), a common desideratum in the design of school choice mechanisms. No FCM mechanism can be stable; however, we find *first choice-stability (FCS)* (i.e., no student forms a blocking pair with her first choice) to be the strongest rank-based relaxation of stability that is compatible with FCM. Focusing on the class of FCM and FCS mechanisms, we show that the Pareto efficient members form a minimally manipulable subset (including the Boston mechanism). Moreover, we identify the Nash equilibrium outcomes of any mechanism in this class when all students act strategically (generalizing (Ergin and Sönmez, 2006)) and when some student report truthfully (in part generalizing (Pathak and Sönmez, 2008)). Our results suggest a novel approach to the design of school choice mechanisms where strategic students self-select into receiving their school in a first step, so that the matching of truthful students can be made independently of the usual incentive constraints.

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

JEL: C78, D47

*We would like to thank Antonio Miralles and participants of the ZEW Market Design Research Seminar for helpful comments on this work. Any errors remain our own. Part of this research was supported by a research grant from the Swiss National Science Foundation.

†North Carolina State University, Department of Economics, 2801 Founders Drive, Raleigh, NC, 27695; e-mail: umutdur@gmail.com; web: https://sites.google.com/site/umutdur/

‡University of Zurich, Department of Informatics, Binzmühlestrasse 14, CH-8050 Zurich; e-mail: {mennle,seuken}@ifi.uzh.ch; web: https://www.ifi.uzh.ch/ce/people/

1. Introduction

School choice gives students an opportunity to choose which public school they would like to attend. Ideally, one would like to give every student their first choice. However, this ideal may not be attainable because schools have limited capacities and some schools are usually more popular than others. Therefore, administrators need to design school choice mechanisms that reconcile students' conflicting interests and produce desirable matchings, with high student welfare being one of the key desiderata.

One simple measure for student welfare is the *share of students who are matched to their first choice school.*[1] This measure is tangible because maximizing it generalizes the desire to match all students with their first choice in a straightforward way. Unsurprisingly, it receives attention from the media in reports on the performance of school choice markets with headlines such as

> *One in six secondary pupils in England doesn't get first school choice,*[2]
>
> *45% of New York City 8th-graders got into top high school choice [...].*[3]

Furthermore, administrators report this measure as part of the public communication about the school choice system they run, e.g., Denver Public Schools prominently feature the share of students (i.e., kindergarten, 6th graders, and 8th graders) who received their first choice on the website that informs parents about the school choice system.[4]

In line with the common use of this measure in the public discussion of school choice, many school choice mechanisms in practice attempt to maximize it. For example, the *Boston mechanism* (Abdulkadiroğlu and Sönmez, 2003; Kojima and Ünver, 2014) matches as many students as possible to their reported first choice in the first round. If all students report their preferences truthfully and if an administrator cares about the share of students who receive their first choice, this mechanism would indeed be an attractive choice. Variants of the Boston mechanism are being used in many school

[1] Chen and Kesten (2016) called this value the *first choice accommodation index.*

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

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

[4] Denver Public Schools Website. Retrieved October 10, 2016: schoolchoice.dpsk12.org/schoolchoice/

choice markets around the world.[5] Arguably, this mechanism owes much of its popularity to the intuitive way in which it attempts to maximize first choices. The common focus on first choices in school choice markets and the popularity of variants of the Boston mechanism in practice motivate our study of *first choice-maximizing (FCM)* mechanisms in this paper.[6]

A second important desideratum in the design of school choice mechanisms is fairness: a student (say i) may have *envy* of another student (say j) if i prefers j's assignment to her own assignment; and i's envy is *justified* if j's school grants priority to i over j.[7] Intuitively, a student who has justified envy may feel as though she has been treated unfairly by the school choice mechanism. On top of that, students with justified envy may have a basis for pursing legal actions against school districts, something that administrators would naturally try to avoid (Abdulkadiroğlu, 2011). *Stability* requires that if any student prefers another school to her matched school, then all seats at that other school must be filled by students who have higher priority. Stability eliminates justified envy and it is therefore an important notion of fairness in school choice markets.

Our first, simple observation in this paper is that no mechanism can be FCM and stable at the same time. Taking FCM as a primary objective, this raises the question whether there exists a relaxed notion of stability that can serve as a useful second-best but is also compatible with FCM. We answer this question in the affirmative: a matching is *first choice-stable (FCS)* if for any student who is not matched to her first choice, all seats at her first choice are taken by students with higher priority. In fact, FCS turns out to be the strongest rank-based relaxation of stability that is compatible with FCM. Following this insight, our main focus in this paper is on the class of FCM and FCS mechanism; typical members of this class include all common variants of the Boston mechanism. We make three main contributions:

First, we study the incentive properties of FCM and FCS mechanisms and show that FCM and strategyproofness are incompatible. While this means that all mechanisms in

[5]Examples inlcude Minneapolis, Seattle, Lee County (Kojima and Ünver, 2014), San Diego, Freiburg (Germany), and throughout the German state of Nordrhein Westfalen (Basteck, Huesmann and Nax, 2015).

[6]Specifically, we say that a mechanism is *first choice-maximizing* if it assigns a maximal share of the students to their reported first choice.

[7]Priorities are a common feature of school choice markets; for example, schools may grant priority to students who already have siblings at a given school, who live near that school, or because they have previously achieved higher grades (Dur et al., 2014).

this class are susceptible to strategic manipulation, we show that the Pareto optimal members form a minimally manipulable subset when comparing the mechanisms by their vulnerability to manipulation (Pathak and Sönmez, 2013). Thus, Pareto efficiency and robustness to manipulation go hand-in-hand for FCM and FCS mechanisms. This teaches us that market designers should not settle for Pareto *in*efficient (FCM and FCS) mechanisms because they can be unambiguously dominated by another member of this class in terms of student welfare without deteriorating the incentives for truthtelling.

Second, despite the fact that the Pareto efficient members are least manipulable, FCM precludes full strategyproofness. Therefore, we identify the Nash equilibria of the induced preference revelation game. We find that the Nash equilibrium outcomes under any FCM and FCS mechanism correspond to the matchings that are stable with respect to the true preferences.[8] Thus, the best conceivable equilibrium outcome in terms of student welfare is the student-optimal stable matching. However, this matching can be implemented in the dominant strategy equilibrium of the Student-Proposing Deferred Acceptance (DA) mechanism (Gale and Shapley, 1962). Therefore, if we expect all students to act strategically, our results shows that no FCM and FCS mechanism improves student welfare over DA.

For our third contribution, we relax the assumption that all students act strategically; following Pathak and Sönmez (2008), we identify the Nash equilibrium outcomes when only a subset of the students act strategically while all others report their preferences truthfully. We show that, from the perspective of the strategic students, these Nash equilibrium outcomes correspond to the matchings that are stable with respect to the true preferences and the *augmented* priorities.[9] This correspondence allows us to identify the marginal impact on student welfare of any student's decision on whether to act strategically or to report her preferences truthfully. Most notably, we show that more students receive their true first choice when fewer students act strategically (in the equilibria that lead to the respective strategic student-optimal augmented-stable matchings). This insight is useful for administrators who care about assigning true

[8] Our result generalizes the finding of Ergin and Sönmez (2006), who proved this correspondence for the specific Boston mechanism.

[9] Our result generalizes the finding of Pathak and Sönmez (2008), who proved such a correspondence for the Boston mechanism, except that our identification leaves open how the truthful students are matched. This limitation is owed to the fact that the properties FCM and FCS only specify the treatment of first choices but leave substantial freedom for the treatment of all other choices, which affects only non-strategic students in equilibrium.

first choice: if some students report their preferences truthfully, then FCM and FCS mechanisms can improve student welfare over DA in equilibrium. In environments with only few strategic students, this may justify the use of carefully designed FCM and FCS mechanisms.

In this paper we shed light on the consequences of focusing on first choices in the design of school choice markets. We demonstrate the severe limitations that FCM imposes on fairness and incentives for truthtelling. At the same time, we find that FCM and FCS mechanisms may lead to more assigned true first choices when only some students act strategically; and we conclude with a recipe for improving student welfare in equilibrium by choosing a different mechanisms from this class.

2. Preliminaries

2.1. Formal Model

A *school choice problem* is a tuple $(I, S, q, P, \succ)$, where

- $I = \{i_1, \ldots, i_n\}$ is the set of n *students*,
- $S = \{s_1, \ldots, s_m\}$ is the set of m *schools*,
- $q = (q_s)_{s \in S}$ is the vector of *school capacities*, where q_s is the number of seats available at school s,
- $P = (P_i)_{i \in I}$ is the *preference profile* where P_i is the strict *preference order* of student i over the schools in S and being unassigned option denoted by $\varnothing$,
- $\succ = (\succ_s)_{s \in S}$ is the *priority profile* where $\succ_s$ is the *priority order* of school s over students in I.

We assume that there is at least one seat at each school (i.e., $q_s \geqslant 1$ for all $s \in S$) and that any student can receive her outside option (i.e., $q_\varnothing = n$). For any preference order P_i, we denote by R_i the corresponding *weak preference order* (i.e., $s \; R_i \; s'$ if either $s \; P_i \; s'$ or $s = s'$). Throughout the paper, we fix the set of students, the set of schools and school capacities and we denote a problem with $(P, \succ)$.

A *matching* is a function $\mu : I \to S \cup \{\varnothing\}$ such that $|\mu^{-1}(s)| \leqslant q_s$ for every school $s \in S$ (i.e., matchings is *feasible*), where $\mu(i)$ is the school to which student i is assigned

and $\mu^{-1}(s)$ is the set of students who are assigned a seat at the schools s. Throughout the paper, we simplify notation by writing μ_i and μ_s for $\mu(i)$ and $\mu^{-1}(s)$, respectively.

For two matchings μ and ν, μ *Pareto dominates* ν if $\nu_i \; R_i \; \mu_i$ for all students $i \in I$, and this dominance is *strict* if in addition $\nu_{i'} \; P_i \; \mu_{i'}$ for at least one student $i' \in I$. A matching μ is *Pareto efficient* if it is not strictly Pareto dominated by any other matching, μ is *individually rational* if $\mu_i \; R_i \; \varnothing$ for each $i \in I$, and μ is *wasteful* if there exists some student $i \in I$ and some school $s \in S$ such that $|\mu_s| < q_s$ but $s \; P_i \; \mu_i$, otherwise, μ is called *non-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 $i \; \succ_s \; i'$; and μ is *free of justified envy* if no student has justified envy under μ. Finally, μ is *stable* if it is individual rational, non-wasteful, and free of justified envy.

Observe that for any unstable matching μ, there must exist at least one pair $(i, s) \in I \times (S \cup \{\varnothing\})$ such that $s \; P_i \; \mu_i$ and $|\mu_s| < q_s$ (i.e., μ is wasteful), or $s = \varnothing$ (i.e., μ is not individual rational), or there exists some other student $i' \in I$ with $\mu_{i'} = s$ and $i \; \succ_s \; i$ (i.e., i has justified envy). Any such student-school pair is called a *blocking pair*. Obviously, μ is stable if and only if there exist no blocking pairs. We employ this equivalent definition of stability to simplify definitions and proofs.

A *school choice 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 assigned under the matching $\varphi(P, \succ)$. Mechanism φ is *Pareto efficient/individual rational/non-wasteful/free of justified envy/stable* if it selects a matching that is *Pareto efficient/individual rational/non-wasteful/free of justified envy/stable* for any problem. φ *Pareto dominates* another school choice mechanism ψ if the matching $\varphi(P, \succ)$ Pareto dominates the matching $\psi(P, \succ)$ for any problem $(P, \succ)$, and φ *strictly Pareto dominates* ψ if this dominance is strict from at least one problem. φ is *strategyproof* if for any problem $(P, \succ)$, any student $i \in I$, and any preference order $P_i' \neq P_i$ we have that

$$\varphi_i(P, \succ) \; R_i \; \varphi_i\left((P_i', P_{-i}), \succ\right), \tag{1}$$

where P_{-i} are the preference orders of all students except i.

2.2. Notation

We introduce some notation that we use throughout the paper. Given the problem $(P, \succ)$, let $\text{choice}_{P_i}(k)$ be the kth *choice* of student i (i.e., the school that i ranks in k^{th} position under the preference order P_i), and let $\text{rank}_{P_i}(s)$ be the *rank* of s in the preference order P_i (i.e., the position in which i ranks s; formally, $\text{rank}_{P_i}(s) = |\{s' \in S : s' \; R_i \; s\}|$ and $\text{choice}_{P_i}(\text{rank}_{P_i}(s)) = s$). Then let $I(s, k, P) \subseteq I$ be the set of students who rank school s as their k^{th} choice; formally, $I(s, k, P) = \{i \in I : \text{choice}_{P_i}(k) = s\}$, and let μ^k be the set of students who receive their k^{th} choice in the matching μ; formally, $\mu^k = \{i \in I : \text{choice}_{P_i}(k) = \mu_i\}$. In particular, $\varphi^k(P, \succ)$ denotes the set of students who receive their k^{th} choice under the mechanism φ at the problem $(P, \succ)$.

2.3. School Choice Mechanisms

Before we continue, we briefly describe a number of well-known school choice mechanisms. While these mechanisms may be implemented in various ways in practice, we give their canonical descriptions as direct revelation mechanisms (i.e., we describe how they determine a matching for any given problem $(P, \succ)$).

The *Boston mechanism* (*BM*) determines a matching in rounds. In the first round, all students apply to their reported first choice. Each school then accepts applications from students according to their priority until all applications are accepted or until all seats are filled. Students whose applications to their first choice are not accepted, enter the second round where they apply to their second choice. Again, schools accept applying students into open seats according to priority and reject all remaining applications once their capacity is exhausted. This process continues (where students who were rejected in round $k-1$, apply to their k^{th} choice in the kth round) until no school receives an application. BM is Pareto efficient but it is neither stable nor strategyproof (Abdulkadiroğlu and Sönmez, 2003).

The *adaptive Boston mechanism* (*aBM*) is similar to BM, except that students who were rejected in round $k-1$, apply to their *most preferred school out of all the schools with positive remaining capacity* in round k. This mechanism has also be studied by researchers (Alcalde, 1996; Miralles, 2008; Dur, 2015; Harless, 2015; Mennle and Seuken, 2015*b*), and variants of this mechanism are used in practice in, e.g., in Freiburg (Germany), throughout German state of Nordrhein Westfalen, and in Amsterdam (until

2014) (Basteck, Huesmann and Nax, 2015; de Haan et al., 2015). Like BM, aBM is Pareto efficient but neither stable nor strategyproof. A striking feature of this mechanism is that it prevents students from making futile applications (i.e., applications to exhausted schools, which have no chance of success, independent of the student's priority).

Under the *Deferred Acceptance mechanism* (*DA*), students also apply to schools in rounds according their reported preference order. However, under DA, acceptances are *tentative* rather than permanent; students can be rejected from a school where they were previously accepted if another student applies who has higher priority. DA has become increasingly popular because of its incentive properties (it is strategyproof for students) and stability of the resulting matchings (Gale and Shapley, 1962; Abdulkadiroğlu and Sönmez, 2003). It mechanism is now being used in Boston, New York, Mexico City, Wake County, and Amsterdam with varying priority structures.

Under the *Top Trade Cycles mechanism* (*TTC*), students and schools from a graph. Each student points to her most preferred school with unfilled seats and each school points to the student who has highest priority at that school. In each step, a cycle of this graph is selected and implemented (i.e., each student in the cycle is matched to the school to which she is pointing, and the seat and the student are removed from the mechanism). Students and school then adjust where they are pointing and the process continues. Abdulkadiroğlu and Sönmez (2003) proposed TTC as an alternative that is strategyproof and Pareto efficient (but not stable).

3. First Choice-Maximization and First Choice-Stability

As we have argued in the introduction, the share of students who receive their first choice is a common measure for student welfare in practice. In this section, we formally define first choice-maximality as a desideratum and we investigate its implications on the design of school choice mechanisms in terms of fairness and incentives for truthtelling.

Definition 1. Given a problem $(P, \succ)$, a matching μ is *first choice-maximizing* if there exists no other matching ν such that a strictly larger number of students get their first choice under ν than under μ. A mechanism φ is *first choice-maximizing* (*FCM*) if for any problem $(P, \succ)$ the matching $\varphi(P, \succ)$ is first choice-maximizing.

The Boston mechanism and its adaptive variant, aBM, are FCM and this is true,

independent of the particular priority structures prevalent in different school choice markets. Rank Value mechanism (Featherstone, 2011) may also be FCM if the rank valuation puts sufficient weight on assigning first choices (e.g., if $v(1) = (n+1) \cdot v(2)$). Neither of these mechanisms is stable or strategyproof. [10] The strategyproof and stable DA is not FCM, and neither is the strategyproof and Pareto efficient TTC.

In fact, FCM is incompatible with stability and with strategyproofness in general, as the following Proposition 1 shows.

Proposition 1. *If φ is FCM, then it is neither strategyproof nor stable.*

Proof. Suppose towards contradiction that φ is FCM and strategyproof and consider the following example: there are two schools $S = \{a, b\}$ with one seat each and there are three students $I = \{1, 2, 3\}$ with preference orders $a \; P_i \; b \; P_i \; \varnothing$ for all $i \in I$. Let $1 \succ_s 2 \succ_s 3$ for all $s \in S$. By feasibility, $\varphi(P, \succ)$ must leave one student unassigned; without loss of generality, let this student be 3 (i.e., $\varphi_i(P, \succ) = \varnothing$). Suppose that 3 reports $b \; P_3' \; a \; P_3' \; \varnothing$ instead. Then FCM implies that φ must assign student 3 to school b (i.e., $\varphi_3((P_3', P_{-3}), \succ) = b \; P_3 \; \varnothing = \varphi_3(P, \succ)$), a contradiction to strategyproofness.

Now, suppose φ is stable. In problem $((P_3', P_{-3}), \succ)$, there exists a unique stable matching in which 3 is unassigned and only one student is assigned to her top choice. Hence, φ cannot be FCM. □

Proposition 1 shows that imposing FCM in the design of school choice mechanisms may be costly: any mechanism with this property is unfair (in the sense that it may induce justified envy) and students may benefit by misrepresenting their preferences. Despite this observation, FCM holds a prominent place in the discussion of school choice mechanisms. Thus it would be interesting to identify a weaker form of stability that is compatible with FCM. We answer this question in the affirmative: like any FCM mechanism, BM may produce unstable matchings; but it is easy to see that no student forms a blocking pair with her first choice school. This motivates our definition of first choice-stability.

Definition 2. A matching μ is *first choice-stable* if there exists no blocking pair (i, s) such that s is i's first choice. A mechanism φ is *first choice-stable* (*FCS*) if for any problem $(P, \succ)$ the matching $\varphi(P, \succ)$ is first choice-stable.

[10] FCM is a natural relaxation of the axiom that a mechanism *favors higher ranks* (Kojima and Ünver, 2014) in the sense that a mechanism is FCM if and only if it favors *first* ranks.

Since stability implies FCS, DA is FCS. It is also straightforward to see that both BM and aBM satisfy this property because first choices are assigned only in the first round and by priority. In contrast, TTC and the Rank Value mechanism violate FCS.

The proof of Proposition 1 can be easily adapted to show that FCM is incompatible with k^{th} *choice-stability* for any rank $k \geqslant 2$.[11] In other words, FCM is incompatible with any other *rank-based* relaxation of stability except FCS. Thus, FCS is in a well-defined sense the strongest notion of stability that is compatible with FCM.

In this paper, we therefore focus on the class of mechanisms that are both FCM and FCS. To get a better intuition of what the members of this class look like, we now give a procedural characterization of these mechanisms.

Definition 3. A *first choices first-algorithm* is a procedure to compute a matching for any given problem $(P, \succ)$ in the following way:

Step 1. Each student i applies to her respective first choice school; each school s accepts applicants into open seats according to $\succ_s$ until there are no more applicants or all q_s seats are filled; all remaining applicants (if any) are rejected.

Step 2. Rejected students are matched to open seats by an arbitrary procedure but without changing the matchings that were made in Step 1.

A mechanism φ is a *first choices first-mechanism* (*FCF-mechanism*) if it is outcome-equivalent to some mechanism ψ that uses a first choice first-algorithm to determine matchings. In this case, ψ is called an *FCF-implementation* of φ.

Proposition 2. *A mechanism φ is FCM and FCS if and only if φ is an FCF-mechanism.*

Proof. Recall that $\varphi^k(P, \succ)$ denotes the set of students who receive their k^{th} choice under the mechanism φ at the problem $(R, \succ)$. We use the following Lemma 1 to prove sufficiency. A formal proof of the Lemma is given in Appendix A.

Lemma 1. *If φ is FCM and FCS, then $\varphi^1(P, \succ) = BM^1(P, \succ)$ for any problem $(P, \succ)$.*

In word, Lemma 1 means that the students who get their first choice under φ are the same as the students who get their first choice under BM.

Sufficiency. By Lemma 1, BM and φ assign the same students to their first choices. However, BM assigns all first choices in the first round and only in this round. Since

[11] A matching is k^{th} *choice-stable* if there exists no blocking pair (i, s) such that s is i's k^{th} choice.

FCF-mechanisms are not restricted in the way in which the assign schools beyond the first step, φ is an FCF-mechanism.

Necessity. If φ is an FCF-mechanism, then it assigns the same first choices as BM but possibly more. On the other hand, BM is first choice-maximizing, so φ must be first choice-maximizing as well. Suppose towards contradiction that φ is not first choice-stable. Then there exists a student $i \in I$ who does not get her first choice (say s), even though s either has an open seat or there exists a student $i' \in I$ who gets a seat at s but has lower priority at s than i. Since i ranks s first, i applies to s in the first step of the FCF-implementation of φ. There are two cases: if s has open seats, then i's application is successful, a contradiction. If i's application is rejected, there exist at least q_s other students who also ranked s in first position and have higher priority at s than i, a contradiction. □

Proposition 2 uncovers the key feature that makes a mechanism FCM and FCS, namely the way in which it matches students to their top choices. BM and aBM are two typical examples of mechanisms which exhibit this feature. However, FCF-mechanisms still have substantial freedom with respect to how they match any students who are not matched with their first choices.

Motivated by the observation that FCM is a common proxy for student welfare in the design of school choice mechanisms and that FCS is the strongest rank-based relaxation of stability that is compatible with FCM, we study the class of mechanisms that satisfy both properties in this paper. In the following, we provide insights about the mechanisms in this class in terms of their vulnerability of manipulation and Pareto efficiency (Section 4), Nash equilibrium outcomes when all students act strategically (Section 5) or when only some students do (Section 6).

4. Incentives under FCM and FCS Mechanisms

In this section we study the incentive properties of FCM and FCS mechanism. From Proposition 1 we have already learned that strategyproofness is out of the question. On the other hand, our intuition suggests that aBM has better incentive properties than BM because it prevents students from submitting futile applications and thereby removes some obvious opportunities to manipulate. One may hope that this advantage can be

formalized via the comparison of BM and aBM by their *vulnerability to manipulation*, a concept recently proposed by Pathak and Sönmez (2013). Surprisingly, we find that *all* FCM, FCS, and Pareto efficient mechanisms look alike under this comparison; in particular, this is true for BM and aBM.

Before we formalize this result, we restate the concept for comparing mechanisms by their vulnerability to manipulation.

Definition 4 (Manipulability)**.** A mechanism ψ is *manipulable by student i at the problem* $(P, \succ)$ if there exists a preference order $P_i' \neq P_i$ such that $\psi_i\left((P_i', P_{-i}), \succ\right) \; P_i \; \psi_i\left(P, \succ\right)$. ψ is *manipulable at* $(P, \succ)$ if it is manipulable by some student $i \in I$ at $(P, \succ)$.

Definition 5 (As Manipulable As-Relation)**.** For two mechanisms φ, ψ, ψ is *as manipulable as* φ if whenever φ is manipulable at some problem $(P, \succ)$, then ψ is also manipulable at $(P, \succ)$.

Our next main result (Theorem 1) is that *all* Pareto efficient, FCM, and FCS mechanisms are in the same equivalence class with respect to the as manipulable as-relation.

Theorem 1. *Let φ and ψ be two FCM, FCS, and Pareto efficient mechanisms. Then φ is as manipulable as ψ and vice versa.*

Proof. We use Lemma 1 from the proof of Proposition 2 as well as the following new Lemma 2. A formal proof of Lemma 2 is given in Appendix B.

Lemma 2. *Let φ be an FCM and FCS mechanism, let $(P, \succ)$ be any problem, let $A \subseteq I$ be a subset of the students, and let (P_A', P_{-A}) be a preference profile such that $choice_{P_i'}(1) = \varphi_i(P, \succ)$ for all $i \in A$. Then $\varphi_i\left((P_A', P_{-A}), \succ\right) = \varphi_i(P, \succ)$.*

In word, Lemma 2 shows that when a group of students simultaneously change their first choice to be the school that they receive under the matching $\varphi(P, \succ)$, then this students will receive those same schools under the new matching.[12]

Suppose, φ is manipulable at $(P, \succ)$, then there exists a student $i \in I$ and a preference profile P_i' such that $\varphi_i\left((P_i', P_{-i}), \succ\right) \; P_i \; \varphi_i\left(P, \succ\right)$. Without loss of generality, let $s =$

[12] This corresponds to a relaxed notion of Maskin monotonicity for FCM and FCS mechanisms: the preference profile (P_A', P_{-A}) is a monotonic transformation of P at $\varphi(P, \succ)$; and under any FCM and FCS mechanism the assignment of the students in A may not change. Observe that this property is independent of *rank respecting invariance*, a different relaxation of Maskin monotonicity that Kojima and Ünver (2014) used to characterize BM.

$\varphi_i((P_i', P_{-i}), \succ)$ be the most preferred school under P_i which i can get by misreporting under φ. By Lemma 2, we can choose P_i' such that s is ranked first and we observe that s cannot be the first choice under P_i. There are two possible cases:

Case $\varphi(P, \succ) = \psi(P, \succ)$. By Lemma 1, $\varphi^1((P_i', P_{-i}), \succ) = \psi^1((P_i', P_{-i}), \succ)$. Since $i \in \varphi^1((P_i', P_{-i}), \succ)$ and s is i's first choice under P_i', we have that $i \in \psi^1((P_i', P_{-i}), \succ)$ and $\psi_i((P_i', P_{-i}), \succ) = s$. Hence, $\psi_i((P_i', P_{-i}), \succ) \; P_i \; \psi_i(P, \succ)$, that is, i can manipulate ψ at the problem $(P, \succ)$.

Case $\varphi(P, \succ) \neq \psi(P, \succ)$. Since both $\varphi(P, \succ)$ and $\psi(P, \succ)$ are first choice-stable and Pareto efficient, there exists a student $i \in I$ such that $\varphi_i(P, \succ) \neq \psi_i(P, \succ)$ and $\varphi_i(P, \succ) \; P_i \; \psi_i(P, \succ)$. Let $s' = \varphi_i(P, \succ)$ and let P_i' be a preference order in which s' is ranked first. Lemmas 1 and 2 imply that $\varphi_i((P_i', P_{-i}), \succ) = s'$ and $\varphi^1((P_i', P_{-i}), \succ) = \psi^1((P_i', P_{-i}), \succ)$. Hence, $\psi_i((P_i', P_{-i}), \succ) = s' \; P_i \; \psi_i(P, \succ)$, that is, i can manipulate ψ at the problem $(P, \succ)$.

Symmetrically, it follows that if ψ is manipulable at some problem, then so is φ. □

Theorem 1 means that we cannot distinguish between any two FCM, FCS, and Pareto efficient mechanisms in terms of their incentive properties when using the comparison by vulnerability to manipulation.

Remark 1. A surprising consequence of Theorem 1 is that BM and aBM are manipulable at exactly the same problems. This raises the question whether the intuitive advantages of aBM over BM in terms of incentives can be justified in other ways. In Appendices F and G we show that a strict distinction cannot be recovered via the as *strongly manipulable as*-relation, and in the case of random priorities, even the weak distinction becomes inconclusive. These challenges illustrate the knife-edge nature of the strict distinctions that Dur (2015) and Mennle and Seuken (2015*b*) have obtained.

Theorem 1 leaves open questions about the incentive properties of FCM and FCS mechanisms that are not Pareto efficient; our next result, Proposition 3, addresses these questions.

Proposition 3. *Let φ be an FCM and FCS mechanism that is not Pareto efficient. Then there exists a mechanism ψ with the following properties:*

1. *ψ is FCM, FCS, and Pareto efficient,*

2. *ψ strictly Pareto dominates φ,*

3. *φ is as manipulable as ψ.*

The proof is constructive: if φ is not Pareto efficient, then there exists at least one problem $(P, \succ)$ where the matching $\varphi(P, \succ)$ is strictly Pareto dominated by some other matching μ. We can define a new mechanism $\overline{\varphi}$ to be exactly the same as φ except at the problem $(P, \succ)$ where we set $\overline{\varphi}(P, \succ) = \mu$. It is straightforward to show that $\overline{\varphi}$ is FCM and FCS and that it strictly Pareto dominates φ, and careful inspection reveals that it is manipulable only at problems where φ is also manipulable. Iterated application of this construction yields a Pareto efficient mechanism ψ that satisfies all three properties from Proposition 3. A formal proof of Proposition 3 is given in Appendix C.

Under the imposition of FCM and FCS, Proposition 3 reveals an alignment between Pareto efficiency and incentive properties: for any such mechanism that violates Pareto efficiency, we can dominate this mechanism in terms of student welfare without deteriorating the incentives for truthtelling. In this sense, administrators who are restricted to running FCM and FCS mechanisms have nothing to gain from sacrificing Pareto efficiency in exchange for better incentive properties.

5. Equilibrium under FCM and FCS Mechanisms

In the previous section we have identified the mechanisms that are least manipulable among the class of FCM and FCS mechanisms, namely those that are Pareto efficient. However, neither of these mechanisms is completely strategyproof. Therefore, we study equilibria of the induced preference revelation games under these mechanisms.[13] Formally, given a mechanism φ and a problem $(P, \succ)$, a preference profile $\widehat{P}$ is a *Nash equilibrium* (*of φ at $(P, \succ)$*) if for all students $i \in I$ and all preference orders $P'_i \neq \widehat{P}_i$ we have that $\varphi_i(\widehat{P}, \succ)\ R_i\ \varphi_i((P'_i, \widehat{P}_{-i}), \succ)$. In words, no student can benefit strictly by unilaterally deviating from the equilibrium.

For the special case of the Boston mechanism, Ergin and Sönmez (2006) showed that the Nash equilibrium outcomes correspond precisely to the set of matchings that are stable with respect to the true preferences. Our next result, Theorem 2, generalizes

[13] Following prior work, such as PRIOR WORK, we use *Nash equilibrium under complete information* as the main solution concept.

their finding: we show that the same characterization extends to all FCM and FCS mechanisms.

Theorem 2. *Given a problem $(P, \succ)$ and a mechanism φ that is FCM and FCS, a matching μ is stable at $(P, \succ)$ if and only if there exists a Nash equilibrium $\widehat{P}$ of φ at $(P, \succ)$ with $\mu = \varphi(\widehat{P}, \succ)$.*

Proof. Necessity. Suppose that $\widehat{P}$ is a Nash equilibrium and assume towards contradiction that the matching $\varphi(\widehat{P}, \succ)$ is not stable under the true preferences. Then, there exist a student $i \in I$ and a school $s \in S \cup \{\varnothing\}$ such that $s \; P_i \; \varphi_i(\widehat{P}, \succ)$ and

i. either $\left|\varphi_s(\widehat{P}, \succ)\right| < q_s$,

ii. or there exists $i' \in I$ such that $\varphi_{i'}(\widehat{P}, \succ) = s$ 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 $(P_i', \widehat{P}_{-i})$, i would receive s in the first round. By Lemma 1 we get $\varphi_i((P_i', \widehat{P}_{-i}), \succ) = \text{BM}_i((P_i', \widehat{P}_{-i}), \succ) = s$. Therefore, $\widehat{P}_i$ is not a best response for i, a contradiction to the assumption that $\widehat{P}$ is a Nash equilibrium.

Sufficiency. Let μ be a stable matching at $(P, \succ)$. Consider a preference profile $\widehat{P}$, where each student ranks the school she receives under μ first. Since μ is feasible, BM produces the matching μ in the first round when applied to $(\widehat{P}, \succ)$. Moreover, $\widehat{P}$ is a Nash equilibrium of BM by Theorem 1 in (Ergin and Sönmez, 2006).

By Lemma 1, $\varphi(\widehat{P}, \succ) = \mu$ as well. Assume towards contradiction that $\widehat{P}$ is not a Nash equilbrium of φ. Then there exists a student $i \in I$ and a preference order P_i' such that $s = \varphi_i((P_i', \widehat{P}_{-i}), \succ) \; P_i \; \varphi_i(\widehat{P}, \succ)$. By Lemma 2, we can assume that i ranks s first in P_i'. But then $s = \varphi_i((P_i', \widehat{P}_{-i}), \succ) = \text{BM}_i((P_i', \widehat{P}_{-i}), \succ) \; P_i \; \mu_i = \text{BM}_i(\widehat{P}, \succ)$ by Lemma 1, which implies that $\widehat{P}$ is not a Nash equilibrium of BM either, a contradiction. □

Theorem 2 implies that stability under true preferences is the characterizing feature of the possible Nash equilibrium outcomes of any FCM and FCS mechanism. Consequently, if students are expected to respond optimally to incentives, then all mechanisms in this class lead to the same set of potential outcomes in equilibrium. All FCM and FCS mechanisms are equivalent in this sense.

Remark 2. Jaramillo, Kayi and Klijn (2016) considered the class of rank priority mechanisms (Roth, 1991). Within this class they gave a characterization of the members

whose Nash equilibrium outcomes correspond to the set of stable matchings. Like our Theorem 2, this result generalizes the result of Ergin and Sönmez (2006) as BM falls under their characterization. However, aBM is not a rank priority mechanism; therefore, the result of Jaramillo, Kayi and Klijn (2016) does not imply our Theorem 2.

6. Equilibrium under FCM and FCS Mechanisms When Some Students Are Truthful

In practice, not all students may play their best responses, e.g., because they lack the strategic sophistication or the information that are necessary or they have an exogenous motivation to report truthfully. Following Pathak and Sönmez (2008), we consider school choice markets two groups of students: *non-strategic* students simply report their preferences truthfully while the *strategic* students play best responses. In such markets, we identify the equilibria of FCM and FCS mechanisms.

6.1. Identification of A-Nash Equilibrium

Pathak and Sönmez (2008) have shown that the equilibrium outcomes of BM in markets where only some students act strategically correspond to matchings that are stable but with respect to an augmented priority profile. We formally define these augmented priorities below.

Definition 6 (Augmented Priority)**.** Given a problem $(P, \succ)$ and a set of strategic students A, recall that $I(s, k, P)$ is the set of students who rank school s in k^{th} position under P. For each school $s \in S$, the *augmented priority order* $\hat{\succ}_s$ is constructed as follows (all undefined priorities are determined by transitivity):

1. $i \hat{\succ}_s j$ if $i \in A \cup I(s, 1, P)$ and $j \in I(s, 2, P) \backslash A$,
2. $i \hat{\succ}_s j$ if $i \in I(s, k, P) \backslash A$ and $j \in I(s, k+1, P) \backslash A$ for any $k \geqslant 2$,
3. $i \hat{\succ}_s j$ if $i \succ_s j$ and either $i, j \in A \cup I(s, 1, P)$ or $i, j \in I(s, k, P) \backslash A$ for any $k \geqslant 2$.

In words, augmented priorities give higher priority to strategic students at all schools, and non-strategic students receive higher priority at schools which they rank earlier in their preference order. Intuitively, this reflects the fact that strategic students can snatch

seats at popular schools by ranking them first and that non-strategic students can snatch seats in earlier rounds, even if they do not actually have priority at these schools. Next, we formally define stability of matchings with respect to the augmented priorities.

Definition 7 (Augmented-Stable)**.** Given a problem $(P, \succ)$, a set of strategic students $A \subseteq I$, and the augmented priority profile $\hat{\succ}$, a matching μ is *augmented-stable* if it is stable at $(P, \hat{\succ})$.

Finally, we need to formalize the notion of a Nash equilibrium in which only strategic students play best responses.

Definition 8 (A-Nash Equilibrium)**.** Given a problem $(P, \succ)$, a set of strategic students $A \subseteq I$, and a mechanism φ, a preference profile $(\hat{P}_A, P_{-A})$ is an *A-Nash equilibrium* (*of φ at $(P, \succ)$*) if for all strategic students $i \in A$ and any preference order P_i' we have that

$$\varphi_i((\hat{P}_A, P_{-A}), \succ) \; R_i \; \varphi_i((P_i', \hat{P}_{A \setminus \{i\}}, P_{-A}), \succ). \tag{2}$$

We can now restate the result of Pathak and Sönmez (2008).

Fact 1 (Proposition 1 in (Pathak and Sönmez, 2008))**.** *Given a problem $(P, \succ)$ and a set of strategic students $A \subseteq I$, a matching μ is augmented-stable if and only if there exists an A-Nash equilibrium $(\hat{P}_A, P_{-A})$ of BM at $(P, \succ)$ with $\mu = BM((\hat{P}_A, P_{-A}), \succ)$.*

Our next main result, Theorem 3, yields a similar characterization but for the entire class of FCM and FCS mechanisms. Observe, however, that FCM and FCS mechanisms are only restricted in the way in which they handle first choices (see Proposition 2), while BM is a specific mechanism, which exhaustively prescribes how to handle all possible preferences. To accommodate this freedom, we identify the equilibrium outcomes *up to equivalence from the perspective of the strategic students.* To make our statements formal, we require additional notation: given a problem $(P, \succ)$, a subset $A \subseteq I$ of the students, and matchings μ, ν, we say that μ is *A-equivalent to ν* if $\mu_i = \nu_i$ for all students $i \in A$, denoted $\mu =_A \nu$. With this we can state Theorem 3, our next main result.

Theorem 3. *Given a problem $(P, \succ)$ and a set of strategic students $A \subseteq I$, for any FCM and FCS mechanism φ the following hold:*

1. *If a preference profile $(\hat{P}_A, P_{-A})$ is an A-Nash equilibrium of φ at $(P, \succ)$, then there exists a augmented-stable matching μ such that μ and $\varphi(\hat{P}_A, P_{-A})$ are A-equivalent.*

2. *If μ is a augmented-stable matching, then there exists an A-Nash equilibrium $(\widehat{P}_A, P_{-A})$ of φ at $(P, \succ)$ such that μ and $\varphi(\widehat{P}_A, P_{-A})$ are A-equivalent.*

A formal proof of Theorem 3 is given in Appendix D.

Theorem 3 shows that augmented-stability describes (to a large extent) the equilibrium outcomes of FCM and FCS mechanisms when some students are strategic and others are not. Intuitively, the strategic students coordinate on an outcome in which none of their priorities are violated but (largely) ignoring the priorities and preferences of the non-strategic students. The only influence that non-strategic students have on the matching of strategic students is by receiving their (true) first choice in the first step of the FCF-implementation of the mechanism, which can block strategic students with lower priority from obtaining these seats.

6.2. Student Welfare in A-Dominant Nash Equilibrium

In this section, we focus on the particular A-Nash equilibrium outcome that all strategic student prefer and we study student welfare in these special equilibria. We first state Corollary 1 which follows from our Theorem 3 and the theory of stable matchings (Roth and Sotomayor, 1990).

Corollary 1. *We say that for two matchings μ, ν, μ A-dominates ν if $\mu_i \; R_i \; \nu_i$ for all students $i \in A$, denoted $\mu \; R_A \; \nu$. Then, given a problem $(P, \succ)$ and a set of strategic students $A \subseteq I$, let $\widehat{\mu}$ be the student-optimal augmented-stable matching. Then for any FCM and FCS mechanism φ,*

1. *there exists an A-Nash equilibrium $(\widehat{P}_A, P_{-A})$ with $\widehat{\mu} =_A \varphi((\widehat{P}_A, P_{-A}), \succ)$,*
2. *for any A-Nash equilibrium (P'_A, P_{-A}), $\widehat{\mu}$ weakly A-dominates $\varphi((P'_A, P_{-A}), \succ)$.*

With Corollary 1, we can identify the A-Nash equilibria that are uniformly preferred by all strategic students.

Definition 9 (A-Dominant Nash Equilibrium)**.** Let $\widehat{\mu}$ be the a student-optimal augmented-stable matching (as in Corollary 1). Then any A-Nash equilibrium $(\widehat{P}_A, P_{-A})$ with $\widehat{\mu} =_A \varphi((\widehat{P}_A, P_{-A}), \succ)$ is an *A-dominant Nash equilibrium*, and the matching $\varphi((\widehat{P}_A, P_{-A}), \succ)$ is called an *A-dominant Nash equilibrium matching*.

In the following we study student welfare in the A-dominant Nash equilibrium. If all students are strategic, then Theorem 2 implies that the set of possible equilibrium outcomes under any FCM and FCS mechanism corresponds to the set of stable matchings. Therefore, the A-dominant Nash equilibrium matching is precisely the student-optimal stable matching. However, the strategyproof Deferred Acceptance already implements this matching in dominant strategy equilibrium. Thus, for fully rational students, there is no real upside in terms of student welfare from using any non-strategyproof FCM and FCS mechanism instead of DA. On the other hand, if all students are truthful, then FCM and FCS mechanisms, like BM or aBM, maximize the number of first choices. Therefore, when we expect all students to report their preferences truthfully and if the number of first choices is the main metric for student welfare, then FCM and FCS mechanisms will produce more desirable matchings than DA.

Our next result identifies student welfare between these two extremes (i.e., fully strategic versus truthful students). Specifically, Theorem 4 describes the impact of a single student's decision on whether to act strategically or to report truthfully.

Theorem 4. *Given a problem $(P, \succ)$, an FCM and FCS mechanism φ, strategic students $A \subseteq I$, and some non-strategic student $i' \notin A$, let μ^A be an A-dominant Nash equilibrium matching and let $\mu^{A \cup \{i'\}}$ be an $(A \cup \{i'\})$-dominant Nash equilibrium matching. Then:*

1. *for all $i \in A$: $\mu_i^A \; R_i \; \mu_i^{A \cup \{i'\}}$,*
2. *for all $i \in I$: if $\mu_i^{A \cup \{i'\}} = choice_{P_i}(1)$, then $\mu_i^A = choice_{P_i}(1)$,*
3. *for i': $\mu_{i'}^{A \cup \{i'\}} \; R_{i'} \; \mu_{i'}^A$.*

The intuition behind the proof (in particular of statement 1.) is the following: suppose that i' claims her equilibrium seat in the first round of the FCF-implementation of φ. When she changes her report, that seat is up for grabs by other students. A number of students may be interested in this seat: a strategic students may prefer it to their current seat and a truthful student may rank it first. Among these interested students, one has highest priority. If this student is strategic, then she grabs the seat by adjusting her preference report; if she is truthful, the she simply takes it. We show that this causes a chain of improvements, where each step is beneficial to the participating students, and we show that this chain must end with a truthful student. The full proof of Theorem 4 is given in Appendix E.

Theorem 4 is of great relevance for market designers: consider a situation in which a group of students $A \cup \{i'\}$ is strategic and plays an $(A \cup \{i'\})$-dominant Nash equilibrium, but then i' changes her mind and switches to reporting truthfully instead. Subsequently, the remaining strategic students A adjust their reports and play an A-dominant Nash equilibrium. Then all remaining strategic students A weakly prefer the new matching to the old matching (statement 1.) but i' weakly prefers the old matching (statement 3.). In other words, the strategic students benefit unambiguously, but possibly at the expense of i'; and i' cannot benefit from this change, even when we take into account that strategic students adjust their preference reports in response to the change of report of i'. Finally, it is ambiguous whether the remaining non-strategic students (all but i') prefer the new or the old matching, with a notable exception: if any student, strategic or not, receives her true first choice under the old matching, then this student continues to receive her true first choice under the new matching as well (statement 2.). Thus, if a market designer cares about the number of students who receive their (true) first choice, then student welfare is lowest if all students are strategic, increases as more students become non-strategic, and peaks when all students are truthful in the respective dominant Nash equilibria.

7. Conclusion

We have observed that first choice-maximality (FCM) is a very natural and wide-spread objective in the design of school choice markets. Following the insight that first choice-stability (FCS) is the strongest rank-based relaxation of stability that is compatible with FCM, we have investigated the class of mechanisms that satisfy both properties.

We have shown that within this class, Pareto efficiency and incentives for truthtelling are congruent in the sense that we can dominate any Pareto *in*efficient member by another Pareto efficient member in terms of student welfare and in terms of vulnerability to manipulation. In addition, we have shown that all members of this class can be obtained by using a simple procedure: in an initial step, student are accepted to their first choices according to priority in the same way as under the Boston mechanism; in a second step, the rejected students are matched with the remaining open seats in an arbitrary fashion.

Taking the students' incentives into account, we have identified the Nash equilibrium

outcomes of the preference revelation game associated with any FCM and FCS mechanism in two situations: when all students act strategically, these outcomes correspond exactly to the matchings that are stable with respect the true preferences and priorities; and when only a subset of the students act strategically, then from the perspective of the strategic students, these outcomes correspond to the matchings that are stable with respect to the true preferences and augmented priorities. In particular, we have shown that more students, strategic or not, receive their true first choice (in the strategic-student optimal augmented-stable equilibrium) if fewer students act strategically.

Our results yield two new insights that are relevant for the design of school choice mechanisms: first, suppose that the share of students who receive their first choice is a primary objective and that a substantial share of the students is expected to report their preferences truthfully. Then our results teach us that, in equilibrium, FCM and FCS mechanisms yield more appealing outcomes than the strategyproof Deferred Acceptance mechanism. Second, suppose that we care in addition about student welfare in other ways, e.g., in terms of the *rank distribution* (Featherstone, 2011). We have observed that, in equilibrium, strategic students essentially self-select into receiving their school in the first step of the FCF-implementation of the mechanism. Thus, only truthful students actually participate in the second step. This means that the matchings in the second step can be made under the assumption that preference reports are in fact truthful (i.e., independent of the usual incentive constraints). For example, we may use a mechanism that optimizes the rank distribution, subject to the matchings made in the first step. This is particularly interesting because achieving optimal rank distributions is prohibitively incompatible with good incentive properties (Mennle and Seuken, 2015*a*).

References

Abdulkadiroğlu, Atila. 2011. "Generalized Matching for School Choice." Working Paper.

Abdulkadiroğlu, Atila, and Tayfun Sönmez. 2003. "School Choice: A Mechanism Design Approach." *American Economic Review*, 93(3): 729–747.

Alcalde, Jose. 1996. "Implementation of Stable Solutions to Marriage Problems." *Journal of Economic Theory*, 69(1): 240–254.

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.

Chen, Yan, and Onur Kesten. 2016. "Chinese College Admissions and School Choice Reforms: An Experimental Study." Working Paper.

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, Scott Kominers, Parag Pathak, and Tayfun Sönmez. 2014. "The Demise of Walk Zones in Boston: Priorities vs. Precedence in School Choice." Working Paper.

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. 2011. "A Rank-based Refinement of Ordinal Efficiency and a new (but Familiar) Class of Ordinal Assignment Mechanisms." Working Paper.

Gale, David, and Lloyd Shapley. 1962. "College Admissions and the Stability of Marriage." *The American Mathematical Monthly*, 69(1): 9–15.

Harless, Patrick. 2015. "A School Choice Compromise: Between Immediate and Deferred Acceptance." Working Paper.

Jaramillo, Paula, Cağatay 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. 2015*a*. "Partial Strategyproofness: An Axiomatic Approach to Relaxing Strategyproofness for Assignment Mechanisms." Working Paper.

Mennle, Timo, and Sven Seuken. 2015*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.

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. 1991. "A Natural Experiment in the Organization of Entry-level Labor Markets: Regional Markets for New Physicians and Surgeons in UK." , (81): 441–464.

Roth, Alvin, and Marilda Sotomayor. 1990. *Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis.* Cambridge University Press.

Appendix

A. Proof of Lemma 1

Proof of Lemma 1. BM is first choice-maximizing because it is obviously an FCF-mechanism and φ is first-choice maximizing by assumption. Therefore, $|\varphi^1(P,\succ)| = |\mathrm{BM}^1(P,\succ)|$. Hence, it suffices to show that for any student $i \in I$, if $i \notin \mathrm{BM}^1(P,\succ)$, then $i \notin \varphi^1(P,\succ)$. Suppose towards contradiction that there exists some student $i \in I$ such that $i \notin \mathrm{BM}^1(P,\succ)$ but $i \in \varphi^1(P,\succ)$. Let $s = \mathrm{choice}_{P_i}(1)$ be the school that i ranks first and consider the first round of BM. Since i did not receive s, there must exist at least q_s students whose first choice is s and who have higher priority at s than i. If i receives s under $\varphi(P,\succ)$, then at least one of these students does not receive s under $\varphi(P,\succ)$, a contradiction to first choice-stability of φ. □

B. Proof of Lemma 2

Proof of Lemma 2. First, assume that $\mathrm{choice}_{P_i}(1) \neq \varphi_i(P,\succ)$ for all $i \in A$. By Lemma 1, $\varphi^1(P,\succ) = \mathrm{BM}^1(P,\succ)$. Thus, all seats that were ultimately taken by some student from A under φ were not exhausted in the first round of BM. Therefore, $\mathrm{BM}_i\left((P'_A, P_{-A}),\succ\right) = \varphi_i\left(P,\succ\right)$ for all $i \in A$. However, $\mathrm{BM}^1\left((P'_A, P_{-A}),\succ\right) = \varphi^1\left((P'_A, P_{-A}),\succ\right)$ by Lemma 1, so that $\varphi_i\left(P,\succ\right) = \mathrm{BM}_i\left((P'_A, P_{-A}),\succ\right) = \varphi_i\left((P'_A, P_{-A}),\succ\right)$ for all $i \in A$.

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

C. Proof of Proposition 3

Proof of Proposition 3. Suppose $\varphi(P,\succ)$ is strictly Pareto dominated by some other matching μ at $(P,\succ)$. Define another mechanism $\overline{\varphi}$ to be the same as φ, except that $\overline{\varphi}(P,\succ) = \mu$.

First, observe that $\overline{\varphi}$ must be FCM because it Pareto improves over an FCM mechanism and therefore cannot assign strictly less first choices. However, since φ is FCM, $\overline{\varphi}$ also cannot assign strictly more first choices. Thus, $\overline{\varphi}$ assigns exactly the same first choices as φ, and therefore, it must be FCS.

Second, we verify that φ is as manipulable as $\overline{\varphi}$. Since φ and $\overline{\varphi}$ coincide for all problems except $(P, \succ)$, we only need to consider two kinds of situations:

1. $(P, \succ)$ is the true problem and some student $i \in I$ is considering some misreport $P_i' \neq P_i$,
2. $((P_i', P_{-i}), \succ)$ with $P_i' \neq P_i$ is the true problem and some student $i \in I$ is considering the particular misreport P_i.

In case 1, suppose that i can manipulate $\overline{\varphi}$ by reporting P_i' instead of reporting P_i truthfully in the problem $(P, \succ)$, i.e., $\overline{\varphi}_i((P_i', P_{-i}), \succ)\ P_i\ \overline{\varphi}_i(P, \succ)$. Since μ Pareto dominates $\varphi(P, \succ)$, we get

$$\varphi((P_i', P_{-i}), \succ) = \overline{\varphi}_i((P_i', P_{-i}), \succ)\ P_i\ \overline{\varphi}_i(P, \succ) = \mu\ R_i\ \varphi_i(P, \succ), \tag{3}$$

or equivalently, i can also manipulate φ by reporting P_i' in the problem $(P, \succ)$.

In case 2, suppose that i can manipulate $\overline{\varphi}$ by reporting P_i instead of reporting P_i' truthfully in the problem $((P_i', P_{-i}), \succ)$, i.e., $\mu_i = \overline{\varphi}_i(P, \succ)\ P_i'\ \overline{\varphi}_i((P_i', P_{-i}), \succ)$. Let $P_i'' \neq P_i$ be a preference order in which μ_i is ranked in first position. Since i can obtain μ_i by reporting P_i, i can also obtain μ_i by ranking it first (in particular by reporting P_i''). By construction $\mu_i = \overline{\varphi}_i((P_i'', P_{-i}), \succ) = \varphi_i((P_i'', P_{-i}), \succ)$, which implies that i can obtain μ_i by reporting P_i'' under φ in the problem $(P, \succ)$. Thus, φ must be manipulable in $(P, \succ)$.

Third, 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. $\square$

D. Proof of Theorem 3

Proof of Theorem 3. Statement 1. Let $(\widehat{P}_A, P_{-A})$ be an A-Nash equilibrium, let $\varphi_i((\widehat{P}_A, P_{-A}), \succ)$ be the school that student i receives in this equilibrium, and let P'_i be preference orders with $\text{choice}_{P'_i}(1) = \varphi_i((\widehat{P}_A, P_{-A}), \succ)$ for all $i \in A$. We first show the Claims 1 and 2.

Claim 1. *(P'_A, P_{-A}) is an A-Nash equilibrium and $\varphi((P'_A, P_{-A}), \succ)$ and $\varphi((\widehat{P}_A, P_{-A}), \succ)$ are A-equivalent.*

Proof of Claim 1. $\varphi_i((P'_A, P_{-A}), \succ)$ and $\varphi_i((\widehat{P}_A, P_{-A}), \succ)$ 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 $\tilde{P}_i$ such that $s = \varphi_i((\tilde{P}_i, P'_{A\setminus\{i\}}, P_{-A}), \succ)\ P_i\ \varphi_i((P'_A, P_{-A}), \succ) = \varphi_i((\widehat{P}_A, P_{-A}), \succ)$. By Lemma 2, we can assume $\text{choice}_{\tilde{P}_i}(1) = s$. Since $(\widehat{P}_A, P_{-A})$ 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 $(\widehat{P}_A, P_{-A})$, have higher priority than i at s, and therefore receive a seat at s in $\varphi((\widehat{P}_A, P_{-A}), \succ)$. By construction, these students still rank s first under $(\tilde{P}_i, P'_{A\setminus\{i\}}, P_{-A})$ and thus exhaust s in the first step of the FCF-implementation of φ. This contradicts the assumption that i could obtain s by deviating from P'_i to $\tilde{P}_i$. □

Claim 2. *$DA((P'_A, P_{-A}), \widehat{\succ})$ is augmented-stable at the true preferences P and A-equivalent to $\varphi(\widehat{P}_A, P_{-A})$.*

Proof of Claim 2. Let $\nu = \text{DA}((P'_A, P_{-A}), \widehat{\succ})$. To see A-equivalence, observe that under the problem $((P'_A, P_{-A}), \widehat{\succ})$ the first round of DA and φ (on $((P'_A, P_{-A}), \widehat{\succ})$) match exactly the same students to their reported first choices. Students who enter further rounds under DA are necessarily non-strategic (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 receive a seat at the school they ranked first under P'_A. With Claim 1, this implies that the matchings $\text{DA}((P'_A, P_{-A}), \widehat{\succ})$, $\varphi((P'_A, P_{-A}), \succ)$, and $\varphi((\widehat{P}_A, P_{-A}), \succ)$ are all A-equivalent.

To see augmented-stability of ν, assume towards contradiction that there exists some student $i \in I$ and some school $s \in S$ such that either s has an unfilled seat under ν or there exists another student $i' \in I$ with $s = \text{DA}_{i'}((P'_A, P_{-A}), \succ)\ P_i\ \text{DA}_i((P'_A, P_{-A}), \succ)$

and $i \succ_s i'$. If $i \notin A$ is a non-strategic student, then she cannot be part of a blocking pair. If $i \in A$ and s has an open 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 $((P'_A, P_{-A}), \succ)$. Consequently, i can obtain s (instead of y_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 $\hat{\succ}_s$ than i holds a seat at s, i can claim this seat in the same way, again a contradiction. □

In conclusion, in Claim 2 (using Claim 1), we have constructed the matching ν that is augmented-stable under the true preferences P and A-equivalent to $\varphi((\widehat{P}_A, P_{-A}), \succ)$.

Statement 2. Fix an arbitrary augmented-stable matching μ. Then let $(\widehat{P}_A, P_{-A})$ be a preference profile such that $\text{choice}_{\widehat{P}_i}(1) = \mu_i$ for all $i \in A$. Similar to statement 1, we prove this statement with two Claims 3 and 4.

Claim 3. *The matchings μ and $\varphi((\widehat{P}_A, P_{-A}), \succ)$ are A-equivalent.*

Proof of Claim 3. Assume towards contradiction that there exists a student $i \in A$ such that $\mu_i \neq \varphi_i((\widehat{P}_A, P_{-A}), \succ)$. Since i ranks μ_i first under $\widehat{P}_i$, μ_i must be exhausted in the first step of φ by students who rank μ_i as their first choice and have higher priority than i at μ_i under $\succ$. Any of these q_s students who is strategic (i.e., $i' \in A$) must receive μ_i under the matching μ, otherwise, i' would not rank μ_i first under $\widehat{P}_{i'}$; 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 receives μ_i in the matching μ, one of these q_s students, i' say, will not receive 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. □

Claim 4. *The preference profile $(\widehat{P}_A, P_{-A})$ is an A-Nash equilibrium.*

Proof of Claim 4. Assume towards contradiction that there exists some student $i \in A$ and a preference order P'_i such that $s = \varphi_i((P'_i, \widehat{P}_{A\setminus\{i\}}, P_{-A}), \succ)\ P_i\ \varphi_i((\widehat{P}_A, P_{-A}), \succ) = \mu_i$. 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. □

In conclusion, we have constructed the preference profile $(\widehat{P}_A, P_{-A})$ that is an A-Nash equilibrium and where the matching $\varphi((\widehat{P}_A, P_{-A}), \succ)$ is A-equivalent to μ. $\square$

E. Proof of Theorem 4

Proof of Theorem 4. Statement 1. We construct a preference profile P^* that is an A-Nash equilibrium and the matching $\varphi(P^*, \succ)$ weakly A-dominates the matching $\mu^{A\cup\{i'\}}$. The statement then follows from the fact that the matching $\varphi(P^*, \succ)$ is weakly A-dominated by any A-dominant Nash equilibrium matching by Corollary 1. The preference profile P^* is created as follows: let $P^0 = (P^0_{A\cup\{i'\}}, P_{-(A\cup\{i'\})})$ be the preference profile where each student $i \in A \cup \{i'\}$ ranks the school $\mu_i^{A\cup\{i'\}}$ first and all other students report truthfully. By Claim 1, P^0 is also an $(A \cup \{i'\})$-Nash equilibrium and the matching $\varphi(P^0, \succ)$ is $(A \cup \{i'\})$-equivalent to $\mu^{A\cup\{i'\}}$. Let $P^1 = (P^0_A, P_{i'}, P_{-(A\cup\{i'\})}) = (P^0_A, P_{-A})$ be the same preference profile expect that i' reports $P_{i'}$ truthfully instead of reporting $P^0_{i'}$.

Consider the simple case where $s' = \mu_{i'}^{A\cup\{i'\}}$ is the true first choice of i'; then P^1 is an A-Nash equilibrium and the matching $\varphi(P^1, \succ)$ is A-equivalent to $\mu^{A\cup\{i'\}}$. 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 P^1, i' is rejected from 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 $(A \cup \{i'\})$-Nash equilibrium). Let $\mu^1 = \varphi(P^1, \succ)$ be the resulting matching when i' reports her true preference order. Observe that $\mu_i^1 = \mu_i^{A\cup\{i'\}}$ for all strategic students $i \in A$. We now construct preference profiles $P^k, k = 2, \ldots, K$ in steps, where $P^* = P^K$ is constructed in the last step.

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

Step 1. Let A^k be the (possibly empty) set of strategic students who prefer s^k to their school under μ^k (i.e., $i \in A^1$ if $i \in A$ and $s^k \; P_i \; \mu_i^k$) and let I^k be the (possibly empty) set of non-strategic students who rank s^k as their first choice but do not receive it under μ^k (i.e., $i \in I^k$ if $i \notin A$, $s^k \; P_i \; s$ for all $s \neq s^k$, and $s^k \neq \mu_i^k$).

Step 2. If $A^k \cup I^k = \varnothing$, 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, $i^k \succ_{s^k} i$ for all $i \in A^k \cup I^k, i \neq i^k$).

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

Step 5. Else, $i^k \in A^k$. In this case, define the new preference profile P^{k+1} by setting

$$P_i^{k+1} = \begin{cases} P_i^k, & \text{if } i \in A \backslash \{i^k\}, \\ P_i^{k+1}, & \text{if } i = i^k, \\ P_i, & \text{if } i \in I \backslash A, \end{cases} \tag{4}$$

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

Step 6. Set $\mu^{k+1} = \varphi(P^{k+1}, \succ)$, then increment $k \rightsquigarrow k+1$ 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, P^* is an A-Nash equilibrium, and the matching $\varphi(P^*, \succ)$ A-dominates the matching $\mu^{A \cup \{i'\}}$.*

Proof of Claim 5. First, observe that for any $k \geqslant 1$, all strategic students $i \in A$ weakly prefer 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 and there exist only finitely many possible matchings. Therefore, the process ends after finitely many iterations. The same argument shows that $\varphi(P^*, \succ)$ A-dominates $\mu^{A \cup \{i'\}}$.

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 $P_i' \neq P_i^*$. Set $s = \varphi_i((P_A^*, P_{-A}), \succ)$ and $s' = \varphi_i((P_{A \backslash \{i\}}^*, P_i', P_{-A}), \succ)$. By Lemma 2 we can assume that i ranks s' as her first choice under P_i'.

Suppose that i can obtain s' because it is not exhausted in the first step when φ is applied to P^*. Since $i \in A$, i prefers s to $\mu_i^{A \cup \{i'\}}$ at least weakly, which implies $s' \; P_i \; \mu_i^{A \cup \{i'\}}$. 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' then i), a contradiction.

Conversely, suppose that i can obtain s' because it has higher priority at s' than some other student $\tilde{i}$ who receives s' under the matching $\varphi(P^*, \succ)$. $\tilde{i}$ could not have received s' in the matching $\varphi(P^1, \succ)$, otherwise, i could have benefited by deviating from the A-Nash equilibrium P^1. Thus, $\tilde{i}$ must have received s' at some step k 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. □

Statement 2. Assume towards contradiction that statement 2 is false, i.e, there exists a student i who receives her true first choice, say s, under $\mu^{A\cup\{i'\}}$, but does not receive s under μ^A. Then $i \notin A$, since all strategic students in A prefer μ^A to $\mu^{A\cup\{i'\}}$ by statement 1. On the other hand, there must exist some student $\tilde{i}$ who receives s under μ^A but not under $\mu^{A\cup\{i'\}}$ and this student must have priority over i at s. Since $\tilde{i}$ is not strategic, she ranks s first and therefore competes for s in the first step of φ. By first choice-stability of φ and the fact that i receives s in $\mu^{A\cup\{i'\}}$, $\tilde{i}$ must also receive s in $\mu^{A\cup\{i'\}}$, a contradiction.

Statement 3. Let $(\widehat{P}_A, P_{-A})$ be an A-dominant Nash equilibrium that leads to the matching μ^A. First, suppose that i' receives her first choice under μ^A (i.e., $\mu^A_{i'} = \text{choice}_{P_{i'}}(1)$). Then truthful reporting is a best response for i' to $(\widehat{P}_A, P_{-(A\cup\{i'\})})$. Since $(\widehat{P}_A, P_{-A})$ is an A-Nash equilibrium, it must be an $(A\cup\{i'\})$-Nash equilibrium. Therefore, the matching $\mu^{A\cup\{i'\}}$ $(A\cup\{i'\})$-dominates the matching μ^A, which implies that i' receives her first choice under $\mu^{A\cup\{i'\}}$ as well.

Next, suppose that i' does not receive her first choice under μ^A. Then i' is rejected from her first choice in the first step of φ applied to $((\widehat{P}_A, P_{-A}), \succ)$. 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 $((\widehat{P}_A, P_{-A}), \succ)$, s say. Now, let $\widehat{\succ}^A$ and $\widehat{\succ}^{A\cup\{i'\}}$ denote the augmented priority profiles that arise from $\succ$ when the sets of strategic students are given by A and $A\cup\{i'\}$, respectively, and let $\widehat{\succ}^{A,i'\sim s}$ be the same priority profile as $\widehat{\succ}^A$ except that i' has priority at s over all non-strategic students who do not rank s as their first choice. Next, we prove the two Claims 6 and 7.

Claim 6. *$DA_i(P, \widehat{\succ}^A) = \varphi_i((\widehat{P}_A, P_{-A}), \succ)$ for all $i \in A$ and those $i \notin A$ for whom $\varphi_i((\widehat{P}_A, P_{-A}), \succ) = choice_{P_i}(1)$.*

Proof of Claim 6. $\text{DA}(P, \widehat{\succ}^A)$ is the student-optimal augmented-stable matching (with respect to augmented priorities $\widehat{\succ}^A$). In particular, it is preferred by all strategic students

$i \in A$ to any other augmented-stable matching. Thus, $\mathrm{DA}(P, \hat{\succ}^A)$ is A-equivalent to the A-dominant Nash equilibrium matching $\varphi((\hat{P}_A, P_{-A}), \succ)$.

Now, assume towards contradiction that there exists some non-strategic student $i \notin A$ who receives her first choice $\tilde{s}$ under $\varphi((\hat{P}_A, P_{-A}), \succ)$ but is rejected from $\tilde{s}$ under $\mathrm{DA}(P, \hat{\succ}^A)$. By the above argument, the strategic students who are accepted at $\tilde{s}$ are exactly the same in both matchings. Thus, there exists a non-strategic student $\tilde{i}$ who gets $\tilde{s}$ under $\mathrm{DA}(P, \hat{\succ}^A)$ but does not get $\tilde{s}$ under $\varphi((\hat{P}_A, P_{-A}), \succ)$. 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. □

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

Proof of Claim 7. Consider the set of schools $\tilde{S}$ that are exhausted in the first step of φ when this mechanism is applied to $((\hat{P}_A, P_{-A}), \succ)$. By Claim 6, $\tilde{S}$ coincides with the set of schools that are exhausted exclusively by strategic students and those non-strategic students who rank them as their first choice under the matching $\mathrm{DA}(P, \hat{\succ}^A)$. 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 the matching $\mathrm{DA}(P, \hat{\succ}^{A,i'\sim s})$. Finally, consider the application process when DA is applied to P with priority profile $\hat{\succ}^{A,i'\sim s}$: i' will definitely apply to s at some point because she will have been rejected from all her more preferred schools. Under $\hat{\succ}^{A,i'\sim s}$, only strategic students and those non-strategic students who rank s as their first choice have priority over i' at s. Thus, i' can only be rejected from 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 from s. □

To complete the proof of statement 3 of Theorem 4, we observe that

$$\mu_{i'}^{A\cup\{i'\}} = \mathrm{DA}_{i'}(P, \hat{\succ}^{A\cup\{i'\}}) \; R_{i'} \; \mathrm{DA}_{i'}(P, \hat{\succ}^{A,i'\sim s}) = s \; R_{i'} \; \mu_{i'}^{A}, \tag{5}$$

where the first equality holds because the matching $\mathrm{DA}(P, \hat{\succ}^{A\cup\{i'\}})$ is $(A\cup\{i'\})$-equivalent to any $(A\cup\{i'\})$-dominant Nash equilibrium matching (Theorem 3 and Corollary 1), 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. $\square$

F. Failure of Strong Comparison

Definition 10 (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 $(P, \succ)$, then ψ is also manipulable by the same student i at the problem $(P, \succ)$.

Proposition 4. *BM is not as strongly manipulable as aBM.*

Proof. Consider a problem $(P, \succ)$ with five students $I = \{1, \ldots, 5\}$ and five schools $S = \{a, b, c, d, e\}$ with a single seat each. Let the preference profile be given by

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

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

Under BM and truthful reporting we have $\mathrm{BM}_5(P, \succ) = c$. 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 $\mathrm{aBM}_5(P, \succ) = e$. However, if student 5 ranks school c in first position (i.e., she reports $c \; P_5' \; \ldots$), then $\mathrm{aBM}_5((P_5', P_{-5}), \succ) = c$. 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 4 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 $I = \{1, \ldots, 4\}$ and four schools $S = \{a, b, c, d\}$ with a single seat each. Let the preference profile be given by

$$a \; P_1 \; \ldots, \tag{6}$$

$$b \; P_2 \; \ldots, \tag{7}$$

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

$$a \; P_4 \; c \; P_4 \; \ldots, \tag{9}$$

and let the priorities be the same at all schools such that $1 \succ_s \ldots \succ_s 4$ for all $s \in S$. Under BM and truthful reporting we have $\mathrm{BM}_3(P, \succ) = d$, but if student 3 reports $a \; P_3' \; c \; P_3' \; \ldots$ she will be assigned to c. Under the adaptive Boston mechanism and truthful reporting we have $\mathrm{aBM}_3(P, \succ) = c$. Since a and b are taken by students 1 and 2 in the first round, student 3 cannot improve her assignment by misreporting under aBM. □

G. Failure of Weak Comparison for Random Priorities

G.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 $u_i : S \to \mathbb{R}^+$ 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 $\varphi^{\mathbb{P}}$. $\varphi_i^{\mathbb{P}}(P)$ denotes the random assignment vector to student i if the reported preference profile is P. This is the ith 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 $\succ = (\succ_1, \ldots, \succ_1)$ 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).

G.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 11 (Manipulability (of Random Mechanisms))**.** $\psi^{\mathbb{P}}$ is *manipulable by student i at problem u* if there exists a preference order $P_i' \neq P_i$ such that $\mathbb{E}_{\psi_i^{\mathbb{P}}(P_i', P_{-i})}[u_i] > \mathbb{E}_{\psi_i^{\mathbb{P}}(\boldsymbol{P})}[u_i]$. $\psi^{\mathbb{P}}$ 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 12 (As Manipulable As-relation (for Random Mechanisms))**.** $\psi^{\mathbb{P}}$ 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.

G.3. Failure of the Comparison of $\mathrm{BM}^{\mathbb{U}}$ and $\mathrm{aBM}^{\mathbb{U}}$

Proposition 5.
- *$BM^{\mathbb{U}}$ is not as manipulable as $aBM^{\mathbb{U}}$,*
- *$aBM^{\mathbb{U}}$ is not as manipulable as $BM^{\mathbb{U}}$.*

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

$$\begin{aligned} &a\ P_1^{(1)}\ b\ P_1^{(1)}\ c\ P_1^{(1)}\ d, \\ &a\ P_{2,3}^{(1)}\ c\ P_{2,3}^{(1)}\ b\ P_{2,3}^{(1)}\ d, \\ &b\ P_4^{(1)}\ \dots. \end{aligned}$$

Student 1's assignment vector under $\mathrm{BM}^{\mathbb{U}}$ is $\mathrm{BM}_1^{\mathbb{U}}(P^{(1)}) = (1/3, 0, 0, 2/3)$ 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 $\mathrm{BM}_1^{\mathbb{U}}(P_1', P_{-1}^{(1)}) = (1/3, 0, 1/3, 1/3)$. Since $\mathrm{BM}_1^{\mathbb{U}}(P_1', P_{-1}^{(1)})$ first order-stochastically dominates $\mathrm{BM}_1^{\mathbb{U}}(P^{(1)})$ 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, $\mathrm{aBM}^{\mathbb{U}}$ is 1/3-partially strategyproof (Mennle and Seuken, 2015*a*). 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 $\mathrm{BM}^{\mathbb{U}}$ is manipulable (by student 1) at the problem $u^{(1)}$ but $\mathrm{aBM}^{\mathbb{U}}$ 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

$$a\ P_{1,2}^{(2)}\ e\ P_{1,2}^{(2)}\ c\ P_{1,2}^{(2)}\ d\ P_{1,2}^{(2)}\ f\ P_{1,2}^{(2)}\ b, \tag{10}$$

$$a\ P_{3,4}^{(2)}\ e\ P_{3,4}^{(2)}\ d\ P_{3,4}^{(2)}\ c\ P_{3,4}^{(2)}\ f\ P_{3,4}^{(2)}\ b, \tag{11}$$

$$b\ P_5^{(2)}\ c\ P_5^{(2)}\ \dots, \tag{12}$$

$$b\ P_5^{(2)}\ d\ P_5^{(2)}\ \dots. \tag{13}$$

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 $\mathrm{BM}^{\mathbb{U}}$. 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
$a \succ e \succ d \succ c \succ f \succ b$	-2.1
$a \succ c \succ e \succ d \succ f \succ b$	-0.4
$a \succ c \succ d \succ e \succ f \succ b$	-0.3
$a \succ d \succ e \succ c \succ f \succ b$	-9.5
$a \succ d \succ c \succ e \succ f \succ b$	-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 choice 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 $\mathrm{BM}^{\mathbb{U}}$ is not manipulable.

Second, we study the incentives to manipulate under $\mathrm{aBM}^{\mathbb{U}}$ for the same problem: under truthful reporting, student 1's assignment vector is $\mathrm{aBM}_1^{\mathbb{U}}(P^{(2)}) = (1/4, 1/4, 1/8, 1/8, 1/4, 0)$ for the schools a, e, c, d, f, b, respectively. If student 1 reported $P_1' : a \succ c \succ d \succ e \succ f \succ b$ instead, her assignment vector would be

$\mathrm{aBM}_1^{\mathbb{U}}(P_1', P_{-1}^{(2)}) = (1/4, 0, 71/120, 3/40, 1/12, 0)$. This means an increase in expected utility from 40.375 to 41.475, a strict improvement.

This concludes the proof. $\square$