When Manipulations are Harm[less]ful?

Mustafa Oğuz Afacan[*] Umut Mert Dur[†]

Abstract

We say that a mechanism is harmless if no student can ever misreport his preferences so that he does not hurt but someone else. We consider a large class of rules which includes the Boston, the agent-proposing deferred acceptance, and the school-proposing deferred acceptance mechanisms (sDA). In this large class, the sDA happens to the unique harmless mechanism. We next provide two axiomatic characterizations of the sDA. First, the sDA is the unique stable, non-bossy, and independent of irrelevant student mechanism. The last axiom is a weak variant of consistency. As harmlessness implies non-bossiness, the sDA is also the unique stable, harmless, and independent of irrelevant student mechanism.

JEL classification: C78, D63, D78, D47, D82.

Keywords: harmful manipulation, harmless, matching, mechanism, characterization.

[*]Faculty of Arts and Social Sciences, Sabanci University, Orhanli, Tuzla, 34956, Istanbul, Turkey; Phone: +90 216 483 9326; Fax: +90 216 483 9250 email: mafacan@sabanciuniv.edu.

[†]Poole School of Management, North Carolina State University, 27695, Raleigh, NC, USA; Phone +1 919 513 2878 email: udur@ncsu.edu.

1 Introduction

Initiated by Gale and Shapley (1962), matching theory has been fruitful both in theory and practice. Fortunately, theoretical findings have been put into practice, and some important real-life markets have been successfully redesigned by researchers under the guidance of theory. Examples include the National Residency Matching Program in the USA, placing doctors at hospitals, and The New York City and Boston student placement systems.[1]

In the redesign of matching markets, one of the main concerns has been agents' incentives in reporting their preferences to relevant matching intuitions. More specifically, both researchers and market practitioners put high premium on market designs giving right incentives to agents to submit their true preferences regardless of everything. This property is known as strategy-proofness in the literature. In terms of positive economics, in compliance with the Wilson doctrine, strategy-proofness is highly desirable because it provides participants a very simple optimal strategy. Moreover, from normative point of view, it sustains fairness among sophisticated and unsophisticated participants by leveling the playing field. The unfairness of non-strategy-proof Boston mechanism (henceforth, BM) has been well-documented by Roth et al. (2006). By using the 2001-02 Boston Public Schools (BPS) data, they show the presence of both sophisticated and unsophisticated students and observe that the latter (group) have indeed hurt the former.[2] In fact, the following memo from BPS Superintended Payzant to the School Committee on May 25, 2005 reveals the importance BPS staff gives to strategy-proofness due to its fairness aspect:

> *"A strategy-proof algorithm levels the playing field by diminishing the harm done*

[1] See Balinski and Sönmez (1999), Abdulkadiroğlu et al. (2005), Roth et al. (2005), and Roth and Peranson (1999).

[2] They show that many of unsophisticated students were unassigned, which would not have been the case if they had chosen a wise preference submission strategy. Similarly, by using data from Wake County School District, Dur et al. (2015) show that sophisticated students benefit under the BM at the expense of sincere students.

to parents who do not strategize or do not strategize well."

The above empirical finding of Roth et al. (2006) is confirmed both in theory and in lab by Pathak and Sönmez (2008) and Chen and Sönmez (2006). Especially, Pathak and Sönmez (2008) show that the Pareto-dominant Nash equilibrium under the BM favors sophisticated students at the expense of sincere ones, providing a theoretical support for the unfairness of the BM stemming from its lack of strategy-proofness.

In this study, especially motivated by the fairness implication of strategy-proofness, we consider manipulable mechanisms[3] and study whether or not a preference manipulation by a student can harm other students. Specifically, we say that a mechanism is harmless no other students are harmed if a student manipulates the mechanism.We first show that the school-proposing deferred acceptance mechanism(sDA) is harmless, whereas the commonly used agent-proposing deferred acceptance mechanism (aDA) is not.

Next, we conduct an axiomatic analysis to deepen our understanding of the set of stable and harmless mechanisms. We say that a mechanism is independent of irrelevant student whenever a new student who finds every school unacceptable joins the problem, the other students' assignments under the mechanism do not change. A mechanism is non-bossy (Satterthwaite and Sonnenschein (1981)) if no student can ever change someone else's assignment without changing his own assignment. We show that a mechanism is stable, non-bossy, and independent of irrelevant student if and only if it is the sDA. As harmlessness implies non-bossiness, as a corollary of this result, we also obtain that the sDA is the unique stable, harmless, and independent of irrelevant students mechanism. To the best of our knowledge, altoughsDA is widely used in real life matching practices,[4] there has been no axiomatization of the sDA other than being the

[3]We say that a mechanism is manipulable if a student can change the mechanism outcome without hurting herself. That is, manipulable mechanisms includes ones where at some problem, the misreporting student's assignment does not change but someone else. Such situations may call for strategic misreporting, because hurt students may incentivize misreporting through offering transfers

[4]Centralized college admission in Turkey, secondary school assignment in France and Finland, and high school assignment in Ireland are done through variants of sDA.

student-pessimal stable mechanism (Gale and Shapley (1962)); thereby those characterizations shed a further light on our sDA understanding.

Lastly, in order to investigate possibly other harmless mechanisms, we consider a very large class of mechanisms, including the BM, the aDA, and the sDA. A subclass of it, which is in use in China for college admissions, was first introduced by Chen and Kesten (2015). Within this very large class, the sDA happens to be the unique harmless rule as well.

2 Related Literature

This paper is broadly related to the extensive manipulations in matching markets literature. In two-sided matching markets where both sides are active agents, Roth (1982) shows that no stable mechanism is strategy-proof. On the other hand, whenever either side has commonly known preferences in the one-to-one matching environment, a stable and strategy-proof rule exists (Dubins and Freedman (1981) and Roth (1982)). Sönmez (1997, 1999) demonstrate that stable mechanisms are vulnerable to capacity and pre-arrangement manipulations, respectively. In contrast to stable rules, the BM and the Top Trading Cycles mechanism (attributed to David Gale) are immune to capacity manipulations (Kesten (2012)). Some of these negative results are overturned in large markets. Kojima and Pathak (2009) prove that under some regularity conditions, the scope of profitable preference and capacity manipulations diminishes under the aDA as the market becomes large.

Pathak and Sönmez (2008) conduct an equilibrium analysis under the BM with sophisticated and sincere students. Among other results, they report that sophisticated students weakly prefer the Pareto-dominant Nash equilibrium outcome of the BM to the dominant-strategy outcome of the aDA. Other papers on manipulations in matching markets include Kojima (2011, 2006), Ergin (2002), Konishi and Ünver (2006), Pathak and Sönmez (2013), and Afacan (2013b).

The paper is also related to the axiomatic characterization line of literature. Gale and Shapley (1962) show that the sDA is the worst stable mechanism for the student side, in other word, it is the student-pessimal stable mechanism. No paper has offered a different characterization of the sDA. This paper provides two axiomatizations of it in terms of non-bossiness and harmless manipulability. As opposed to the sDA, there are various axiomatizations of the aDA, including Kojima and Manea (2010), Morrill (2013a), Ehlers and Klaus (2014), Balinski and Sönmez (1999), and Alcalde and Barbera (1994). Other well-known and in use school choice mechanisms BM and Top Trading Cycles have been characterized in the literature as well by Kojima and Ünver (2014), Afacan (2013a), Abdulkadiroglu and Che (2010), Dur (2015), and Morrill (2013b).

3 Model

A **school choice problem** (Abdulkadiroğlu and Sönmez (2003)), or simply a problem, consists of:

- a finite set of students $I = \{i_1, i_2, ..., i_n\}$,
- a finite set of schools $S = \{s_1, s_2, ..., s_m\}$,
- a quota vector $q = (q_s)_{s\in S}$ where q_s is the number of available seats in school s,
- a list of preference relations $P = (P_i)_{i\in I}$ where P_i is the strict preference of student i over the schools and being unassigned option denoted by $\emptyset$,
- a list of priority orders $\succ = (\succ_s)_{s\in S}$ where $\succ_s$ is the strict priority relation of school s over I.

Let $q_\emptyset = \infty$. In all sections except Section 4.1 (Characterization section) we fix I,S, q, and $\succ$, and represent a problem with P. Let R_i be the at-least-as-good-as relation associated with P_i

for all $i \in I$. We write $\mathcal{P}$ for the set of all possible strict preferences of students. We say that school s is **acceptable** to student i if $sP_i\emptyset$. Otherwise, it is called unacceptable.

A **matching** $\mu : I \to S$ is a function such that $|\mu(i)| \leq 1$ and $|\mu^{-1}(s)| \leq q_s$ for all $i \in I$ and $s \in S$. Let $\mathcal{M}$ be the set of all matchings.

A matching $\mu \in \mathcal{M}$ **Pareto dominates** another matching $\nu \in \mathcal{M}$ if $\mu(i)R_i\nu(i)$ for each student $i \in I$ and $\mu(j)P_j\nu(j)$ for at least one student $j \in I$. A matching μ is **Pareto efficient** if there does not exist another matching $\nu \in \mathcal{M}$ which Pareto dominates μ.

A matching μ is **non-wasteful** if there exists no student school pair (i, s) such that $|\mu^{-1}(s)| < q_s$ and $sP_i\mu(i)$. A matching μ is **individually rational** if $\mu(i)R_i\emptyset$ for all $i \in I$. A matching μ is **fair** if there does not exist a student school pair (i, s) where $s\ P_i\ \mu_i$ and $i \succ_s j$ for some $j \in \mu^{-1}(s)$. A matching is **stable** if it is non-wasteful, individually rational, and fair. A stable matching μ is student-optimal stable matching if it Pareto dominates any other stable matching.

A **mechanism** Φ is a procedure which selects a matching for each problem. The matching selected by mechanism Φ in problem P is denoted by $\Phi(P)$ and the assignment of each student $i \in I$ is denoted by $\Phi_i(P)$.

A mechanism Φ is **Pareto efficient (stable)** if for any problem P its outcome $\Phi(P)$ is Pareto efficient (stable).

A mechanism Φ is **strategy-proof** if there do not exist a student $i \in I$ and a preference relation P' such that $\Phi_i(P', P_{-i})\ P_i\ \Phi_i(P)$. We say a mechanism is non-strategy-proof if it is not strategy-proof.

A mechanism Φ is **bossy** if there exist a problem P, student $i \in I$, and a preference relation P' such that $\Phi_i(P', P_{-i}) = \Phi_i(P)$ and $\Phi(P', P_{-i}) \neq \Phi(P)$. We say a mechanism is **non-bossy** if it is not bossy.

A mechanism Φ is **group-strategy-proof** if there exist no problem P, a group of students $I' \subseteq I$, and a false preference profile $P'_{I'} = (P'_i)_{i \in I'}$ for I' such that for every $i \in I'$, $\Phi_i(P'_{I'}, P_{-I'})\ R_i\ \Phi_i(P)$, with strictly holding for some student in I'. Pápai (2000) shows that a mechanism is group-strategy-proof if and only if it is non-bossy and strategy-proof.

A mechanism Φ is **manipulable** if there are a problem P, a student i and a preference relation P' such that $\Phi_i(P', P_{-i}) R_i \Phi_i(P)$ and $\Phi(P', P_{-i}) \neq \Phi(P)$. Our class of manipulable mechanisms includes ones where at some problem, the misreporting student's assignment does not change but someone else. Such situations may call for strategic misreporting, because badly affected students may incentivize misreporting through offering transfers. Note that, if a mechanism is non-strategy-proof and/or bossy, then it is manipulable.

Now we are ready to define the key notions that we use in our analysis. We say a mechanism Φ is **weakly harmfully manipulable** if there are a problem instance P, student i, and P' such that $\Phi_i(P', P_{-i}) R_i \Phi_i(P)$ and, for some student $j \neq i$, $\Phi_j(P) P_j \Phi_j(P', P_{-i})$. A mechanism ψ is **harmfully manipulable** if there are a problem instance P, student i, and P' such that $\Phi_i(P', P_{-i}) P_i \Phi_i(P)$ and, for some student $j \neq i$, $\Phi_j(P) P_j \Phi_j(P', P_{-i})$. A mechanism is **harmless** if it is not weakly harmfully manipulable. Similarly, a mechanism is **weakly harmless** if it is not harmfully manipulable.

4 Results

In the literature, mechanisms are compared based on their vulnerability to manipulation. In particular, a strategy-proof (or group-strategy-proof) mechanism has been considered mostly better than ones which are not strategy-proof (or group-strategy-proof). However, being a strategy-proof mechanism does not imply being a harmless mechanism.

Proposition 1 *(i). If a mechanism is group-strategy-proof, then it is harmless.*

(ii). If a mechanism is strategy-proof, then it is weakly harmless.

(iii). If a mechanism is bossy, then it is weakly harmfully manipulable.

Proof. Parts (i) and (ii) directly follow from the definition.

(iii). If a mechanism Φ is bossy, then there exist a problem P, a student i and preference relation P' such that $\Phi_i(P', P_{-i}) = \Phi_i(P)$ and $\Phi(P', P_{-i}) \neq \Phi(P)$. Then, there exists at least one student $j \in I \setminus \{i\}$ such that $\Phi_j(P', P_{-i}) \neq \Phi_j(P)$. Due to the strict preferences either $\Phi_j(P', P_{-i})\ P_j\ \Phi_j(P)$ or $\Phi_j(P)\ P_j\ \Phi_j(P', P_{-i})$. Hence, j will be harmed either in problem P when i reports P' or in problem (P', P_{-i}) when i reports P_i. ■

Since the set of group-strategy-proof mechanisms is a strict subset of the set of strategy-proof mechanisms,[5] Proposition 1 implies that some of the strategy-proof mechanisms can be weakly harmfully manipulable.

Proposition 2 *(i).If a manipulable mechanism is harmless, then it is Pareto inefficient.*
(ii). Given a mechanism Φ, if there exists a problem P such that $\Phi(P)$ is Pareto efficient and someone has a manipulation making him strictly better off, then Φ is harmfully manipulable.
(iii). Given a mechanism Φ, if there exists a problem P such that $\Phi(P)$ is Pareto efficient and someone has a manipulation changing the outcome but leaving his assignment unchanged, then Φ is weakly harmfully manipulable.

Proof. (i). Suppose not. That is, there exists a manipulable mechanism Φ which is harmless and Pareto efficient. Suppose under problem P, there exist a student i and a preference relation P' such that $\Phi(P)$ is Pareto efficient, $\Phi_i(P', P_{-i}) R_i \Phi_i(P)$ and $\Phi(P', P_{-i}) \neq \Phi(P)$. Since Φ is harmless, then $\Phi_j(P', P_{-i}) R_j \Phi_j(P)$ for all $j \in I \notin \{i\}$. Then, $\Phi_k(P', P_{-i}) R_k \Phi_k(P)$ for all $k \in I$ and $\Phi(P', P_{-i}) \neq \Phi(P)$. However, this contradicts with the Pareto efficiency of $\Phi(P)$.

(ii). Suppose not and assume that at problem P, student i strictly benefits from reporting P'. That is, $\Phi_i(P', P_{-i}) P_i \Phi_i(P)$ and $\Phi_j(P', P_{-i}) R_j \Phi_j(P)$ for all $j \in I \notin \{i\}$. Then, $\Phi(P', P_{-i})$ Pareto dominates $\Phi(P)$. However, this contradicts the Pareto efficiency of $\Phi(P)$.

(iii). As the same as above, suppose not and assume that at problem P, student i does not hurt by reporting P'. That is, $\Phi_i(P', P_{-i}) R_i \Phi_i(P)$ and $\Phi_j(P', P_{-i}) R_j \Phi_j(P)$ for all $j \in I \notin \{i\}$.

[5] The aDA is strategy-proof but not group-strategy-proof.

Then, $\Phi(P', P_{-i})$ Pareto dominates $\Phi(P)$. However, this contradicts the Pareto efficiency of $\Phi(P)$. ■

As immediate two corollaries to Proposition 2 are stated below.

Corollary 1 (i). *If a mechanism is Pareto efficient and someone has a manipulation at some problem making him strictly better off, then it is harmfully manipulable.*
(ii). *If a mechanism is Pareto efficient and someone has a manipulation at some problem leaving him indifferent, then it is weakly harmfully manipulable.*

In Corollary 1, we state that any Pareto efficient and manipulable mechanism is (weakly) harmfully manipulable; hence any such rule is not (weakly) harmless. A very well-known class mechanisms is stable mechanisms (Gale and Shapley (1962); Dubins and Freedman (1981); Roth (1982); Balinski and Sönmez (1999)). In this class, the aDA[6] is bossy and any other stable mechanism is not strategy-proof. Hence, any stable mechanism is manipulable. Since the aDA is bossy (Ergin (2002)), it is weakly harmfully manipulable. Moreover, because the aDA is strategy-proof, it is weakly harmless (Proposition 1) at the same time.

Corollary 2 *The aDA is weakly harmfully manipulable; thereby it is not harmless. Yet, the aDA is weakly harmless.*

Note that, neither Proposition 1 nor Proposition 2 says anything about whether the sDA[7] is harmful or not. In the following theorem, we show that the sDA is harmless.

Theorem 1 *The sDA is harmless.*

Proof. See the Appendix B. ■

[6] Formal definitions of the BM and the aDA are provided in Fact 2 in Section 4.2.
[7] The formal definition of the sDA is relegated to Appendix A.

4.1 Characterization

In Theorem 1, we show that the sDA is harmless. A natural question arises is whether there exists another harmless mechanism in the class of stable mechanisms. By using an intuitive and weak axiom, we prove that the sDA is the unique harmless and stable mechanism satisfying that weak property. In particular, we show a stronger result by using non-bossiness instead of harmlessness in our characterization. Before presenting the characterization result, we introduce a new axiom.

Consider a problem $(I, \succ, P)$ where there exists a student $i \in I$ preferring $\emptyset$ to all schools in S. We say student i is an **irrelevant student** in problem $(I, \succ, P)$. Let $\bar{\succ}$ be the priority order obtained by removing i from all schools' priority orders while keeping the relative order of the other students. A mechanism Φ is **independent of irrelevant student** if it selects the same outcome when the independent student is removed with all of his priorities: $\Phi_j(I, \succ, P) = \Phi_j(I \setminus \{i\}, \bar{\succ}, P_{-i})$ for all $j \in I \setminus \{i\}$.

Theorem 2 *The sDA is the unique mechanism which is stable, non-bossy and independent of irrelevant student.*

Proof. See the Appendix B. ■

The three axioms used in Theorem 2 are independent:

A mechanism which is non-bossy and independent of irrelevant student: Top Trading Cycles mechanism is non-bossy (Pápai (2000)) and independent of irrelevant student.

A mechanism which is stable and independent of irrelevant student: The aDA is stable and independent of irrelevant student.

A mechanism which is stable and non-bossy: A mechanism which selects the aDA outcome whenever the priority structure is acyclic otherwise selecting the sDA outcome is stable and non-bossy. The aDA becomes non-bossy under acyclic priorities (Ergin (2002)), and the sDA is already non-bossy (due to Theorem 1 and Proposition 1); hence the mechanism is non-bossy. Its stability directly comes from its definition. However, as the rule changes depending on the priority structure, the presence of an irrelevant student's priority matters here; thereby it is not independent of irrelevant student.

Theorem 2 implies that any stable mechanism satisfying independence of irrelevant agent which is not equivalent to the sDA is bossy. When we combine this observation with Proposition 1 and Theorem 1, we have the following corollary.

Corollary 3 *The sDA is the unique mechanism which is stable, harmless, and independent of irrelevant student.*

In the rest of the paper, we consider a very large class of rules, including the BM, the aDA, and the sDA, and investigate their harmlessness property. It will turn out that the even in this large class of rules, the sDA is the unique harmless one.

4.2 The Class of Student-Proposing Parallel Mechanisms

In this subsection, we consider a particular class of mechanisms, which is first considered by Chen and Kesten (2015). They call it "application-rejection" class of mechanisms where each member is identified by a single parameter $e \in \{1, 2, .., \infty\}$. Here, we call this class the student-proposing parallel mechanisms.

In what follows now, we first describe that class of mechanisms. For a given e,

Round 1.

* Each student applies to her first choice. Among the current applicants, each school tentatively accepts the highest priority ones up to its capacity and rejects the rest.

* Each rejected student in the previous step and yet to apply her e^{th} choice applies to her next best choice. Any other rejected student in the previous step remains unassigned in this round. Among the tentatively accepted and current applicants, each school tentatively accepts the highest priority ones up to its capacity and rejects the rest.

* This round terminates whenever every student is tentatively assigned or had all of her top e applications rejected. By the end of the round, all tentative assignments are finalized, and tentatively assigned students leave the problem with their assigned seats.

In general,

Round $t \geq 2$

* Each unassigned student from Round t-1 applies to his next best acceptable school. Each school tentatively accepts the highest priority students among the applicants up to its left capacity and rejects the rest.

*Each rejected student in the previous step and yet to apply his acceptable $(e \times t)^{th}$ choice applies to his next best acceptable school. Any other rejected student in the previous step remains unassigned in this round. Among the current applicants and the tentatively accepted ones, each school tentatively accepts the highest priority ones up to its available capacity and rejects the rest.

* The round terminates whenever every student is tentatively assigned or had all of her top $(e \times t)^{th}$ applications rejected. By the end of the round, all tentative assignments are finalized, and tentatively assigned students leave the problem with their assigned seats.

The algorithm terminates whenever every student is matched or had rejection from all of his acceptable choices.

By following the notations of Chen and Kesten (2015), let us write Φ^e for the above mechanism.

Fact 1 *[see Chen and Kesten (2015)]*

(i) For $e = 1$, $\Phi^e = BM$.

(ii) For $e = \infty$, $\Phi^e = aDA$

Chen and Kesten (2015) show that for any finite e, Φ^e is non-strategy-proof. In the following result, we prove that for any finite e, whenever a student has profitable deviation under Φ^e at a problem, then at least one other student is hurt. That is, Φ^e is harmfully manipulable for any finite e.

Theorem 3 *For any finite e, if there exists a student who can deviate profitably under Φ^e at problem P, then it is harmfully manipulable at P.*

Proof. See the Appendix B. ■

Theorem 3 above, along with the fact that for any finite e Φ^e is non-strategy-proof, shows that Φ^e is harmfully manipulable.

Corollary 4 *For any finite e, Φ^e is harmfully manipulable.*

The only strategy-proof member of the student-proposing class of parallel mechanisms is the aDA. This, along with the above results, shows that the aDA is the only rule in that class which is not harmfully manipulable. However, it is weakly harmfully manipulable (see Corollary 2). Therefore, any member of the student-proposing class of parallel mechanisms is weakly harmfully manipulable.

Next we introduce a new notion which helps us to compare any two harmfully manipulable mechanisms a la Pathak and Sönmez (2013).

Definition 1 *A mechanism Φ is more (weakly) harmfully manipulable than ϕ if, whenever the latter is (weakly) harmfully manipulable at a problem P, then so is the former at P; yet the converse is not true.*

By using this notion, in the next result, we show that as e decreases Φ^e becomes more harmfully manipulable.

Proposition 3 *Φ^e is more harmfully manipulable than $\Phi^{e'}$ for any $e < e'$.*

Proof. The proof follows from Theorem 3 and Theorem 1 of Chen and Kesten (2015). In particular, Theorem 1 of Chen and Kesten (2015) shows that Φ^e is more manipulable than $\Phi^{e'}$ for any $e < e'$ in the sense that whenever some student has profitable deviation at some problem P under $\Phi^{e'}$, then the same is true for Φ^e at the same problem P, while the converse is not true. This result, along with the Theorem 3, proves Proposition 3. ■

However, we cannot compare the mechanisms in terms of weakly harmful manipulability.

Proposition 4 *There is no comparison in the above sense for the weak type of harmful manipulation.*

Proof. See Appendix B. ■

4.3 The Class of School-Proposing Parallel Mechanisms

In this subsection, we introduce a new class of mechanisms including the sDA.

As in the above student-proposing class of mechanisms, each member is specified with a parameter $e \in \{1, ..., \infty\}$. For a given $e \in \{1, ..., \infty\}$, we define ϕ^e as follows:

Round 1.

* Schools offer to their respective top choice student. Each offer receiving student tentatively holds the best offer and rejects the rest.

* Each rejected school in the previous step or still having an available seat after giving one seat to the offer accepting student in the previous step proposes to its next best student if it did not offer to its top e^{th} choice. Each offer receiving student tentatively accepts the best offer among the current offers and the tentatively held one and rejects the rest.

* The round terminates whenever every school offers to its top e choice or had as many accepted offers as its capacity. The tentatively held offers are finalized and the students holding offers leave the problem with their assigned seats. Each school capacity is decreased by the

number of students getting seats from it.

In general,

Round $t \geq 2$

* Each school still having a left capacity offers to its next best remaining student. Each offer receiving student tentatively holds the best one and rejects the rest.

* Each rejected school in the previous step or still having an available seat after giving one seat to the offer accepting student in the previous step offers to its next best student if it did not offer to its top $(e \times t)^{th}$ choice. Each offer receiving student tentatively accepts the best offer among the current offers and the tentatively held one and rejects the rest.

* The round terminates whenever every school offers to its top $(e \times t)^{th}$ choice or had as many accepted offers as its capacity. The tentatively held offers are finalized and the students holding offers leave the problem with their assigned seats. Each school capacity is decreased by the numbers of students getting seats from it.

The algorithm terminates whenever every school fills out its capacity or made offer to every student. We call this class of mechanisms, where each member is induced by different e, "School-Proposing Parallel Mechanisms.

In the following proposition, we show that for $e = \infty$, we obtain the sDA.

Proposition 5 *$\phi^{\infty} = sDA$.*

Proof. In the above class, schools simultaneously make offers in decreasing order of their priorities with a difference that they propose to a single student at each step. However, at the end, schools do propose to the same sets of students under both ϕ^{∞} and the sDA; hence they are the same. ■

sDA is a member of the above class of school-proposing parallel mechanisms; and it is known that the sDA is not strategy-proof. Indeed, the following result shows that no member of this class is strategy-proof.

Proposition 6 *For any e, ϕ^e is non-strategy-proof.*

Proof.

For a finite e, let us consider a problem instance consisting of $I = \{i_1, ..., i_{e+2}\}$ and $S = \{s_1, s_2\}$, with $q_{s_1} = e + 1$ and $q_{s_2} = 1$. The priority orders are such that (i) student i_{e+1} is the top $(e+1)^{th}$ priority student at school s_1, and (ii) student i_{e+1} is the top priority one at school s_2. Student preferences are such that every school is acceptable to every student, and student i_{e+1} prefers school s_1 to school s_2.

Then, mechanism ϕ^e at the true preference profile places student i_{e+1} at school s_2. Now, let student i_{e+1} misreport his preferences by declaring school s_1 as the only acceptable school. At this false preference profile, mechanism ϕ^e places student i_{e+1} at school s_1; thereby he would be better off.

The above shows that ϕ^e is non-strategy-proof for any finite e. From Proposition 6, we already know that sDA coincides with ϕ^∞; hence for $e = \infty$, it is non-strategy-proof as well, which finishes the proof.

■

Now, we are ready to show that sDA is the unique harmless mechanism in this class.

Theorem 4 *For any finite e, ϕ^e is harmfully manipulable; thereby it is not even weakly harmless.*

Proof. We prove by constructing an example for each finite e. There are $e + 2$ schools and $e+2$ students: $S = \{s_1, s_2, ..., s_{e+2}\}$ and $I = \{i_1, i_2, ..., i_{e+2}\}$. $q_s = 1$ for all $s \in S$. For $k \leq e+1$, each school s_k ranks student i_k at the top: $i_k \succ_{s_k} i$ for all $i \in I \setminus \{i_k\}$. The priority order of school s_{e+2} is given as: $i_1 \succ_{s_{e+2}} i_2 \succ_{s_{e+2}} ... \succ_{s_{e+2}} i_{e+2}$. For all $k \leq e$, each student i_k considers only s_k acceptable, i.e., $\emptyset P_{i_k} s$ for all $s \in S \setminus \{s_k\}$. The preferences of student i_{e+1} and i_{e+2} are given as follows: $s_{e+2} P_i s_{e+1} P_i \emptyset P_{i_{e+1}} s$ for all $s \in S \setminus \{s_{e+1}, s_{e+2}\}$ and $i \in \{i_{e+1}, i_{e+2}\}$.

Under true-telling, in the first round, the offer of each school s_k is accepted by i_k for any $k \leq e+1$. In the following round(s), s_{e+2} proposes to i_{e+2} and i_{e+2} is assigned to s_{e+2}.

However, if i_{e+1} reports only s_{e+2} as acceptable, then i_{e+1} will be assigned to s_{e+2} and i_{e+2}will be assigned to s_{e+1}. Hence, i_{e+2} is hurt when i_{e+1} manipulates the mechanism. ■

Remark 1. In our whole analysis above, we look at whether any student is hurt whenever someone misreports his preferences while all the rest is sincere. However, this analysis does not say anything for the case where more than one student deviates from truth telling while assuming that the rest sticks to their true preferences. In this remark, we will formally deal with this question and show that the sDA is the unique rule among the class of rules we consider above such that no student is hurt whenever more than one student deviates from truth telling while assuming that the rest sticks to their true preferences.

For a mechanism Φ, and any preference profile P, let $B_i(P_{-i}, \Phi) = \{P'_i \in \mathcal{P} : \Phi_i(P'_i, P_{-i}) R_i \Phi_i(P)\}$. That is, it is the set of strategies (referring to preferences as strategies as we consider game among students in this section) of student i giving at least as good outcome as the true preferences while all the others are sincere under Φ. Note that $P_i \in B_i(P_{-i}, \Phi)$. We say that a mechanism Φ is strongly harmless if there exist no problem P, student j, and P' such that $\Phi_j(P) P_j \Phi_j(P')$ where $P'_i \in B_i(P_{-i}, \Phi)$ for any $i \in I$. It is easy to see that strong harmlessness implies harmlessness. For any $e \in \{1, .., \infty\}$, let ψ^e and ϕ^e be the student and school-proposing parallel rule corresponding to the given e, respectively. As ψ^e for any e, and ϕ^e for any finite e are at least weakly harmfully manipulable, they are in particular not strongly harmless. However, the sDA happens to be strongly harmless as well.

Proposition 7 *$sDA_i(P') R_i sDA_i(P)$ where $P'_i \in B_i(P_{-i})$ for any $i \in I$.*

Proof. See Appendix B. ■

5 Conclusion

This paper concerns with the cost of manipulation in terms of harm given to sincere students. To this end, we say that a mechanism is (weakly) harmless if no student ever can misreport his preferences so that he is (better off) not worse off but someone else (becomes worse off). We consider a very large class of rules admitting the well-known BM, the aDA, and the sDA. In this class, the sDA turns out to be the unique harmless rule. Next, in order to improve our understanding of the class of harmless and stable rules, we provide an axiomatic characterization of the sDA. The sDA is the unique stable and harmless mechanism satisfying a very mild consistency requirement.

Appendix A

Below describes the sDA.

Round 1. Each school offers to the top priority student group up to its capacity. Each offer receiving student tentatively accepts the most preferred one and rejects the rest.

In general,

Round k. Each rejected school offers to the next top priority student group for its rejected seats. Each offer receiving student tentatively accepts the most preferred offer among the current ones and the tentatively held one in the previous round and rejects the rest.

The sDA terminates whenever every school tentatively fills out its all seats or has all offers exhausted. The assignments at the terminal round realize as the final outcome of the sDA. The student-offering version of this algorithm is the aDA.

Appendix B

Proof. [Proof of Theorem 1] Let us consider a problem P where there exist a student $i \in I$ and a preference order P_i' such that $sDA_i(P_i', P_{-i}) R_i sDA_i(P)$. We claim that no student is worse off at $P' = (P_i', P_{-i})$, i.e., $sDA_j(P') R_j sDA_j(P)$ for all $j \in I$. Let ψ denote the sDA and $\psi_i(P') = a$ and $\psi_i(P) = b$ (schools a and b may be the same).

Let P_i'' be the preference relation where the only acceptable school is a and $P'' = (P_i'', P_{-i})$.

Lemma 1 $\psi(P'') = \psi(P')$.

Proof. [Proof of Lemma 1]For the proof of this Lemmata, we will show that students reject the same set of schools in the course of ψ at both problems P' and P''. This in turn implies that students receives the same set of offers at both problems; thereby the outcomes are the same.

Let $O_s^k(P')$ and $O_s^k(P'')$ be the set of offers school s makes in step k at problems P' and P'', respectively. By definition, $O_s^1(P') = O_s^1(P'')$ for any $s \in S$. Then, again by definition, any student other than i rejects the same set of offers while student i may reject more offers at P'' in the first round. As in the first round, no rejected offer at P' is tentatively accepted at P'', hence $O_s^2(P') \subseteq O_s^2(P'')$ for any $s \in S$. The same arguments above applies to the second stage offers and so on. At the end of the ψ at problem P'', no student fails to receive an offer he gets at P'. This shows that $\psi_j(P'')R_j\psi_j(P')$ for any student $j \in N$. Hence, by the definition of P'', we have $\psi_i(P'') = \psi_i(P') = a$.

We now claim that $\psi_j(P'') = \psi_j(P')$ for any $j \in I \setminus \{i\}$ as well. Assume for a contradiction that for some student $j \in I \setminus \{i\}$, $\psi_j(P'')P_j\psi_j(P')$. From above, we know that $O_s^1(P') = O_s^1(P'')$ for any $s \in S$. If student i rejects the same set of offers in the first round, then we would have $O_s^2(P') = O_s^2(P'')$ for any $s \in S$ as well (note that the preferences of the other students are the same). Similarly, if he rejects the same set of offers in the second stage, then we would have $O_s^3(P') = O_s^3(P'')$ for any $s \in S$ and so on and so forth. In this case, the outcomes would have been the same. Therefore, our supposition implies that there exists some stage at which student i rejects an offer at P'', yet he does not reject it at P'. Moreover, he does not reject it at any later stage at P' as well, because if he were to reject, then that school would make offer to the next priority students and so on; thereby nothing would change in the working of ψ. This, however, contradicts $\psi_i(P'') = \psi_i(P') = a$. ■

Lemma 1 shows that $\psi(P'') = \psi(P')$. It is now enough to show that $\psi_j(P'')R_j\psi_j(P)$ for each student $j \in I$. We already have $\psi_i(P'')R_i\psi_i(P)$. By the definition of P'' and by the same arguments as above, we have $O_s^k(P) \subseteq O_s^k(P'')$ for every step k and every school $s \in S$. This easily shows that $\psi_j(P'')R_j\psi_j(P)$, which completes the proof.

■

Proof. [Proof of Theorem 2] We first show that the sDA satisfies non-bossiness and independence of irrelevant student.

Non-bossiness: Suppose the sDA is bossy. By Proposition 1, the sDA is weakly harmfully manipulable and therefore it is not harmless. However this contradicts with Theorem 1. Hence, the sDA is non-bossy.

Independence of irrelevant student: Since in any problem the sDA selects an individually rational matching, one can alternatively define the sDA in which each school $s \in S$ does not offer to the students considering s unacceptable. Under this definition, removal of the irrelevant student does not affect the matching selected by the sDA. Hence, the sDA is independent of irrelevant student.

Uniqueness: Next we show that there does not exist another mechanism satisfying these three axioms. On the contrary, let ϕ be a mechanism satisfying all the desired properties and it selects a different outcome than school optimal stable matching in problem $(I, \succ, P)$. Denote the school optimal stable matching in problem $(I, \succ, P)$ and $\phi(I, \succ, P)$ with μ and ν, respectively. Since $\nu \neq \mu$, both μ and ν are stable and the fact that any other stable matching is weakly preferred to μ by all students, $\nu(i) R_i \mu(i)$ for all $i \in I$ and $\nu(j) P_j \mu(j)$ for some student $j \in I$. Moreover, due to rural hospital theorem, there exists at least two distinct students j_1 and j_2 strictly preferring ν to μ. Let $\mu(j_1) = s$ and $\nu(j_1) = \tilde{s}$. Due to stability, $|\mu^{-1}(\tilde{s})| = q_{\tilde{s}}$ and $k \succ_{\tilde{s}} j_1$ for all $k \in \mu^{-1}(\tilde{s})$. That is, there exist at least $q_{\tilde{s}}$ students in I with higher priority than j_1 for school $\tilde{s}$ according to $\succ_{\tilde{s}}$.

Now we add a student $\bar{i}$ to problem $(I, \succ, P)$ and construct a new problem $(\bar{I}, \bar{\succ}, \bar{P})$ such that

- $\bar{I} = I \cup \{\bar{i}\}$,
- $\bar{P}_{\bar{i}} : \emptyset \bar{P}_{\bar{i}} s$ for all $s \in S$ and $\bar{P}_i = P_i$ for all $i \in I$,
- $i \succ_s j \iff i \bar{\succ}_s j$ for all $i, j \in I$ and $s \in S$,
- $i \bar{\succ}_s \bar{i}$ for all $i \in I$ and $s \in S \setminus \{\tilde{s}\}$,
- $i \succ_{\tilde{s}} j_1 \iff i \bar{\succ}_{\tilde{s}} \bar{i}$ for all $i \in I$.

Due to independence of irrelevant student, $\phi_i(\bar{I}, \bar{\succ}, \bar{P}) = \nu(i)$ for all students in $i \in I$. Let $\bar{\mu}$ be a matching such that $\bar{\mu}(i) = \mu(i)$ for all $i \in I$ and $\bar{\mu}(\bar{i}) = \emptyset$. One can easily verify that $\bar{\mu}$ is the school optimal stable matching in problem $(\bar{I}, \bar{\succ}, \bar{P})$.

Consider preference profile $\hat{P}_{\bar{i}} : \tilde{s}\hat{P}_{\bar{i}}\emptyset\hat{P}_{\bar{i}}s$ for all $s \in S \setminus \tilde{s}$. In problem $(\bar{I}, \bar{\succ}, (\hat{P}_{\bar{i}}, \bar{P}_{-\bar{i}}))$ $\bar{\mu}$ will be the outcome of the sDA (j_1 does not get offer from $\tilde{s}$ in problem $(I, \succ, P)$ and therefore $\bar{i}$ does not get offer from $\tilde{s}$ in problem $(\bar{I}, \bar{\succ}, (\hat{P}_{\bar{i}}, \bar{P}_{-\bar{i}})))$ hence it is the school optimal stable matching of problem $(\bar{I}, \bar{\succ}, (\hat{P}_{\bar{i}}, \bar{P}_{-\bar{i}}))$. Due to rural hospital theorem, the set of students assigned to a school in any stable in problem $(\bar{I}, \bar{\succ}, (\hat{P}_{\bar{i}}, \bar{P}_{-\bar{i}})))$ and $(I, \succ, P)$ and $(\bar{I}, \bar{\succ}, \bar{P})$ is the same and each school fills the same number of seats under any stable matching of these three problems. That is, $\phi_{\bar{i}}(\bar{I}, \bar{\succ}, (\hat{P}_{\bar{i}}, \bar{P}_{-\bar{i}})) = \emptyset$. Due to stability, $\phi_{j_1}(\bar{I}, \bar{\succ}, (\hat{P}_{\bar{i}}, \bar{P}_{-\bar{i}})) \neq \tilde{s}$. Otherwise, priority of $\bar{i}$ at $\tilde{s}$ would be violated. Hence, $\phi(\bar{I}, \bar{\succ}, (\hat{P}_{\bar{i}}, \bar{P}_{-\bar{i}})) \neq \phi(\bar{I}, \bar{\succ}, \bar{P})$. Since $\phi(\bar{I}, \bar{\succ}, (\hat{P}_{\bar{i}}, \bar{P}_{-\bar{i}})) \neq \phi(\bar{I}, \bar{\succ}, \bar{P})$ and $\phi_{\bar{i}}(\bar{I}, \bar{\succ}, (\hat{P}_{\bar{i}}, \bar{P}_{-\bar{i}})) = \phi_{\bar{i}}(\bar{I}, \bar{\succ}, \bar{P}) = \emptyset$, ϕ cannot be non-bossy.

■

Proof. [Proof of Theorem 3] Before moving to proof, we introduce some notation and definition. For a student i with preferences P_i, we write $Rank(i, s|P_i)$ for the ranking of school s at student i's preferences P_i. That is, $Rank(i, s|P_i) = \ell$ means that school s is the top ℓ^{th} choice of student i having P_i. Let us next define $r(i, s|P_i) = \lceil Rank(i, s|P_i)/e \rceil$. In the course of the proof, $r(i, s|P_i)$ will give us the round in which student i applies to school s in the course of Φ^e given that he is rejected from all of his better schools.

From Chen and Kesten (2015), we know that for any e, Φ^e is non-strategy-proof. Hence, let us assume that some student has a profitable deviation at problem P under Φ^e. That is, there exists a student i_1 with false preferences P'_{i_1} such that $\Phi^e(P'_{i_1}, P_{-i_1})P_{i_1}\Phi^e(P)$. For ease of notation, let $P' = (P'_{i_1}, P_{-i_1})$ and $a_1 = \Phi^e_{i_1}(P)$ and $a_2 = \Phi^e_{i_1}(P')$.

As $a_2P_{i_1}a_1$, school a_2 has no excess capacity at $\Phi^e(P)$. By the definition of Φ^e, for any $j \in \Phi^e_{a_2}(P)$, either $r(j, a_2|P_j) < r(i_1, a_2|P_{i_1})$ or $r(j, a_2|P_j) = r(i_1, a_2|P_{i_1})$ and $j \succ_{a_2} i_1$. In words, any student who is matched with school a_2 at $\Phi^e(P)$ either applies to school a_2 in a round before student i_1 does or they apply in the same round while he has higher priority than student i_1 at school a_2.

Student i_1 being assigned to school a_2 at $\Phi^e(P')$ implies that some student $i_2 \in \Phi^e_{a_2}(P)$ is placed at some other school, say a_3. If $a_2 P_{i_2} a_3$, then we are done as it means that student i_3 is harmed by student i_1's misreporting. Hence, let us assume that $a_3 P_{i_2} a_2$. That is, student i_2 is better off at $\Phi^e(P')$. By the same reasoning as above, it means that there exists some student $i_3 \in \Phi^e_{a_3}(P)$ placed at some other school, say a_4, at $\Phi^e(P')$. If $a_3 P_{i_3} a_4$, then we are done as student i_3 is getting worse off at P'. On the other hand, if the converse is the case, then it gives us another student-school pair as the same as before. By continuing in the same manner, if we eventually come up with a student being worse off at P', then we are done. Therefore, let us assume that any student we consider throughout this process becomes better off. We depict it with the following diagram:

$$i_1 \to a_2 \to i_2 \to a_3 \to i_3 \to a_4$$

Above, students are pointing to their more preferred assignments at $\Phi^e(P')$; and schools are pointing the students who are matched with them at P, yet not at P'. As everything is finite, continuing in the same manner as above would give us a cycle, implying that student i_1 belongs to a cycle of the following form:

$$i_1 \to a_2 \to i_2 .....i_{n-1} \to a_n \to i_1.$$

That is, there exist distinct students $i_1, .., i_{n-1}$ and schools $a_n = a_1, a_2, ...a_{n-1}$ such that

(i) $a_{k+1} P_{i_k} a_k$ for any $k \in \{1, .., n-1\}$ and

(ii) $\Phi^e_{i_k}(P) = a_k$ for any $k \in \{1, .., n-1\}$.

Claim 1 *For any $k \in \{1, .., n-1\}$, $r(i_k, a_k|P_{i_k}) = r(i_k, a_{k+1}|P_{i_k}) = r(i_{k+1}, a_{k+1}|P_{i_{k+1}})$. That is, all students in the above cycles apply to their respective schools within the same round in the course of $\Phi^e(P)$.*

Proof. [Proof of Claim 1]By construction, we have $a_{k+1}P_{i_k}a_k$ and $\Phi^e_{i_k}(P) = a_k$ for any $k \in \{1, ..., n-1\}$. This along with the definition of Φ^e implies that $r(i_k, a_{k+1}|P_{i_k}) \geq r(i_{k+1}, a_{k+1}|P_{i_{k+1}})$ (where $i_n = i_1$). Let us suppose that it holds strictly for some k'. That is, $r(i_{k'}, a_{k'+1}|P_{i_{k'}}) > r(i_{k'+1}, a_{k'+1}|P_{i_{k'+1}})$. If we apply the above inequality for every pair in the cycle, we end up with $r(i_{k'-1}, a_{k'-1}|P_{i_{k'-1}}) < r(i_{k'}, a_{k'+1}|P_{i_{k'+1}})$. This is turn implies that $r(i_{k'-1}, a_{k'}|P_{i_{k'-1}}) < r(i_{k'}, a_{k'}|P_{i_{k'}})$, contradicting the definition of Φ^e. If we apply the fact $r(i_k, a_{k+1}|P_{i_k}) = r(i_{k+1}, a_{k+1}|P_{i_{k+1}})$ to every student-school pair appearing in the cycle, we prove the claim. ∎

We therefore show that every student i_k in the above cycle applies to schools a_k and a_{k+1} in the same round in the course of $\Phi^e(P)$. By definition, within each round of Φ^e, the usual agent proposing DA is used. The above improvement cycle, hence, implies that there exists some student (out of the above cycle), say j, causing rejection cycle at P in the round in which the above students apply to their respective schools. On the other hand, as this rejection cycle does not occur at P', it means that $r(j, a_2|P_j) = r(i_1, a_2|P_{i_1})$ and $j \succ_{a_2} i_1$ (student i_1 prevents this cycle from occurring by misreporting) and $r(i_1, a_2|P'_i) < r(j, a_2|P_j)$.

At false preference profile P', student i_1 applies to school a_2 in round $r(i_1, a_2|P'_{i_1})$ and is placed at a_2. On the other hand, student i_1 and j apply to it in a later round at P, given by $r(i_1, a_2|P_{i_1})$. This implies that school a_2 does not exhaust its capacity by the end of round $r(i_1, a_2|P_{i_1}) - 1$ in the course of $\Phi^e(P)$. This, however, gives us another manipulative misreporting by student j. Namely, student j can misreport his preferences by announcing: $P''_j:\ a_2, \emptyset$ (as $r(j, a_2|P_j) > 1$, we infer that $P''_j \neq P_j$) and can be matched with school a_2 at $P'' = (P''_j, P_{-j})$ (note that all the students belonging the above cycle apply to their respective schools in some round $t \geq 2$).

Above shows that Φ^e is manipulable by student j at problem P and $\Phi^e_j(P'') = a_2 P_j \Phi^e(P)$. We now repeat the very first analysis above for student j. That is, student j being matched with school a_2 at P'' implies that some other student, say $k \in \Phi^e_{a_2}(P)$, is assigned to a different school. If that student is worse off, then we are done. If not, then we continue in the same manner as before. If we end up with a student who is worse off at P'', then we are done. Otherwise, we can create a cycle as above including student j and conclude that some student causes a rejection cycle causing student j to be rejected by school a_2 at P. This, however, is impossible as student j himself initiates a rejection cycle at P by applying to school a_2, which brings him anything while causing the students in the cycle above to be worse off. ■

Proof. [Proof of Proposition 4] We prove by comparing two extreme members of the class: the BM and the aDA. Consider a problem instance consisting of $I = \{i_1, i_2, i_3\}$ and $S = \{a, b, c\}$, each with unit capacity. Let the preference and priority order profiles be as follows:

$$P_{i_1} : a, c, b, \emptyset;\ P_{i_2} :\ a, b, c, \emptyset;\ P_{i_3} :\ c, a, b, \emptyset.$$

$$\succ_a:\ i_3, i_1, i_2;\ \succ_b:\ i_2, i_3, i_1;\ \succ_c:\ i_1, i_2, i_3.$$

Under the BM, $BM_{i_1}(P) = a$, $BM_{i_2}(P) = b$, $BM_{i_3}(P) = c$. It is easy to see that at P, the BM is neither profitably manipulable (that is, no student receives better school by misreporting) nor bossy. Let us consider the aDA (that is, $e = \infty$). $DA(P) = BM(P)$ at P, however, student i_2 can manipulate DA mechanism under problem P. Consider $P'_{i_2} : c, b, a, \emptyset$. Then, $DA_{i_1}(P'_{i_2}, P_{-i_2}) = c$, $DA_{i_2}(P'_{i_2}, P_{-i_2}) = b$, and $DA_{i_3}(P'_{i_2}, P_{-i_2}) = a$.

For the converse, consider the following profiles:

$$\tilde{P}_{i_1} : a, c, b, \emptyset;\ \tilde{P}_{i_2} :\ a, c, b, \emptyset;\ \tilde{P}_{i_3} :\ a, \emptyset.$$

$$\succ_a:\ i_3, i_1, i_2;\ \succ_b:\ i_2, i_3, i_1;\ \succ_c:\ i_1, i_2, i_3.$$

In problem $\tilde{P}$, $BM_{i_1}(\tilde{P}) = c$, $BM_{i_2}(\tilde{P}) = b$, $BM_{i_3}(\tilde{P}) = a$. Student i_2 can strategize by reporting school c as his top choice. In this case, he would receive school c, while student i_1 would be matched with school b and become worse off. However, in problem $\tilde{P}$, it is easy to verify that the aDA is non-manipulable.

■

Proof. [Proof of Proposition 7] Let $P_i' \in B_i(P_{-i})$ and $P^* = (P_i', P_{-i})$. Let $O_i(P)$ be the set of schools offering to student i at P (in the course of sDA) and $sDA_i(P) = s$. As student i does not lose at P^*, that is, $sDA_i(P^*)P_i sDA_i(P)$, it implies that at P^*, student i is not matched with any school in $O_i(P) \setminus \{c\}$. This along with everything else (except the preferences of student i) being the same at both P and P^* implies that any school $c' \in O_i(P)$ continues to propose to student i at P^* as well and he rejects any school $c' \in O_i(P) \setminus \{s\}$.

Above shows that we can without loss of generality assume that for any school $s' \in O_i(P) \setminus \{s\}$, $\emptyset P_i' s'$. We now want show that $sDA_j(P')P_j sDA_j(P)$ where $P' = (P_j')_{j \in I}$ and for any $j \in I$, $P_j' \in B_j(P_{-j})$ (recall that, by our supposition, $sDA_i(P_i', P_{-i})P_i sDA_i(P)$ for any $i \in I$).

By our above arguments, we have $\emptyset P_k' s'$ for any $s' \in O_k(P) \setminus \{sDA_k(P)\}$ for $k \in I$. As schools propose to students in decreasing order of their priorities in the course of sDA and schools that are rejected at P are to be rejected at P' as well, every student k receives an offer from $sDA_k(P)$ (the steps of sDA leading school $sDA_k(P)$ to offer to student k will be the same at both P and P'). Hence, the result follows.

■

References

Abdulkadiroglu, A. and Y.-K. Che (2010): "The Role of Priorities in Assigning Indivisible Objects: A Characterization of Top Trading Cycles," *mimeo.*

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

Abdulkadiroğlu, A., P. A. Pathak, and A. E. Roth (2005): "The New York City High School Match," *American Economic Review Papers and Proceedings*, 95, 364–367.

Afacan, M. O. (2013a): "Alternative Characterizations of Boston Mechanism," *Mathematical Social Sciences*, 66(2), 176–179.

——— (2013b): "Application Fee Manipulations in Matching Markets," *Journal of Mathematical Economics*, 49(6), 446–453.

Alcalde, J. and S. Barbera (1994): "Top dominance and the possibility of strategy-proof stable solutions to matching problems," *Economic Theory*, 4(3), 471–435.

Balinski, M. and T. Sönmez (1999): "A Tale of Two Mechanisms: Student Placement," *Journal of Economic Theory*, 84, 73–94.

Chen, Y. and O. Kesten (2015): "From Boston to Chinese Parallel to Deferred Acceptance: Theory and Experiments on a Family of School Choice Mechanisms," *forthcoming, Journal of Politicial Economy.*

Chen, Y. and T. Sönmez (2006): "School choice: an experimental study," *Journal of Economic Theory*, 127, 202–231.

Dubins, L. E. and D. A. Freedman (1981): "Machiavelli and the Gale-Shapley Algorithm," *American Mathematical Monthly*, 88, 485–494.

Dur, U. M. (2015): "A Characterization of the Top Trading Cycles mechanism in the School Choice Problem," *mimeo.*

Dur, U. M., R. G. Hammond, and T. Morrill (2015): "Identifying the Harm of Manipulable School-Choice Mechanisms," *mimeo.*

Ehlers, L. and B. Klaus (2014): "Strategy-Proofness makes the Difference: Deferred-Acceptance with Responsive Priorities," *Mathematics of Operations Research*, 39, 949–966.

Ergin, H. I. (2002): "Efficient Resource Allocation on the Basis of Priorities," *Econometrica*, 88, 485–494.

Gale, D. and L. S. Shapley (1962): "College Admissions and the Stability of Marriage," *American Mathematical Monthly*, 69, 9–15.

Kesten, O. (2012): "On two kinds of manipulation for school choice problems," *Economic Theory*, 51(3), 677–693.

Kojima, F. (2006): "Mixed strategies in games of capacity manipulation in hospitalintern markets," *Social Choice and Welfare*, 27, 25–28.

——— (2011): "Robust Stability in Matching Markets," *Theoretical Economics*, 6, 257–267.

Kojima, F. and M. Manea (2010): "Axioms for Deferred Acceptance," *Econometrica*, 78(2), 633–653.

Kojima, F. and P. A. Pathak (2009): "Incentives and Stability in Large Two-Sided Matching Markets," *American Economic Review*, 99(3), 608–627.

Kojima, F. and U. Ünver (2014): "The "Boston" School-Choice Mechanism," *Economic Theory*, 55(3), 515–544.

Konishi, H. and U. Ünver (2006): "Games of Capacity Manipulation in Hospital-Intern Markets," *Social Choice and Welfare*, 27, 3–24.

Morrill, T. (2013a): "An Alternative Characterization of the Deferred Acceptance Algorithm," *International Journal of Game Theory*, 42, 19–28.

——— (2013b): "An Alternative Characterization of Top Trading Cycles," *Economic Theory*, 54, 181–197.

Pápai, S. (2000): "Strategyproof Assignment by Hierarchical Exchange," *Econometrica*, 68, 1403–1433.

Pathak, P. and T. Sönmez (2008): "Leveling the Playing Field: Sincere and Sophisticated Players in the Boston Mechanism," *American Economic Review*, 98(4), 1636–1652.

Pathak, P. A. and T. Sönmez (2013): "School Admissions Reform in Chicago and England: Comparing Mechanisms by their Vulnerability to Manipulation," *American Economic Review*, 103(1), 80–106.

Roth, A. E. (1982): "The Economics of Matching: Stability and Incentives," *Mathematics of Operations Research*, 7, 617–628.

Roth, A. E., P. A. Pathak, A. Abdulkadiroğlu, and T. Sönmez (2005): "The Boston Public School Match," *American Economic Review*, 95(2), 368–371.

——— (2006): "Changing the Boston School Choice Mechanism: Strategy-proofness as Equal Access," *mimeo.*

Roth, A. E. and E. Peranson (1999): "The Redesign of the Matching Market for American Physicians: Some Engineering Aspects of Economic Design," *American Economic Review*, 89(4), 748–780.

Satterthwaite, M. A. and H. Sonnenschein (1981): "Strategy-Proof Allocation Mechanisms at Differentiable Points," *The Review of Economic Studies*, 48(4), 587–597.

Sönmez, T. (1997): "Manipulation via Capacities in Two-Sided Matching Markets," *Journal of Economic Theory*, 77(1), 197–204.

——— (1999): "Can Pre-arranged Matches Be Avoided in Two-Sided Matching Markets?" *Journal of Economic Theory*, 86, 148–156.