

# On Two Kinds of Manipulation for School Choice Problems

Onur Kesten\*

This version: January 2008

## Abstract

Many school districts in the US employ centralized clearing houses for the placement of students to public schools. A popular mechanism from two-sided matching theory, the Gale-Shapley student-optimal stable mechanism, has recently replaced two deficient mechanisms that were in use in NYC and Boston. An important potential threat against any school choice mechanism is the tendency of schools to circumvent the procedure via two kinds of strategic manipulations introduced by Sönmez (1997, 1999): *manipulation via underreporting capacities* and *manipulation via pre-arranged matches*. This paper studies the extent of the vulnerability of a school choice mechanism, and in particular, that of the student-optimal stable mechanism, to the two manipulations, and identifies conditions under which they can be avoided. *Journal of Economic Literature* Classification Numbers: C71, C78, C79, D61, D71, D78

*Key words:* School choice; Student-optimal stable mechanism; Manipulation; Acyclicity

## 1 Introduction

Starting in 1987 with Minnesota, several US states began using centralized *school choice programs* that offer students flexibility over choosing schools they wish to attend. In a school choice program, each student submits a list of *preferences* over schools to a central placement authority, such as the school district, which then decides which school each student will attend by also taking into consideration the *priorities* of students for each school. Several criteria may be used for determining a priority order for a school. For example, under the current Boston public school system, a student is given priority for a particular school before others, if he is in the school's walk zone and has a sibling attending that school, and the remaining priorities are determined based on which one of the two criteria a student satisfies, and so on.

A standard *school choice problem* or simply, a problem, contains two pieces of information: a *preference profile* of students and a *priority profile* of schools. At a *matching*, each student is assigned to at most one school, and the number of students assigned to a particular school does not exceed the capacity of that school. A *school choice mechanism* or simply, a mechanism, is a systematic way of selecting a matching for a given problem.

A very closely related problem is the well-known *two-sided matching problem* of Gale and Shapley (1962).<sup>1</sup> The main difference is that in two-sided matching, the 'priorities of students for schools' are replaced by the 'preferences of schools over (groups of) students.' In the school choice context however, priorities are imposed by local/state laws, and schools have no say on the way priority orders are determined. For a school choice problem a matching is *stable* if (1) each student prefers his own assignment to not having any (real) assignment, and (2) there is no unmatched student-school pair  $(i, x)$  such that

---

\*Tepper School of Business, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213, USA; e-mail: okesten@andrew.cmu.edu. Any remaining errors are my own responsibility.

<sup>1</sup>See Roth and Sotomayor (1990) for a thorough analysis of this problem.

student  $i$  prefers school  $x$  to his assignment, and student  $i$  has higher priority than at least one student who is assigned a seat at school  $x$ .<sup>2</sup>

Many school districts determine each school’s priority order either completely independently of that particular school’s own evaluation of the students (such as in Boston), or offer only a limited number of schools such an opportunity (such as in New York City). Therefore, priorities cannot be seen and interpreted to be representative of the preferences of schools. This observation led the standard approaches to the school choice problem to consider schools as merely “objects” to be consumed, and ignore the strategic (as well as the welfare aspects) of the problem for the school side. A recent study on the New York City public high school system by Abdulkadiroğlu, Pathak, and Roth (2005) shows that this may not always be the case. In particular, in New York City substantial number of schools were seen to conceal their capacities to get matched with the students they preferred under the school choice mechanism that was in use between 1990’s and 2002.

Shortly before the school choice problems came to widespread attention, Sönmez (1997, 1999) introduced and studied two interesting kinds of strategic behaviour for two-sided matching problems. The first one is the behaviour Abdulkadiroğlu, Pathak, and Roth (2005) observed in New York City: “manipulation via capacities” (Sönmez, 1997): A school may circumvent the centralized procedure by underreporting its capacity. The second one is the possibility of “manipulation via pre-arranged matches” (Sönmez, 1999): A student-school pair may commit themselves to a mutual agreement prior to the centralized procedure, according to which the student does not participate in the procedure, and in return is offered a seat in the school, which results in a strict gain for at least one of the two parties while hurting neither. It turns out that no stable mechanism survives any of the two manipulations (Sönmez 1997, 1999).

The pioneering work of Abdulkadiroğlu and Sönmez (2003) on school choice problems examines some of the real-life student placement mechanisms, and offers two alternative competing mechanisms as attractive replacements.<sup>3</sup> One of these mechanisms, the Gale and Shapley *student-optimal stable mechanism* inherits a number of appealing properties from two-sided matching theory such as stability and strategy-proofness<sup>4</sup>. Among school choice mechanisms, the student-optimal stable mechanism has taken an early lead over its competitors by recently replacing two controversial school choice mechanisms in New York City (which has the largest public school system in the country with over a milion students) and Boston (which has over 60,000 students enrolled in the public school system; see Abdulkadiroğlu, Pathak, Roth, and Sönmez, 2005 and Ergin and Sönmez, 2006).

In this paper we study the two manipulation notions for school choice problems. In our formulation of the problem, we extend a standard school choice problem to also contain the preferences of each school which are now assumed to be its *private* information and possibly different than the corresponding priorities. Hence, the central planner is unable to observe schools’ preferences, and from his perspective they are quite unrestricted.

We show that between the two manipulation notions, capacity manipulations are easier to avoid, while it is essentially impossible to achieve immunity againts manipulations via pre-arranged matches (Proposition 1 and Theorem 2). Furthermore, quite interestingly, the popular student-optimal stable mechanism is prone to capacity manipulations unlike some other less popular mechanisms (Corollary 1). Motivated by its increasing popularity, we further investigate the causes of the manipulability of the student-optimal stable mechanism, and identify a sufficient and necessary condition on the priority

---

<sup>2</sup>The counterpart of stability in our context is the elimination of “justified envy.” An allocation suffers from the *justified envy* of a student if there is some school  $x$  which he prefers to his placement, and there is some other student who has lower priority for school  $x$  than him but was placed to school  $x$ . Here we will not concentrate on this aspect of the problem. See Abdulkadiroğlu and Sönmez (2003) and Kesten (2004) for more on this.

<sup>3</sup>To the best of our knowledge, Balinski and Sönmez (1999) is the first to formulate and study the school choice problem in an abstract setting.

<sup>4</sup>It requires that no agent ever benefits by misrepresenting his preferences. See Pápai (2000) for a characterization of a large class of strategy-proof mechanisms and Ehlers (2002) for a study of group strategy-proof mechanisms.

profile and the vector of minimum capacities (that can be reported), that ensures its immunity to capacity manipulations (Theorem 1).

The paper is organized as follows: In the next section we introduce the model, and describe the student-optimal stable mechanism. Section 3 discusses manipulations via capacities, and Section 4 manipulations via pre-arranged matches. We briefly conclude in Section 5. All the proofs are given in the Appendix.

## 2 The Model

Let  $N \equiv \{1, 2, \dots, n\}$ ,  $n \geq 3$ , denote the finite set of students. Let  $X$  denote the finite set of schools. Each school  $x \in X$  has a capacity  $q_x$  which is the maximum number of students it can be assigned. Let  $q \equiv (q_x)_{x \in X}$  be the capacity vector.

Each student  $i \in N$  is equipped with a complete, transitive, and strict preference relation  $R_i$  over  $X \cup \{\emptyset\}$  where  $\emptyset$  represents not being assigned to any school in  $X$ . Let  $\mathcal{R}_i$  denote the class of all such preferences for student  $i \in N$ . For each  $i \in N$ , let  $P_i$  denote the strict relation associated with  $R_i$ .

For each school there is a complete and transitive *priority order* over all students, which are determined according to state/local laws and certain criteria of school districts (such as proximity of residence, possible specific needs of a student etc.). In the US public school system, for each school, the priority between any two students who are identical in every relevant aspect is determined by a lottery. Hence, for each school the associated priority order is *strict*. Let  $\succ_x$  denote the priority order for school  $x$ . For example,  $i \succ_x j$  means that student  $i$  has higher priority for school  $x$  than student  $j$ . Let the upper contour set of  $\succ_x$  at student  $i$  be denoted by  $U_x(i) \equiv \{j \in N \mid j \succ_x i\}$ . The collection of priority orders  $\succ \equiv (\succ_x)_{x \in X}$  is called a *priority profile*. The pair  $(\succ, q)$  is called a *priority structure*. We assume that a priority profile is given as a primitive of model.

A *matching* is a function  $\mu : N \cup X \rightarrow 2^{N \cup X}$  such that:

1. For all  $i \in N$ ,  $|\mu(i)| \leq 1$  and  $\mu(i) \subseteq X$ .
2. For all  $x \in X$ ,  $|\mu(x)| \leq q_x$  and  $\mu(x) \subseteq N$ .
3. For all  $(i, x) \in N \times X$ ,  $\mu(i) = \{x\}$  if and only if  $i \in \mu(x)$ .

Given a priority structure  $(\succ, q)$ , a matching  $\mu$  is *stable* if there is no unmatched student-school pair  $(i, x)$  such that  $x P_i \mu(i)$ , and no  $j \in \mu(x)$  such that  $i \succ_x j$ .

Each school  $x \in X$  is equipped with a *strict* and *responsive* (Roth, 1985) preference relation  $R_x$  over  $2^N$ . Each school's preferences are its private information, and hence are not observable by the central planner. Preference relation  $R_x$  is responsive if for all  $M \subset N$ ,

1. For all  $i \in N \setminus M$ ,  $M \cup \{i\} P_x M$  if and only if  $\{i\} P_x \emptyset$ .
2. For all  $i, j \in N \setminus M$ ,  $M \cup \{i\} P_x M \cup \{j\}$  if and only if  $\{i\} P_x \{j\}$ .

Let  $\mathcal{R}_x$  denote the class of all such preferences for school  $x \in X$ .

A (*school choice*) *problem* is a pair  $((R_s)_{s \in N \cup X}, q)$  or simply,  $(R, q)$  that specifies the preferences of each student, the unobservable preferences of each school, and the capacity of each school. Let  $\mathcal{E} \equiv \prod_{s \in N \cup X} \mathcal{R}_s \times \mathbb{N}_+^{|X|}$  be the class of all problems.

Given a priority structure  $(\succ, q)$ , the matching outcome is determined according to the “mechanism” chosen based *only* on the submitted preferences of students. A (*school choice*) *mechanism*  $\varphi$  associates to each (school choice) problem  $(R, q)$  a matching  $\varphi(R, q)$ . A mechanism is stable if its outcome is stable at each problem. For ease of notation, we will sometimes use  $\varphi_i(R, q)$  to refer to  $\varphi(R, q)(i)$ .

### 2.1 Student-optimal Stable Mechanism

Gale and Shapley (1962) proposed the student-optimal stable mechanism to find the stable allocation that is most favorable to each student for any given two-sided matching problem. Since then this mech-

anism (as well as its dual for the school side) has been the central mechanism in two-sided matching. Its outcome can be calculated via the following *deferred acceptance algorithm* for a given problem:

*Step 1: Each student applies to his favorite acceptable school. For each school  $x$ , those  $q_x$  applicants (all applicants if fewer than  $q_x$ ) who have the highest priority for school  $x$  are tentatively placed to school  $x$ . The remaining applicants are rejected.*

*Step  $k$ ,  $k \geq 2$ : Each student who was rejected from a school at step  $k - 1$  applies to his next favorite acceptable school. For each school  $x$ , those  $q_x$  students who have the highest priority for school  $x$  among the new applicants and those who were tentatively placed to it at an earlier step are tentatively placed to school  $x$ . The remaining applicants are rejected.*

The algorithm terminates when each student is either tentatively placed to a school, or has been rejected by every school acceptable to him. We denote the student-optimal stable mechanism (SOSM) associated with the priority profile  $\succ$  by  $\varphi^\succ$ .

SOSM has proven quite successful not only in two-sided matching theory but also in a number of entry-level labor markets the biggest of which is the re-designed US National Resident Matching Program (Roth and Peranson, 1999). One main reason (aside from its stability property) behind its success is the fact that it makes it a dominant strategy for students to state true preferences (Roth (1982); Dubins and Freedman (1981)).

SOSM has so far been the leading mechanism for school choice applications as well. Since the first study on this problem by Abdulkadiroglu and Sönmez (2003), it has often been advocated for school choice problems. These efforts recently paid off, and it has been adapted by the New York City (in 2002) and Boston (in 2005) public school systems who have been suffering from congestion (NYC) and incentive (Boston) problems.

### 3 Manipulation via Capacities

Sönmez (1997) considers a two-sided matching application with medical interns on one side of the market and hospitals on the other. He introduces an interesting kind of manipulation that a hospital may engage in by withholding its capacity. We consider the same possibility for the school choice problem where preferences of each school are now its private information, and the capacity reported by each school is the only parameter it has an influence on among those that get included in the calculation of the outcome of the central clearing house. In fact, this is not a merely theoretical concern. Indeed, Abdulkadiroglu, Pathak, and Roth (2005) report that public schools in NYC were withholding capacity under the old public school system<sup>5</sup> that was in use from 1990s till 2003.

For each school  $x \in X$ , we now interpret  $q_x$  as the capacity reported by school  $x$ . In many school districts, each school is typically required to admit a certain number of students. Therefore it is not realistic to assume that a school's capacity can be arbitrarily small. We incorporate this possibility into our model via the following restriction: For each  $x \in X$ , let  $\underline{q}_x$  denote the minimum capacity imposed on school  $x$ . That is, school  $x$  can never report a capacity less than  $\underline{q}_x$ . Let  $\underline{q} \equiv (\underline{q}_x)_{x \in X}$  be the minimum capacity vector. Let  $\varphi$  be a mechanism:

**Non-manipulability via Capacities:** For all  $(R, q) \in \mathcal{E}$ , all  $x \in X$ , and all  $\underline{q}_x \leq q'_x < q_x$ , we have  $\varphi(R, q)(x) R_x \varphi(R, q'_x, q_{-x})(x)$ .

---

<sup>5</sup>It was a centralized system which operated through three rounds of application processing, and subsequently suffered from congestion problems. See Abdulkadiroglu, Pathak, and Roth (2005) for more.

Sönmez (1997) shows that the hospital-optimal stable mechanism that was used by the National Resident Matching Program is not immune to manipulation via capacities, and moreover, there is no stable mechanism that is non-manipulable via capacities in the two-sided matching context.

For school choice problems it turns out that non-manipulability via capacities requires that if a student is matched to a school when that school underreports its capacity, then the student should continue to be matched to that school when the school truthfully reports its capacity. All proofs are deferred to the Appendix.

**Proposition 1.** *A school choice mechanism  $\varphi$  is non-manipulable via capacities for a problem  $(R, q)$  if and only if  $\varphi(R, q'_x, q_{-x})(x) \subset \varphi(R, q)(x)$  for all  $q_x \leq q'_x < q_x$ .*

Surprisingly, the most popular mechanism in two-sided matching theory, which is also gaining popularity among school districts, is not immune to capacity manipulations. On the contrary, its closest competitor, the top trading cycles mechanism introduced in Abdulkadiroglu and Sönmez (2003) as well as the mechanism it recently replaced in Boston survive this manipulation test.

**Corollary 1.** *The student-optimal stable mechanism is manipulable via capacities, whereas the top trading cycles mechanism and the Boston mechanism are not.*<sup>6</sup>

The student-optimal stable mechanism is currently in use in a number of entry-level labor markets (see Roth and Rothblum (1999) for a list of these markets). Due to its increasing popularity also among the public school systems in the US, we next study this mechanism more closely and analyze the causes and the extent of its manipulability. We first consider an example that illustrates how it can be manipulated:

**Example 1:** Let  $N = \{1, 2, 3\}$ ,  $X = \{a, b\}$  where (true) capacities are  $q_a = 2$  and  $q_b = 1$ . Also let

$\succ_a$	$\succ_b$
1	3
3	2
2	1

$$\begin{aligned} & \{1, 2, 3\} P_a \{1\} P_a \{2, 3\} P_a \{3\} P_a \{2\}, \\ & b P_1 a P_1 \emptyset, a P_2 b P_2 \emptyset, a P_3 \emptyset, \end{aligned}$$

where a school not specified in a student's preference relation is ranked lower than those specified. It is straightforward to calculate that  $\varphi^\succ(R, q)(a) = \{2, 3\}$ . Now let  $q'_a = 1$ . Observe that when school  $a$  underreports its capacity as  $q'_a = 1$ , student 2 is now rejected from school  $a$ . Next he applies to his second favorite school which is  $b$ , and causes student 1 to be rejected from school  $b$ . This, in turn, causes school  $a$  to admit student 1 and reject student 3. Thus,  $\varphi^\succ(R, q'_a, q_{-a})(a) = \{1\}$ , and school  $a$  gains by underreporting its capacity.  $\diamond$

---

<sup>6</sup>The Boston mechanism is a member of a large class of *priority matching mechanisms* studied by Roth (1991). These mechanisms have been in use in the entry-level labor markets in the UK. Any such priority matching mechanism that satisfies the monotonicity requirement of Proposition 1 is also immune to capacity manipulations.

Ergin (2002) identifies conditions under which the student-optimal stable mechanism can achieve a number of appealing properties it lacks otherwise. He introduces a notion of “acyclicity” for priority structures, and shows that it is sufficient as well as necessary to recover properties such as Pareto efficiency and group strategy-proofness. We argue that acyclicity (adopted appropriately to our setting) also plays a key role for the student-optimal stable mechanism to avoid capacity manipulations.

A given priority structure  $(\succ, \underline{q})$  contains a cycle if the following two conditions are satisfied:

*Loop condition:* There are distinct  $i, j, k \in N$  and  $x, y \in X$  such that  $i \succ_x j \succ_x k$  and  $k \succ_y i$ .

*Scarcity condition:* There exist (possibly empty) disjoint sets of students  $N_x, N_y \subset N \setminus \{i, j, k\}$  such that  $N_x \subset U_x(j)$ ,  $N_y \subset U_y(i)$ ,  $|N_x| = \underline{q}_x - 1$  and  $|N_y| = \underline{q}_y - 1$ .

The priority structure  $(\succ, \underline{q})$  is *acyclical* if it has no cycles.

Acyclicity restrictions (the loop condition and the scarcity condition) apply jointly on the priority profile and the minimum capacity vector. For the student-optimal stable mechanism, loosely speaking, acyclicity prevents those situations in which a school, by underreporting capacity, initiates a new rejection chain that eventually leads to the application of a completely new student to this school (recall Example 1). For example, in the presence of a cycle in the priority structure, some school, say  $x$ , can simply refuse to be assigned some low priority student, say  $k$ , by concealing an open slot. Such refusal may however give rise to a subsequent rejection of some high priority student, say  $i$ , from another school, say  $y$ , causing him to apply to school  $x$ .

Two extreme examples of acyclic structures are the following: If the minimum capacity a school can report is  $n$ , then school seats are always in abundance (hence the scarcity condition is not satisfied), and there are no cycles. On the other hand, if the priority orders are identical for each school, then the loop condition is not satisfied, and the structure is again acyclic regardless of minimum capacities. In general, as minimum capacities get larger, formation of cycles becomes harder. We refer the interested reader to Ergin (2002) for a nice visual characterization of an acyclic structure.

Next is our main result. It turns out that the student-optimal stable mechanism is immune to manipulation via capacities so long as the priority structure  $(\succ, \underline{q})$  is acyclical.

**Theorem 1.** *The student-optimal stable mechanism  $\varphi^\succ$  associated with a priority profile  $\succ$  is immune to manipulation via capacities if and only if the priority structure  $(\succ, \underline{q})$  is acyclical.*

An acyclic priority structure completely eliminates capacity manipulations. The presence of a cycle in a priority structure however introduces the possibility of manipulation. Consequently, as also suggested by the proof of Theorem 1, the more cycles a priority structure has, the more room there is for manipulation opportunities.

Example 1 as well as the proof of Theorem 1 hint at a second observation about capacity manipulations: Any school successfully gaining by underreporting capacity *should* indeed have preferences that have certain degree of ‘correlation’ with the school’s priority order. The connection between capacity manipulations and the acyclicity of a priority structure shown by Theorem 1 also entails an important practical advice suggested by the next result.

**Corollary 2.** *Fix a priority profile  $\succ$ . Consider a school district that uses the student-optimal stable mechanism  $\varphi^\succ$  and needs to choose between two possible minimum capacity vectors  $\underline{q}$  and  $\underline{q}'$  such that  $\underline{q} \succeq \underline{q}'$ . If a school can manipulate by underreporting its capacity when  $\underline{q}$  is imposed, then the same school can also manipulate when  $\underline{q}'$  is imposed. However, the converse is not necessarily true.*

Corollary 2 implies that the Boston and NYC school districts can simply reduce the chances of capacity manipulations by setting higher minimum capacity restrictions on schools.

Konishi and Unver (2006) consider capacity manipulation games in the hospital-intern market application of the model. They analyze the pure strategy equilibria of the game-form under the school-optimal (hospital-optimal in their context) and the student-optimal (intern-optimal in their context) stable mechanisms. They show that neither game may have a pure strategy equilibrium.

Now an admissible capacity of each school  $x$  is a nonnegative integer no more than its minimum capacity  $q_x$ . Thus,  $x$  can report  $q'_x \in Q_x \equiv \{q_x, q_x + 1, \dots, q_x\}$ . Define the set of admissible capacities as  $Q = \prod_{x \in X} Q_x$ . A *capacity reporting game under a mechanism*  $\varphi$  is described by a strategic form game  $(X, (Q_x, R_x)_{x \in X})$ . A pure strategy Nash equilibrium of a game  $(X, (Q_x, R_x)_{x \in X})$  is a strategy profile at which no school has a profitable deviation.

Theorem 1 has direct implications for capacity reporting games. We know that under the acyclicity restriction no school benefits by underreporting its capacity. Then the following is immediate.<sup>7</sup>

**Corollary 3.** *The capacity reporting game under the student-optimal stable mechanism  $\varphi^{\succ}$  associated with a priority profile  $\succ$  has a pure strategy Nash equilibrium if the priority structure  $(\succ, \underline{q})$  is acyclical.*

## 4 Manipulation via Pre-arranged Matches

Sönmez (1999) studies the vulnerability of two-sided matching markets to an alternative form of manipulation: *manipulation via pre-arranged matches*: A student-school pair agree on an arrangement before the formal procedure according to which the student is enrolled at the school, and the student does not participate in the procedure. This arrangement is successful if (at least) one of the two parties gains and neither one suffers as a result. Such a behavior was observed in the US hospital-intern market and the Canadian lawyer market where stable mechanisms were in use. Sönmez (1999) shows for two-sided matching problems that there is no mechanism that is both stable and non-manipulable via pre-arranged matches.

We now study this second kind of manipulation for school choice problems. In the present context, since the population of students might change, a (*school choice*) *problem* is defined by a triplet  $(N, R, q)$ . The rest of the definitions and notations identically apply. Let  $\varphi$  be a mechanism:

**Non-manipulability via pre-arranged matches:** *Given a problem  $(N, R, q)$ , there is no school-student pair  $(x, i)$  such that  $x R_i \varphi_i(N, R, q)$ , and*

$$(\{i\} \cup \varphi_x(N \setminus \{i\}, R_{-i}^{N \setminus \{i\}}, q_{-x}, q_x - 1)) R_x \varphi_x(N, R, q).$$

*with at least one of the relations holding strictly.*

It turns out that manipulations via pre-arranged matches are essentially impossible to avoid. The next result makes this point precise.

**Theorem 2.** *Suppose that there is a school  $x \in X$  such that  $n > q_x$ . Then no mechanism is immune to manipulations via pre-arranged matches.*

Theorem 2 shows that mechanism design is almost helpless about the potential pre-arrangements of agents from the two-sides of the problem. This result simply suggests that essentially the only way to prevent this kind of behaviour would be via establishing explicit laws and regulations within school

<sup>7</sup>Also see Kojima (2006) and Kojima and Pathak (2007).

districts that discourage schools from enrolling students on their own, and avoiding systems that would allow decentralized behaviour prior to the centralized procedure.

**Corollary 4.** *All three of the student-optimal stable mechanism, the top trading cycles mechanism, and the Boston mechanism are manipulable via a pre-arranged match.*

## 5 Conclusion

Our results show that immunity to capacity manipulations is easier to achieve than to those via-pre arranged matches. We have shown that while the top trading cycles mechanism (TTCM) and the monotone priority matching mechanisms such as the Boston mechanism are readily immune to capacity maneuvers, certain restrictions are needed to make this possible for the student-optimal stable mechanism (SOSM). At a first glance these results may seem in odd conflict with those of Kesten (2006) where SOSM has been shown to have the edge over TTCM in terms of other desirable properties such as population/resource monotonicity<sup>8</sup> and consistency. One way to reconcile the two results is that the monotonicity property of SOSM (in particular, its resource monotonicity), when viewed from the perspective of schools, translates into a chance of profitable manipulation for the school side given the stability of SOSM. It is also worthwhile to note that ‘cyclicity of a priority structure’ has once again proven to be the main reason behind another vulnerable aspect of SOSM.

## 6 The Appendix

**Proof of Proposition 1:** The ‘if’ part simply follows from the responsiveness of schools’ preferences. To see the ‘only if’ part suppose there exist a mechanism  $\varphi$ , a problem  $(R, q)$ , a school  $x \in X$ , and  $q_x \leq q'_x < q_x$  such that  $\varphi(R, q'_x, q_{-x})(x) \setminus \varphi(R, q)(x) \neq \emptyset$ . But because schools’ preferences are unobservable by the mechanism designer, the outcome is independent of the preferences of school  $x$ . Then, one can easily construct preferences such that  $\varphi(R, q'_x, q_{-x})(x) P_x \varphi(R, q)(x)$ .

**Q.E.D.**

**Proof of Corollary 1:** Example 1 shows that SOSM does not satisfy the property given in Proposition 1. To see that the Boston mechanism and TTCM indeed do, simply observe that the steps and assignments of students for problems  $(R, q'_x, q_{-x})$  and  $(R, q)$  are identical under each mechanism, until right after the last seat of school  $x$  is assigned.

**Q.E.D.**

**Proof of Theorem 1:**

(i) *Acyclicity of  $\succ \implies$  Non-manipulability via capacities:*

Suppose that there are  $(R, q) \in \mathcal{E}$ ,  $x \in X$ , and  $q_x \leq q'_x < q_x$  such that  $\varphi^\succ(R, q'_x, q_{-x})(x) P_x \varphi^\succ(R, q)(x)$ . Let  $\mu \equiv \varphi^\succ(R, q)$  and  $\mu' \equiv \varphi^\succ(R, q'_x, q_{-x})$ . If  $q_x > |\mu(x)|$ , then  $\mu \neq \mu'$  implies  $q'_x < |\mu(x)|$ . Otherwise,  $q_x = |\mu(x)|$  and since  $q'_x < q_x$ , clearly  $q'_x < |\mu(x)|$ . Thus,  $|\mu'(x)| \leq q'_x < |\mu(x)|$ . This means  $|\mu(x) \setminus \mu'(x)| \geq 1$ . Moreover, since  $\mu'(x) P_x \mu(x)$ , we have  $\mu'(x) \setminus \mu(x) \neq \emptyset$ . Recall the way the outcome of  $\varphi^\succ$  is calculated: Each student applies to his favorite choice and if rejected, applies to his next favorite choice and so on. Note that the only difference between  $(R, q)$  and  $(R, q'_x, q_{-x})$  is that  $q'_x < q_x$ . This means that each student applies to the same schools at  $(R, q'_x, q_{-x})$  as he did at  $(R, q)$ , and to possibly more, since now  $q'_x < q_x$ . This is also equivalent to saying that all students are made weakly worse off.<sup>9</sup> Moreover, all students in the set  $(\mu'(x) \setminus \mu(x)) \cup (\mu(x) \setminus \mu'(x))$  are made strictly worse off at  $(R, q'_x, q_{-x})$ .

<sup>8</sup>The resource/population monotonicity property requires that all students are affected in the same direction (in welfare terms) whenever the set of schools/students shrinks or expands.

<sup>9</sup>This is the ‘resource monotonicity’ property of SOSM.

We prove this part by constructing a matching  $\mu''$  that Pareto dominates  $\mu'$  for the problem  $(R, q'_x, q_{-x})$  when we consider only students' welfare (i.e., we assume that schools are now irrelevant for welfare considerations just as in Ergin (2002)). Then by Ergin (2002) this means that priority structure  $(\succ, q'_x, q_{-x})$  has a cycle according to the definition in Ergin (2002). Since  $\underline{q}_x \leq q'_x$  and  $\underline{q}_y \leq q_y$  for any  $y \in X \setminus \{x\}$ , the priority structure  $(\succ, \underline{q})$  has a cycle according to our definition. A contradiction.

McVitie and Wilson (1970) showed that the outcome of the student-optimal stable mechanism is independent of the sequence according to which students make their applications. Therefore one can alternatively calculate  $\varphi^\succ(R, q'_x, q_{-x})$  as follows: Initially run the DA algorithm for the problem  $(R, q)$ . Then decrease the quota for  $x$  from  $q_x$  to  $q'_x$ . This will cause only those  $q_x - q'_x$  students with lowest priority to be rejected from  $x$ . Next each rejected student will apply to his next choice, and the algorithm will eventually terminate when there are no more rejections.

Let  $i_0 \in \mu'(x) \setminus \mu(x)$ . Also let  $i_0 \in \mu(y_0)$  for some  $y_0 \in X \setminus \{x\}$ .<sup>10</sup> Since student  $i_0$  is worse off at  $\mu'$ , there exists  $i_1 \in \mu'(y_0) \setminus \mu(y_0)$  such that  $i_1 \neq i_0$ , who did not apply to  $y_0$  at  $(R, q)$  (otherwise, he would be accepted by  $y_0$ ). Let  $i_1 \in \mu(y_1)$  for some  $y_1 \in X \setminus \{x, y_0\}$ . Since student  $i_1$  is worse off at  $\mu'$ , there exists  $i_2 \in \mu'(y_1) \setminus \mu(y_1)$  who did not apply to  $y_1$  at  $(R, q)$ . Let  $i_2 \in \mu(y_2)$  for some  $y_2 \in X \setminus \{x, y_0, y_1\}$ . Continuing in this fashion, there must exist  $i_k \in \mu'(y_{k-1}) \setminus \mu(y_{k-1})$  for some  $y_{k-1} \in X \setminus \{x, y_0, \dots, y_{k-2}\}$  with  $k \geq 1$  and  $x \equiv y_{-1}$ , who is the *first* student (according to the steps of the DA algorithm) in this sequence to apply to his assignment at  $(R, q'_x, q_{-x})$  to which he did not apply at  $(R, q)$ . By the above described calculation of  $\varphi^\succ(R, q'_x, q_{-x})$ , this student has to belong to  $\mu(x) \setminus \mu'(x)$  or otherwise,  $\mu' = \mu$ . Then let  $\mu''$  be such that if  $i \in \{i_0, i_1, \dots, i_k\}$ ,  $\mu''(i) = \mu(i)$ , and otherwise  $\mu''(i) = \mu'(i)$ . (i.e.,  $\mu''$  is obtained from  $\mu'$  by simply assigning each  $i_t$  with  $t \in \{0, 1, \dots, k\}$  to  $\mu'(i_{t+1})$  where  $i_{k+1} \equiv i_0$ .) Since  $k \geq 1$ ,  $\mu''$  clearly Pareto dominates  $\mu'$  at  $(R, q'_x, q_{-x})$  from the view of the student side.

(ii) *Non-manipulability via capacities*  $\implies$  *Acyclicity of  $\succ$* :

Suppose by contradiction that a priority structure  $(\succ, \underline{q})$  contains a cycle. In particular, there are (1)  $i, j, k \in N$  and  $x, y \in X$  such that  $i \succ_x j \succ_x k$  and  $k \succ_y i$  and (2)  $N_x, N_y \subset N \setminus \{i, j, k\}$  such that  $N_x \subset U_x(j)$ ,  $N_y \subset U_y(i)$ ,  $|N_x| = \underline{q}_x - 1$ ,  $|N_y| = \underline{q}_y - 1$ , and  $N_x \cap N_y = \emptyset$ . Let  $(R, q)$  be a problem satisfying the following: Let  $q \geq \underline{q}$  be any capacity vector with  $q_x > \underline{q}_x$  and  $q_y = \underline{q}_y$ . For all  $m \in N_x$  and all  $z \in X \setminus \{x\}$ ,  $x P_m z P_m \emptyset$ . For all  $m \in N_y$  and all  $z \in X \setminus \{y\}$ ,  $y P_m z P_m \emptyset$ . For all  $z \in X \setminus \{x, y\}$ ,  $y P_i x P_i \emptyset P_i z$  and  $x P_k y P_k \emptyset P_k z$ . For all  $z \in X \setminus \{x\}$ ,  $x P_j \emptyset P_j z$ . For all  $m \in N \setminus (N_x \cup N_y \cup \{i, j, k\})$  and all  $z \in X$ , let  $\emptyset P_m z$ . Finally, suppose  $N_x \cup \{i\} N_x \cup \{j, k\}$ .

Clearly,  $\varphi^\succ(R, q)(x) = N_x \cup \{j, k\}$ . Now let  $q'_x = \underline{q}_x$ . Then,  $\varphi^\succ(R, q'_x, q_{-x})(x) = N_x \cup \{i\}$ . Thus, school  $x$  successfully manipulates  $\varphi^\succ$  via underreporting its capacity at  $(R, q)$ .

**Q.E.D.**

**Proof of Corollary 2:** Suppose  $(\succ, \underline{q})$  is imposed and there exist a problem  $(R, q) \in \mathcal{E}$ , and a school  $x \in X$  that can manipulate  $\varphi^\succ$  by underreporting its capacity at  $(R, q)$ . Then by Theorem 1  $(\succ, \underline{q})$  has a cycle. Since  $\underline{q} \geq \underline{q}'$ , the priority structure  $(\succ, \underline{q}')$  shares the same cycle. Then since  $\underline{q}'_x \leq \underline{q}_x$ , school  $x$  can manipulate  $\varphi^\succ$  at  $(R, q)$  in the same way if  $(\succ, \underline{q}')$  is imposed instead. For the converse case, consider Example 1. Note that school  $a$  can manipulate if  $\underline{q}_a = \underline{q}_b = 1$  however this would not be possible if  $\underline{q}_a$  were raised to two.

**Q.E.D.**

**Proof of Theorem 2:**

Consider any mechanism  $\varphi$ . Let a priority structure  $(\succ, q)$  be given. Take a school  $x \in X$  with  $n > q_x$ . Let  $R$  be any preference profile satisfying the following: Choose any  $N'_x \subset N$  with  $|N'_x| = q_x + 1$  and for all

<sup>10</sup>Note that  $y_0 \neq \emptyset$  since  $\varphi^\succ$  is *individually rational*. Otherwise student  $i_0$  would have never applied to school  $x$  at  $(R, q'_x, q_{-x})$ .

$i \in N'_x$  and all  $z \in X \setminus \{x\}$ , let  $x P_i z P_i \emptyset$ . For all  $i \in N \setminus N'_x$  and all  $z \in X$ , let  $\emptyset P_i z$ . Let  $\mu \equiv \varphi(N, R, q)$ . By feasibility, there is  $m \in N'_x$  such that  $\mu(m) \neq x$ . Suppose the pair  $(x, m)$  makes a pre-arrangement. Let  $\mu' \equiv \varphi(N \setminus \{m\}, R_{-m}^{N \setminus \{m\}}, q_{-x}, q_x - 1)$ . Simply letting  $\{m\} \cup \mu'(x) P_x \mu(x)$  is sufficient to show that the pair  $(x, m)$  can successfully manipulate at  $(N, R, q)$ .

**Q.E.D.**

## References

1. Abdulkadiroglu, Atila, P. Pathak, and A. Roth (2005), The New York City High School Match, *American Economic Review, Papers and Proceedings*, **95**, 2, 364-367.
2. Abdulkadiroglu, A., P. Pathak, A. Roth, and T. Sönmez (2005), The Boston Public School Match, *American Economic Review, Papers and Proceedings*, **95**, 2, 368-371.
3. Abdulkadiroğlu, A. and T. Sönmez, School choice: A mechanism design approach, *American Economic Review*, **93**, 729-747.
4. Balinski, M. and T. Sönmez (1999), A tale of two mechanisms: Student placement, *Journal of Economic Theory*, **84**, 73-94.
5. Dubins, L.E. and D.A. Freedman (1981), Machiavelli and the Gale-Shapley algorithm, *American Mathematical Monthly*, **88**, 485-494.
6. Ehlers, L. (2002), Coalitional strategy-proof house allocation, *Journal of Economic Theory*, **105**, 298-317.
7. Ergin, H. (2002), Efficient resource allocation on the basis of priorities, *Econometrica*, **70**, 2489-2497.
8. Ergin, H. and T. Sönmez (2006), Games of school choice under the Boston mechanism, *Journal of Public Economics*, **90**, 215-237.
9. Gale D. and L. Shapley (1962), College admissions and the stability of marriage, *American Mathematical Monthly*, **69**, 9-15.
10. Kesten, O. (2006), On two competing mechanisms for priority based allocation problems, *Journal of Economic Theory*, **127**, 155-171.
11. Konishi, H. and Unver, U. (2006), Games of capacity manipulation in hospital-intern markets, *Social Choice and Welfare*, **27**, 3-24.
12. Kojima, F. (2006), Mixed Strategies in Games of Capacity Manipulation in Hospital-Intern Market, *Social Choice and Welfare* **27**, 25-28.
13. Kojima, F. and Pathak, P. (2007), Incentives and Stability in Large Two-Sided Matching Markets, mimeo.
14. McVitie D. and B. Wilson (1970), Stable marriage assignment for unequal sets, *BIT*, **10**, 295-309.
15. Pápai, S., Strategy-proof assignment by hierarchical exchange, *Econometrica* **68** (2000), 1403-1433.
16. Roth, A. (1985), The college admissions problem is not equivalent to the marriage problem, *Journal of Economic Theory*, **36**, 277-288.

17. Roth, A. and E. Peranson. (1999), The effects of a change in the NRMP matching algorithm, *American Economic Review*, **89**, 748-780.
18. Roth, A. (1982), The economics of matching: Stability and incentives, *Mathematics of Operations Research*, **7**, 617-628.
19. Roth, A and U.G. Rothblum (1999), Truncation strategies in matching markets –In search of advice for participants, *Econometrica* **67**, 21-43.
20. Roth, A. and M. Sotomayor (1990), Two-sided matching, New York: Cambridge University Press.
21. Roth, A. (1991), A Natural Experiment in the Organization of Entry-Level Labor Markets: Regional Markets for New Physicians and Surgeons in the United Kingdom, *American Economic Review*, **81**, 415-440
22. Sönmez, T. (1997), Manipulation via capacities in two-sided matching markets, *Journal of Economic Theory*, **77**, 197-204.
23. Sönmez, T. (1999), Can pre-arranged matches be avoided in two-sided matching markets, *Journal of Economic Theory*, **86**, 148-156.