The B.E. Journal of Theoretical Economics

Contributions

Volume 7, *Issue* 1 2007 *Article* 32

When Can Manipulations be Avoided in Two-Sided Matching Markets? – Maximal Domain Results

Fuhito Kojima*

*Harvard University, kojima@fas.harvard.edu

Recommended Citation
Fuhito Kojima (2007) "When Can Manipulations be Avoided in Two-Sided Matching Markets? – Maximal Domain Results," *The B.E. Journal of Theoretical Economics*: Vol. 7: Iss. 1 (Contributions), Article 32.
Available at: http://www.bepress.com/bejte/vol7/iss1/art32

When Can Manipulations be Avoided in Two-Sided Matching Markets? – Maximal Domain Results*

Fuhito Kojima

Abstract

In two-sided matching markets, stable mechanisms are vulnerable to various kinds of manipulations. This paper investigates conditions for the student-optimal stable mechanism (SOSM) and the college-optimal stable mechanism (COSM) to be immune to manipulations via capacities and pre-arranged matches. For SOSM, we find that strongly monotone preferences in populations and weakly maximin preferences are the maximal domains of college preferences that guarantee immunity to manipulations via capacities and pre-arranged matches, respectively. In contrast, COSM is susceptible to both manipulations whenever colleges have multiple positions.

KEYWORDS: two-sided matching, stability, manipulation via capacities, manipulation via pre-arranged matches

*I thank Hideo Konishi for his question that inspired the current research. I am also grateful to the Co-Editor, an anonymous referee, Aytek Erdil, Onur Kesten, Parag Pathak, Alvin Roth, Tayfun Sonmez, Satoru Takahashi, Utku Unver, Yosuke Yasuda and participants at INFORMS Annual Meeting 2006 for comments and discussion.

1. Introduction

The theory of two-sided matching considers matching between two types of agents, for example colleges and students. The theory has interested researchers for its theoretical appeal and relevance in real-life applications. A matching is stable if there is no individual agent who is matched to unacceptable partner, and there is no pair of agents who prefer each other to their current partners. A mechanism is stable if it produces stable matchings with respect to reported preferences of participants. Empirical studies have shown that stable mechanisms often succeed whereas unstable ones often fail in the real world applications.[1]

Although stable mechanisms have a number of desirable properties, they are vulnerable to various types of manipulations by participants. Dubins and Freedman (1981) and Roth (1982) show that any stable mechanism is manipulable via preference lists: reporting a preference list different from the true preference list may make a participant better off than reporting a true preference list. Sönmez (1997b) and Sönmez (1999) show that other types of manipulations are possible when colleges can admit more than one student. First, any stable mechanism is manipulable via capacities: colleges may sometimes benefit by underreporting their quotas. Second, any stable mechanism is manipulable via pre-arranged matches: a college and a student may benefit by agreeing to match with each other outside of the formal matching process, before receiving their allocations from the centralized matching mechanism. Manipulation of matching mechanisms seems not only theoretical possibility but a real concern in the field. For instance, Abdulkadiroğlu, Pathak, and Roth (2005) report that underreporting of capacities and matching outside the formal process was one of the major concerns in the school choice program in NYC before it was redesigned recently.

This paper focuses on manipulations via capacities and via pre-arranged matches and investigates when we can expect that stable mechanisms are immune to these manipulations. More specifically, we consider the student-optimal stable mechanism (SOSM) and the college-optimal stable mechanism (COSM), which are by far the most popular stable mechanisms, and present conditions on college preferences under which these manipulations are impossible.

Our first focus is on manipulation via capacities. A preference relation of a college is strongly monotone in population if it prefers a larger number of acceptable students than a smaller number of students irrespective of desirability of each student. If a college preference is strongly monotone in population,

[1] See Roth (2002) for evidence.

then that college can not manipulate SOSM via capacities (Konishi and Ünver 2006). We find that strong monotonicity is the *weakest* condition, or a *maximal domain*, for such a conclusion. More specifically, we show that if a college preference is not strongly monotone in population, then there exist preferences and capacities of other colleges and students under which that college can manipulate SOSM via capacities. A much more negative result holds for COSM: if a college has more than one position, then there exists preferences of other colleges and students under which that college can manipulate COSM via capacities.

Then we investigate manipulations via pre-arranged matches. A preference relation of a college is weakly maximin if it prefers a group of students to another if both fill its quota and the least preferred student of the former is more desirable than the least preferred student of the latter. We show that, if a college preference is weakly maximin, then that college can not manipulate SOSM via pre-arranged matches.[2] Moreover, we find that if a college preference is not weakly maximin, then there exists preferences and capacities of other colleges and students under which that college can manipulate SOSM via pre-arranged matches. The result for COSM is again more negative: if a college has more than one position, then there exists preferences and capacities of other colleges and students under which that college can manipulate COSM via pre-arranged matches.

The conditions that prevent the possibility of manipulation are quite restrictive. Hence this paper suggests that the impossibility of incentive compatible stable mechanisms are not merely a theoretical possibility but potentially a real concern. This point is further discussed in the last section.

SOSM is guaranteed to be immune to these manipulations only when preferences are strongly monotone in population and weakly maximin, each of which is a strong condition. Even so, these are not trivial classes of preferences. By contrast, neither strong monotonicity in population nor weak maximin property guarantees non-manipulability under COSM. Indeed, whenever a college has more than one position, each of these manipulations is possible depending on preferences of other market participants. These observations may give some support to using SOSM rather than COSM for real-life applications. Interestingly National Resident Matching Program (NRMP), which is an organization providing matching of medical residents and hospitals in the U.S., changed its

[2]To our knowledge, the weak maximin property is a new condition, while it is related to maximin properties proposed by Sotomayor (2004) and Baïou and Balinski (2000). As opposed to the case of manipulations via capacities, no condition has previously been proposed that guarantees non-manipulability via pre-arranged matches.

matching mechanism from a variant of COSM to a variant of SOSM in the late 1990's. Also, the current school assignment mechanism in NYC builds on SOSM.

Manipulations via capacities and via pre-arranged matches are introduced by Sönmez (1997b) and Sönmez (1999), who show impossibility theorems mentioned above for these manipulations. Konishi and Ünver (2006) and Kojima (2006b) further analyze equilibria of games under SOSM and COSM in which strategies of colleges are announcement of their quotas. Konishi and Ünver (2006) is closely related to the present paper as they introduce strong monotonicity in population and show that it implies non-manipulability via capacities under SOSM.

Our paper is also closely related to Kesten (2006), who also investigates conditions under which SOSM is immune to manipulations via capacities and pre-arranged matches. Although the paper and ours are similar in spirit, there are several differences. First, Kesten (2006) assumes that colleges have private information about their preference lists, which may not agree with the (publicly known) priority structure with which SOSM is carried out. By contrast, the current paper follows the previous literature and considers situations where preference lists are publicly known and stable mechanisms are based on them. Second, Kesten (2006) derives a condition on the *entire priority structure of all the colleges* such that *no college* can manipulate. In contrast, this paper obtain conditions in terms of preferences of an *individual college* under which that *particular college* cannot manipulate. Our results and his are logically unrelated.

More generally, manipulations in two-sided matching markets have been studied from several directions, following seminal papers by Dubins and Freedman (1981) and Roth (1982) mentioned above. Roth (1984), Gale and Sotomayor (1985), Ma (1995), Shin and Suh (1996), Alcalde (1996) and Sönmez (1997a), among others, analyze equilibrium outcomes in games under SOSM. Motivated by applications in such real-life markets as NRMP, Roth and Peranson (1999), Immorlica and Mahdian (2005) and Kojima and Pathak (2006) investigate two-sided markets with a large number of participants. These papers show that scopes of these manipulations become small as the number of market participants increases. Equilibria of games induced by some unstable mechanisms also attract attention because of their widespread use in the field. Ergin and Sönmez (2006) characterize the set of Nash equilibrium outcomes under a popular unstable mechanism called the Boston mechanism, which Kojima (2006a) also analyzes.

The paper proceeds as follows. Section 2 presents the model. Section 3 investigates manipulations via capacities. Section 4 investigates manipulations via pre-arranged matches. Section 5 concludes. Proofs are in the Appendix.

2. Model

A market is tuple $\Gamma = (S, C, \succ, q)$. S and C are finite and disjoint sets of students and colleges. We denote $N = S \cup C$. $\succ = (\succ_i)_{i \in N}$. For each student $s \in S$, $\succ_s$ is a strict preference relation over C and being unmatched (being unmatched is denoted by $\emptyset$). For each college $c \in C$, $\succ_c$ is a strict preference relation over the set of subsets of students. If $j \succ_i \emptyset$, then j is said to be **acceptable** to i. The non-strict counterparts of $\succ_i$ is denoted by $\succeq_i$. $q = (q_c)_{c \in C}$ is a profile of non-negative integers, called quotas.

For each college $c \in C$, its preference relation $\succ_c$ is **responsive** if (i) for any $s, s' \succ_c \emptyset$, and any $S' \succ_c \emptyset$ with $s, s' \notin S'$, we have $s \cup S' \succeq_c s' \cup S' \Leftrightarrow s \succeq_c s'$, (ii) for any $s \in S$ and any $S' \succ_c \emptyset$ with $s \notin S'$, we have $s \cup S' \succeq_c S' \Leftrightarrow s \succeq_c \emptyset$ (Roth 1985).[3] That is, the ranking of a student is independent of her colleagues. We assume that college preferences are responsive throughout the paper.

A **matching** μ is a correspondence from $C \cup S$ to $C \cup S$ such that (i) for every $s \in S$, $|\mu(s)| \leq 1$, and $\mu(s) = \emptyset$ if $\mu(s) \notin C$, (ii) $\mu(c) \subseteq S$ and $|\mu(c)| \leq q_c$ for every $c \in C$, and (iii) $\mu(s) = c$ if and only if $s \in \mu(c)$. For any matchings μ and μ' and for any $i \in N$, we write $\mu \succ_i \mu'$ if and only if $\mu(i) \succ_i \mu'(i)$.

Given a matching μ, we say that it is **blocked** by (s, c) if $c \succ_s \mu(s)$ and either (i) $s \succ_c s'$ for some $s' \in \mu(c)$ or (ii) $|\mu(c)| < q_c$ and $s \succ_c \emptyset$. A matching μ is **individually rational** if for every $i \in N$ and any $j \in \mu(i)$, $j \succeq_i \emptyset$. A matching μ is **stable** if it is individually rational and is not blocked. A mechanism is a systematic way of assigning students to colleges. A stable mechanism is a mechanism that yields a stable matching for any market. We denote a generic mechanism by ϕ We write the matching that is realized in $(S, C, \succ, q)$ when ϕ is run by $\phi(S, C, \succ, q)$.

We consider the following **student-optimal stable mechanism (SOSM)**, denoted by ϕ^S, which is analyzed by Gale and Shapley (1962).[4]

- Step 1: Each student applies to her first choice college. Each college rejects the lowest-ranking students in excess of its capacity and all unacceptable students among those who applied to it, keeping the rest of students temporarily (so students not rejected at this step may be rejected in later steps.)

[3] We sometimes abuse notation and denote a singleton set $\{x\}$ simply by x.

[4] SOSM is known to produce a stable matching that is unanimously most preferred by every student among all stable matchings (Gale and Shapley (1962)).

In general,

- Step t: Each student who was rejected in Step (t-1) applies to her next highest choice. Each college considers these students *and* students who are temporarily held from the previous step together, and rejects the lowest-ranking students in excess of its capacity and all unacceptable students, keeping the rest of students temporarily (so students not rejected at this step may be rejected in later steps.)

The algorithm terminates either when every student is matched to a college or every unmatched student has been rejected by every acceptable college. The algorithm always terminates in a finite number of steps. Gale and Shapley (1962) show that the resulting matching is stable.
We also consider the following **college-optimal stable mechanism (COSM)**, denoted by ϕ^C.

- Step 1: For each $c \in C$, c makes offers to its q_c most preferred acceptable students as long as there are such students. Each student keeps an acceptable college she prefers most if such a college exists, and rejects every other college (so colleges not rejected at this step may be rejected in later steps.)

In general,

- Step t: For each $c \in C$, c makes offers to its q_c most preferred acceptable students who has not yet rejected c as long as there are such students. Each student keeps an acceptable college she prefers most, and rejects every other college (so colleges not rejected at this step may be rejected in later steps.)

The algorithm terminates either when every college has filled their positions with students or every college with vacant positions has made an offer to every student acceptable to it. The algorithm always terminates in a finite number of steps. Similarly to SOSM, COSM produces a stable matching.

3. Manipulation via capacities

We say that $c \in C$ can **manipulate ϕ via capacities** in $(S, C, \succ, q)$ if there exists $q'_c \in \{0, 1, \ldots, q_c - 1\}$ such that

$$\phi(S, C, \succ, q'_c, q_{-c}) \succ_c \phi(S, C, \succ, q).$$

A mechanism ϕ is **manipulable via capacities** if there exists a market $(S, C, \succ, q)$ and some $c \in C$ such that c can manipulate ϕ via capacities in

$(S, C, \succ, q)$. Sönmez (1997b) shows that any stable mechanism is manipulable via capacities.[5]

Given the above impossibility result, a natural direction is to give a restriction on preferences under which a college c cannot manipulate via capacities. It is clear that c cannot manipulate via capacities if $|C| = 1$. Thus, assume $|C| \geq 2$. Also, we assume that there are at least q_c students that are acceptable to c.

Our first focus is on SOSM. Consider the following class of preferences due to Konishi and Ünver (2006).

Definition 1 (Konishi and Ünver (2006)). Preference $\succ_c$ is **strongly monotone in population** if, if $q_c \geq |S'| > |S''|$ and every $s \in S'$ is acceptable, then $S' \succ_c S''$.

A college with strongly monotone preference in population prefers a larger number of acceptable students to a smaller number of students, irrespective of preferability of each student. A college may have such preference when filling positions is much more important than student quality.

Result 1 (Theorem 5 of Konishi and Ünver (2006)) **.** *Assume every college has responsive preferences. If $\succ_c$ is strongly monotone in population, then c cannot manipulate ϕ^S via capacities.*

For the original proof, see Konishi and Ünver (2006). We give an alternative proof in the Appendix, where proofs of all the propositions are relegated.

Strong monotonicity in population is a demanding condition. A natural question is whether there are weaker conditions than strong monotonicity in population such that we can obtain a similar conclusion. The answer turns out to be negative. More specifically, we prove that the class of strongly monotone preferences in population is a *maximal domain*, that is, if $\succ_c$ violates strong monotonicity in population, then there exist preferences and capacities of other colleges and students that allow for c to profitably manipulate. In this particular sense, strong monotonicity in population is "necessary" for guaranteeing that SOSM is immune to manipulations via capacities.

Proposition 1. *Assume every college has responsive preferences. Fix S, C, $c \in C$, $\succ_c$ and q_c. If $\succ_c$ is not strongly monotone in population, then there exist preferences and capacities of other colleges and students $(\succ_{-c}, q_{-c})$ such that c can manipulate ϕ^S via capacities. Preferences of colleges other than c can be taken as strongly monotone in population.*

[5] We consider only underreporting of capacities, as it is easily seen that colleges never benefit by overreporting quotas.

Next we consider COSM. Konishi and Ünver (2006) present an example in which a college can manipulate COSM via capacities even though its preference is strongly monotone in population. Indeed, the next proposition shows that COSM is potentially manipulable whenever colleges have more than one position.

Proposition 2. *Assume every college has responsive preferences.*

(1) *If $q_c = 1$, then c cannot manipulate ϕ^C via capacities.*
(2) *Fix S, C, $c \in C$, $\succ_c$ and q_c. If $q_c > 1$, then there exist preferences and capacities of other colleges and students $(\succ_{-c}, q_{-c})$ such that c can manipulate ϕ^C via capacities. Capacities of colleges other than c can be set to one.*

4. Manipulation via pre-arranged matches

When colleges have more than one position to fill, there is concern for manipulation not only within but also outside the formal matching mechanism. Sönmez (1999) introduces the idea of manipulation via pre-arranged matches. Suppose that c and s arrange a match before the centralized matching procedure is executed. Then s does not participate in the formal mechanism and c participates with the number of positions reduced by one.

Motivated by the above scenario, we say that $c \in C$ can **manipulate ϕ via pre-arranged matches** in $(S, C, \succ, q)$ if there exists $s \in S$ such that

$$\phi(S \setminus s, C, \succ_{-s}, q_c - 1, q_{-c})(c) \cup s \succeq_c \phi(S, C, \succ, q)(c), \text{and}$$
$$c \succeq_s \phi(S, C, \succ, q)(s),$$

where at least one of the above relations holds strictly.[6] A mechanism ϕ is **manipulable via pre-arranged matches** if there exists a market $(S, C, \succ, q)$ and some $c \in C$ such that c can manipulate ϕ via pre-arranged matches in $(S, C, \succ, q)$. In words, a college can manipulate via pre-arranged matches if it can find a student such that both parties have incentives to do so: both of them are made weakly better off and at least one of them is made strictly better off with pre-arranging matches than without doing so. Sönmez (1999) shows that any stable mechanism is manipulable via pre-arranged matches.[7]

Matching outside the centralized mechanism is prohibited in some markets. Even so, the student and college can effectively "pre-arrange" a match by

[6] College preferences over individual students in the market without s is the restriction of $\succ_c$ on $2^{S \setminus s}$. We however abuse notation and continue to use $\succ_c$.

[7] Although manipulations via capacities and via pre-arranged matches look similar, they are logically unrelated. That is, a college can manipulate a stable mechanism via capacities but not via pre-arranged matches in some markets, and vice versa.

listing each other on the top of their preference lists under stable mechanisms. Thus manipulation via pre-arranged matches may be an important concern even in such markets.

Given that no stable mechanism is non-manipulable via pre-arranged matches, a natural question is, as for manipulations via capacities, what conditions guarantee non-manipulability. Specifically this section investigates conditions under which a college cannot manipulate SOSM or COSM via pre-arranged matches. It is clear that c cannot manipulate via pre-arranged matches if $|C| = 1$ or, equivalently, $C = \{c\}$. Thus, assume $|C| \geq 2$. Also, we assume that there are at least $q_c + 1$ students that are acceptable to c.

As in the last section, we begin our analysis by focusing on SOSM. Consider the following class of college preferences.

Definition 2. Preference $\succ_c$ is **weakly maximin** if the following holds: If $|S'| = |S''| = q_c$, every $s \in S'$ is acceptable, and the least preferred student in S' is strictly preferred to the least preferred student in S'', then $S' \succ_c S''$.

Colleges with weakly maximin preferences compare the worst students of two groups of students that fill its quotas, and prefer the group whose worst student is better. One could think of such preferences as an extreme form of risk aversion. The condition may be satisfied by a certain factory in which the productivity of its assembly line crucially depends on the performance of the worst worker. Note that the condition does not impose any restriction when not all the positions are filled by students.

Sotomayor (2004) and Baïou and Balinski (2000) propose conditions related to our maximin preferences. In Sotomayor (2004)'s definition (see also Section 6.3 of Echenique and Oviedo (2006) who discuss the related concept of strong substitutability in a model without the responsiveness assumption), $\succ_c$ is **maximin** if, for any S', S'' such that $S'' \succ_c S'$ and every $s \in S'$ is acceptable, we have that (i) every $s \in S''$ is acceptable to c, and (ii) if $|S'| = q_c$, then the least preferred student in $S'' \setminus S'$ is preferred to the least preferred student in $S' \setminus S''$. In an alternative definition by Baïou and Balinski (2000), $\succ_c$ is **maximin** if, for any S' and S'', $S' \succ_c S''$ if $|S'| \geq |S''|$ and the least preferred student in S' is strictly preferred to the least preferred student in S''. It is easy to see that either of them is stronger than our weak maximin condition (hence our terminology "weak" maximin property).

Weakly maximin preferences turn out to play an important role under SOSM, as the next proposition asserts. In the propositions below, we maintain the assumption that there are at least another college in addition to c and at least $q_c + 1$ students that are acceptable to c.

Proposition 3. *Assume every college has responsive preferences.*

(1) *If $\succ_c$ is weakly maximin, then c cannot manipulate ϕ^S via pre-arranged matches.*

(2) *Fix S, C, $c \in C$, $\succ_c$ and q_c. If $\succ_c$ is not weakly maximin, then there exist preferences and capacities of other colleges and students $(\succ_{-c}, q_{-c})$ such that c can manipulate ϕ^S via pre-arranged matches. Preferences of colleges other than c can be taken as weakly maximin.*

The outline of part (1) of the proposition is as follows. We show that, in order to profitably manipulate SOSM via pre-arranged matches, a college has to pre-arrange a match with a student who is strictly less preferred to those who would be matched under SOSM without pre-arrangement (the disadvantage of being matched with a less desirable student should be compensated by matching to a better set of students in the centralized matching mechanism after pre-arrangement). When the college has weakly maximin preferences, the cost of matching with such a student is never offset by matching with a better set of students for the remaining positions, implying that the student cannot manipulate SOSM via pre-arranged matches.
For COSM, a very negative result holds as in the case of manipulation via capacities.

Proposition 4. *Assume every college has responsive preferences.*

(1) *If $q_c = 1$, then it cannot manipulate ϕ^C via pre-arranged matches.*

(2) *Fix S, C, $c \in C$, $\succ_c$ and q_c. If $q_c > 1$, then there exist preferences and capacities of other colleges and students $(\succ_{-c}, q_{-c})$ such that c can manipulate ϕ^C via pre-arranged matches. Capacities of colleges other than c can be set to one.*

5. Concluding remarks

This paper investigated conditions under which the student-optimal stable mechanism (SOSM) and the college-optimal stable mechanism (COSM) are immune to manipulations via capacities and pre-arranged matches. For SOSM, we find that strong monotonicity in population and weak maximin property are the weakest conditions on college preferences which guarantee that the college cannot manipulate the mechanism via capacities and via pre-arranged matches, respectively. By contrast, COSM is susceptible to both manipulations whenever the college has more than one position.

Our study is part of a research program to find restrictions on (profiles of) preferences for stable matching mechanisms to have desirable incentive properties. Alcalde and Barberà (1994) take this approach for manipulation via preferences, and Konishi and Ünver (2006), and Kesten (2006) for manipulation via capacities and manipulation via pre-arranged matches. One interpretation of these papers as well as ours is that the conditions that prevent the possibility of manipulation are often quite restrictive, so the impossibility results in the literature are not just a theoretical possibility but potentially a real concern. Large market arguments such as Kojima and Pathak (2006) may be one way to overcome such negative conclusions.[8] We also speculate that an algorithmic approach to matching markets that does not impose restrictions on preferences, such as Echenique and Oviedo (2004), Echenique and Yenmez (2007) and Kojima (2007), may also be useful in strategic issues in stable mechanisms.

As we have seen, scopes of manipulations seem large even for SOSM, as strong monotonicity in population and weak maximin property are certainly strong assumptions. Even so, these assumptions may be reasonable in some situations. One example would be hospital preferences in the hospital-resident matching market in Japan, which recently adopted the doctor-optimal stable mechanism. Under the current system, hospitals receive substantial subsidy from the government for each intern they hire. Thus hospitals have strong incentives to hire as many interns as possible, which may make their preferences strongly monotone in population. Indeed, there is a regulation which restricts the maximal number of interns hospitals can hire, depending on their sizes. In this application, we may as well expect that hospitals will not under-report their capacities. Interestingly, a number of markets such as NRMP and the public school system in NYC adopted mechanisms based on SOSM in addition to the Japanese medical matching market. Our results may give some support to their choices of the mechanisms, as they suggest that manipulations may be more difficult under SOSM than under COSM.

Although logically independent, manipulations via capacity and pre-arranged matches seem to be closely related. In a sense, manipulation via capacities could be seen as pre-arranging matches with a "null student." Further exploring connection of this sort is an interesting direction of future research. One modeling issue there will be how to extend the definition of pre-arranged matches to allow for pre-arranging with more than one student. This generalization seems to be nontrivial, since students offered such generalized pre-arrangement options would be in a strategic situation in which benefit of pre-arrangement depends on decisions of other students.

[8]See Kojima and Manea (2006) for a similar large market argument on a random assignment mechanism studied by Bogolomania and Moulin (2001).

Appendix A: Proofs

A.1. Alternative proof of Result 1. Let $\mu = \phi^S(S, C, \succ, q)$ and $\mu' = \phi^S(S, C, \succ, q'_c, q_{-c})$ where $q'_c < q_c$. Suppose that fewer than or equal to q'_c students who are acceptable to c have applied to c during the execution of SOSM in $(S, C, \succ, q)$. Then c never rejects an acceptable student either under reported quota $q_c > q'_c$ or q'_c and the algorithm proceeds in exactly the same ways in $(S, C, \succ, q)$ and $(S, C, \succ, q'_c, q_{-c})$. Therefore $\mu(c) = \mu'(c)$, implying that c does not benefit strictly by reporting q'_c instead of q_c. Now suppose that at least $q'_c + 1$ acceptable students have applied to c during the execution of SOSM in $(S, C, \succ, q)$. This implies that c is matched to at least $q'_c + 1$ and at most q_c acceptable students at μ by definition of the algorithm.[9] Since $|\mu'(c)| \leq q'_c$ by definition, this implies that reporting q'_c instead of q_c makes c strictly worse off, whose preference is strongly monotone in population.

A.2. Proof of Proposition 1. Suppose that $\succ_c$ is not strongly monotone in population. Then there exist S' and S'' such that $q_c \geq |S'| > |S''|$, no $s \in S'$ is unacceptable to c, and $S'' \succ_c S'$. Then consider T defined as the $|S'| - 1$ most preferred students by c in $S' \cup S''$. By construction and $|S''| \leq |S'| - 1$, and noting that every student in S' is acceptable and $\succ_c$ is responsive, we have that (i) every student in T is acceptable to c, and (ii) $T \succeq_c S'' \succ_c S'$. Let $S' = \{s_1, \ldots, s_{|S'|}\}$ and $T \setminus S' = \{\tilde{s}_1, \ldots, \tilde{s}_{|T \setminus S'|}\}$, where $s_{|S'|} \succ_c s_{|S'|-1} \succ_c \cdots \succ_c s_1$. Note that $T \setminus S' \neq \emptyset$.

Consider the following preferences and capacities of S and $C \setminus \{c\}$. Fix some $c' \neq c$ (such c' exists by assumption that $|C| \geq 2$). Let preferences of colleges different from c be strongly monotone in population, satisfying

$$s_{|S'|} \succ_{c'} \cdots \succ_{c'} s_1 \succ_{c'} \tilde{s}_{|T \setminus S'|} \succ_{c'} \cdots \succ_{c'} \tilde{s}_1 \succ_{c'} \emptyset \succ_{c'} s$$
$$\text{for any } s \notin S' \cup T, q_{c'} = |T \setminus S'|,$$
$$\emptyset \succ_{c''} s \text{ for any } c'' \neq c, c' \text{ and any } s \in S.$$

Student preferences are given by

$$c \succ_s c' \succ_s \emptyset \succ_s c'' \text{ for any } s \in S' \text{ and } c'' \neq c, c',$$
$$c' \succ_s c \succ_s \emptyset \succ_s c'' \text{ for any } s \in T \setminus S' \text{ and } c'' \neq c, c',$$
$$\emptyset \succ_s c'' \text{ for any } s \notin S' \cup T \text{ and any } c'' \in C.$$

First suppose that c reports its quota truthfully. Then ϕ^S proceeds as follows: at the first step, every student in S' applies to c, every student in $T \setminus S'$

[9] Note that SOSM proceeds in an identical manner both in $(S, C, \succ, q)$ and $(S, C, \succ, q'_c, q_{-c})$ until the $(q'_c + 1)$st student applies to c.

applies to c', and no other student applies to any college. No student is rejected, terminating the algorithm. Therefore, at the realized matching μ, we have that $\mu(c) = S'$ and $\mu(c') = T \setminus S'$.

Next let c report capacity $q'_c = |T| = |S'| - 1$. Then ϕ^S proceeds as follows: at the first step, every student in S' applies to c, every student in $T \setminus S'$ applies to c', and no other student applies to any college. c rejects the least preferred student $s_1 \in S'$ since there are one more applicants than its reported capacity $q'_c = |S'| - 1$. s_1 applies to c' in the second step. Then c' accepts s_1 and rejects $\tilde{s}_1 \in T \setminus S'$ according to its capacity and preferences. At the third step, $\tilde{s}_1$ applies to c since c is her second choice college. c accepts $\tilde{s}_1$ and rejects $s_2 \in S' \setminus T$, according to its reported capacity and preferences. The algorithm repeats itself in an analogous way, until every student in T is matched to c. Therefore we have $\mu'(c) = T$ at the realized matching μ', completing the proof.

A.3. Proof of Proposition 2. Part (1). If $q_c = 1$, then $0 \leq q'_c < q_c$ implies $q'_c = 0$. Reporting a capacity $q'_c = 0$ does not make c strictly better off, since c is not matched to any student when reporting $q'_c = 0$, and c is either matched to one acceptable student or not matched to any student when reporting q_c by the definition of COSM.

Part (2). Fix $c' \neq c$ and $s, s' \in S$ such that $s \succ_c s' \succ_c \emptyset$ (there exist such agents by assumption). Let preferences of colleges different from c satisfy

$$s' \succ_{c'} s \succ_{c'} \emptyset \succ_{c'} s'' \text{ for any } s'' \neq s, s', q_{c'} = 1,$$
$$\emptyset \succ_{c''} s'' \text{ for any } c'' \neq c, c' \text{ and any } s'' \in S.$$

Let student preferences be given by

$$c' \succ_s c \succ_s \emptyset \succ_s c'' \text{ for any } c'' \neq c, c',$$
$$c \succ_{s'} c' \succ_{s'} \emptyset \succ_{s'} c'' \text{ for any } c'' \neq c, c',$$
$$\emptyset \succ_{s''} c'' \text{ for any } s'' \neq s, s' \text{ and any } c'' \in C.$$

Suppose that c reports its true quota $q_c \geq 2$. Then COSM proceeds as follows. First note that college c is rejected by every student that c strictly prefers to s, if any, since every such student finds c unacceptable by assumption. Thus consider the step just after c is rejected by every such student. At that step c makes offers to s and s', c' makes an offer to s', and no other college makes any offer. Since $c \succ_{s'} c' \succ_{s'} \emptyset$, s' rejects c' and accepts c. Then at the second step, c' makes an offer to its second choice s, who accepts the offer from c' while rejecting c. The algorithm terminates and the resulting matching μ gives $\mu(c) = s'$ and $\mu(c') = s$.

Now suppose that c reports $q'_c = 1 < q_c$ instead. Then COSM proceeds as follows. At the first step, c makes an offer to s, c' makes an offer to s', and no other college makes any offer. s accepts c since $c \succ_s \emptyset$ and s' accepts c' since $c' \succ_{s'} \emptyset$, which terminates the algorithm. The resulting matching μ' gives $\mu'(c) = s$ and $\mu'(c') = s'$. c benefits by reporting $q'_c < q_c$ since $\mu'(c) = s \succ_c s' = \mu(c)$, completing the proof.

A. 4. Proof of Proposition 3. Part (1). Let $\mu = \phi^S(S, C, \succ, q)$ be the matching under SOSM when no college engages in pre-arranged matches. Obviously c cannot manipulate ϕ^S via pre-arranged matches when $|\mu(c)| < q_c$. Suppose that $|\mu(c)| = q_c$. For any stable mechanism ϕ, if c can manipulate ϕ via pre-arranged matches with student s, then either $s \in \mu(c)$ or $s' \succ_c s$ for every $s' \in \mu(c)$ (Theorem 2 of Sönmez (1999)). In the latter case c cannot manipulate: Since $\succ_c$ is weakly maximin, matching with a student s such that $s' \succ_c s$ for every $s' \in \mu(c)$ cannot be desirable for c. Thus suppose that c can manipulate ϕ^S with some $s \in \mu(c)$. Consider matching μ' given by

$$\mu'(c') = \begin{cases} \mu(c) \setminus s & \text{if } c' = c, \\ \mu(c') & \text{otherwise.} \end{cases}$$

μ' is stable in $(S \setminus s, C, \succ_{-s}, q_c - 1, q_{-c})$ by stability of μ in $(S, C, \succ, q)$.

Since the matching under SOSM is weakly less preferred to any stable matching by colleges (attributed to Conway in Knuth (1996)) and preferences are responsive,

$$\mu(c) = \mu'(c) \cup s \succeq_c \phi^S(S \setminus s, C, \succ_{-s}, q_c - 1, q_{-c})(c) \cup s.$$

Therefore c cannot manipulate. This is a contradiction, completing the proof.

Part (2). Suppose that $\succ_c$ is not weakly maximin. Then there exist S' and S'' such that $|S'| = |S''| = q_c$, every $s \in S'$ is acceptable to c, the least preferred student in S' is preferred to the least preferred student in S'' and $S'' \succ_c S'$. Denote the least preferred student in S'' by s_0. Consider T composed of s_0 and the $q_c - 1$ most preferred students in $S' \cup S''$. By construction, and noting that every student in S' is acceptable to c and $\succ_c$ is responsive, we have that (i) every student in $T \setminus \{s_0\}$ is acceptable to c, and (ii) $T \succeq_c S'' \succ_c S'$. Let $S' = \{s_1, \ldots, s_{q_c}\}$ and $T \setminus S' = \{s_0, \tilde{s}_1, \tilde{s}_2, \ldots, \tilde{s}_{|T \setminus S'|-1}\}$, where $s_{q_c} \succ_c s_{q_c - 1} \succ_c \cdots \succ_c s_1 \succ_c s_0$.

Consider the following preferences and capacities of S and $C \setminus \{c\}$. Fix some $c' \neq c$ (such c' exists by assumption). Let preferences of colleges different from c be weakly maximin, satisfying

$$s_{q_c} \succ_{c'} \cdots \succ_{c'} s_1 \succ_{c'} \tilde{s}_{|T \setminus S'|-1} \succ_{c'} \cdots \succ_{c'} \tilde{s}_1 \succ_{c'} \emptyset \succ_{c'} s \text{ for any } s \notin S' \cup T \setminus s_0,$$

$$q_{c'} = |T \setminus S'| - 1,$$

$$\emptyset \succ_{c''} s \text{ for any } c'' \neq c, c' \text{ and any } s \in S.$$

Student preferences are given by

$$c \succ_s c' \succ_s \emptyset \succ_s c'' \text{ for any } s \in S' \text{ and } c'' \neq c, c',$$
$$c' \succ_s c \succ_s \emptyset \succ_s c'' \text{ for any } s \in (T \setminus S') \setminus s_0 \text{ and } c'' \neq c, c',$$
$$c \succ_{s_0} \emptyset \succ_{s_0} c'' \text{ for any } c'' \neq c,$$
$$\emptyset \succ_s c'' \text{ for any } s \notin S' \cup T \text{ and any } c'' \in C.$$

Suppose that c does not engage in pre-arranged matches. Then ϕ^S proceeds as follows: at the first step, all the students in S' and s_0 apply to c, every student in $(T \setminus S') \setminus \{s_0\}$ applies to c', and no other student applies to any college. c accepts every student in S' and rejects s_0 and c' accepts all the students according to their preferences and quotas. The algorithm terminates since s_0 is the only student who is rejected in this step, and she has applied to every college that is acceptable to her. Therefore, at the realized matching μ, we have that $\mu(c) = S'$ and $\mu(c') = (T \setminus S') \setminus \{s_0\}$. s_0 is unmatched.

Now let c and s_0 pre-arrange matches, and we will verify that c and s_0 are made strictly better off. With the pre-arranged matches, c has capacity $q_c - 1$ for the centralized matching procedure. Then ϕ^S proceeds as follows: at the first step, every student in S' applies to c, every student in $(T \setminus S') \setminus \{s_0\}$ applies to c', and no other student applies to any college. c rejects the least preferred student $s_1 \in S'$. s_1 applies to c' at the second step. Then c' accepts s_1 and rejects $\tilde{s}_1 \in (T \setminus S') \setminus \{s_0\}$. $\tilde{s}_1$ applies to c. c accepts $\tilde{s}_1$, and rejects s_2, and so on. The algorithm repeats itself until every student in $T \setminus \{s_0\}$ is accepted by c. Therefore c is matched to s_0 via pre-arranged matches and to $T \setminus \{s_0\}$ in the centralized procedure. This implies that c is matched to $T \succ_c S' = \mu(c)$ as a whole, completing the proof.

A. 5. Proof of Proposition 4. Part (1). Suppose that $q_c = 1$ and c manipulates ϕ^C via pre-arranged matches with $s \in S$. Then we have $s \succeq_c \mu(c)$ and $c \succeq_s \mu(s)$, with at least one of them holding strictly. Since $q_c = 1$, this implies that $s \succ_c \mu(c)$ and $c \succ_s \mu(s)$, implying that μ is not stable. This is a contradiction, implying that c cannot manipulate ϕ^C if $q_c = 1$.
Part (2). Suppose $q_c > 1$. Fix $c' \neq c$ and $s_1, s_2, \ldots, s_{q_c}, s_{q_c+1}$ be such that

$$s_1 \succ_c s_2 \succ_c \cdots \succ_c s_{q_c} \succ_c s_{q_c+1} \succ_c \emptyset.$$

(such agents exist by assumption.) Let preferences of colleges different from c be given by

$$s_{q_c} \succ_{c'} s_1 \succ_{c'} \emptyset \succ_{c'} s \text{ for any } s \neq s_1, s_{q_c}, q_{c'} = 1,$$
$$\emptyset \succ_{c''} s \text{ for any } c'' \neq c, c' \text{ and any } s \in S.$$

Let student preferences be given by

$$c' \succ_{s_1} c \succ_{s_1} \emptyset \succ_{s_1} c'' \text{ for any } c'' \neq c, c',$$
$$c \succ_s c' \succ_s \emptyset \succ_s c'' \text{ for any } s = s_2, \ldots, s_{q_c+1} \text{ and any } c'' \neq c, c',$$
$$\emptyset \succ_s c'' \text{ for any } s \neq s_1, \ldots, s_{q_c+1} \text{ and any } c'' \in C.$$

Suppose that c does not engage in pre-arranged matches. Then COSM proceeds as follows. At the first step, c makes offers to $s_1, \ldots, s_{q_c}$, c' makes an offer to s_{q_c}, and no other college makes any offer. Since $c \succ_{s_{q_c}} c' \succ_{s_{q_c}} \emptyset$, s_{q_c} rejects c' and accepts c. Then at the second step, c' makes an offer to its second choice s_1, who accepts the offer while rejecting c. At the third step c makes an offer to s_{q_c+1}, who accepts the offer. The algorithm terminates, and the resulting matching μ gives $\mu(c) = \{s_2, \ldots, s_{q_c+1}\}$ and $\mu(c') = s_1$.

Now suppose that c pre-arranges a match with s_{q_c+1}. Then COSM proceeds as follows. At the first step, c makes an offer to $s_1, \ldots, s_{q_c-1}$, c' makes an offer to s_{q_c}, and no other college makes any offer. Every student accepts her offer, terminating the algorithm. The resulting matching μ', including the student matched through pre-arranged matches, gives $\mu'(c) = \{s_1, \ldots, s_{q_c-1}, s_{q_c+1}\}$ and $\mu'(c') = s_{q_c}$. Therefore $\mu'(c) \succ_c \mu(c)$. Moreover s_{q_c+1} has a weak incentive to engage in pre-arranged matches, since s_{q_c+1} is matched to c in either case. Thus c and s_{q_c+1} successfully pre-arrange, completing the proof.

References

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.

Alcalde, J. (1996): "Implementation of stable solutions to the marriage problem," *Journal of Economic Theory*, 69, 240–254.

Alcalde, J., and S. Barberà (1994): "Top Dominance and the Possibility of Strategy-Proof Stable Solutions to Matching Problems," *Economic Theory*, 4, 417–435.

Baïou, M., and M. Balinski (2000): "Many-to-many matching: stable polyandrous polygamy (or polygomous polyandry)," *Discrete Applied Mathematics*, 101, 1–12.

Bogolomania, A., and H. Moulin (2001): "A New Solution to the Random Assignment Problem," *Journal of Economic Theory*, 100, 295–328.

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

Echenique, F., and J. Oviedo (2004): "Core Many-to-One Matchings by Fixed-Point Methods," *Journal of Economic Theory*, 115, 358–376.

——— (2006): "A theory of stability in many-to-many matching," *Theoretical Economics*, 1, 233–273.

Echenique, F., and M. B. Yenmez (2007): "A Solution to Matching with Preferences over Colleagues," *Games and Economic Behavior*, 59, 46–71.

Ergin, H., and T. Sönmez (2006): "Games of School Choice under the Boston Mechanism," *Journal of Public Economics*, 90, 215–237.

GALE, D., AND L. S. SHAPLEY (1962): "College admissions and the stability of marriage," *American Mathematical Monthly*, 69, 9–15.

GALE, D., AND M. A. O. SOTOMAYOR (1985): "Ms. Machiavelli and the stable matching problem," *American Mathematical Monthly*, 92, 261–268.

IMMORLICA, N., AND M. MAHDIAN (2005): "Marriage, Honesty, and Stability," *SODA 2005*, pp. 53–62.

KESTEN, O. (2006): "On Two Kinds of Manipulation for School Choice Problems," Unpublished mimeo.

KNUTH, D. E. (1996): *Stable marriage and its relation to other combinatorial problems.* American Mathematical Society, Providence.

KOJIMA, F. (2006a): "Games of school choice under the Boston mechanism with general priority structures," mimeo.

——— (2006b): "Mixed Strategies in Games of Capacity Manipulation in Hospital-Intern Markets," *Social Choice and Welfare*, 27, 25–28.

——— (2007): "Finding All Stable Matchings with Couples," mimeo.

KOJIMA, F., AND M. MANEA (2006): "Strategy-Proofness of the Probabilistic Serial Mechanism in Large Random Assignment Problems," Unpublished mimeo, Harvard University.

KOJIMA, F., AND P. A. PATHAK (2006): "Incentives and stability in large two-sided markets," mimeo.

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

MA, J. (1995): "Stable matchings and rematching-proof equilibria in a two-sided matching market," *Journal of Economic Theory*, 66, 352–369.

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

——— (1984): "Misrepresentation and stability in the marriage problem," *Journal of Economic Theory*, 34, 383–387.

——— (1985): "The college admission problem is not equivalent to the marriage problem," *Journal of Economic Theory*, 36, 277–288.

——— (2002): "The economist as engineer: Game theory, experimentation, and computation as tools for design economics," *Econometrica*, 70, 1341–1378.

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, 748–780.

SHIN, S., AND S.-C. SUH (1996): "A mechanism implementing the stable rule in marriage problems," *Economics Letters*, 51, 185–189.

SÖNMEZ, T. (1997a): "Games of Manipulation in Marriage Problems," *Games and Economic Behavior*, 20, 169–176.

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

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

SOTOMAYOR, M. A. O. (2004): "Implementation in the many-to-many matching market," *Games and Economic Behavior*, 46, 199–212.