Discussion Paper No. 1247

Stable Matchings and the Small Core in Nash Equilibrium in the College Admissions Problem

by

Jinpeng Ma
Northwestern University
Rutgers University-Camden, NJ 08102
jinpeng@crab.rutgers.edu

Last version October 27, 1997
This version November 29, 1998

Math Center web site:
http://www.kellogg.nwu.edu/research/math

STABLE MATCHINGS AND THE SMALL CORE IN NASH EQUILIBRIUM IN THE COLLEGE ADMISSIONS PROBLEM*

Jinpeng Ma
Department of Economics
Rutgers University-Camden, NJ 08102
jinpeng@crab.rutgers.edu.

Last version October 26, 1997

This version November 29, 1998

Abstract

Both rematching proof and strong equilibrium outcomes are stable with respect to the true preferences in the marriage problem. We show that not all rematching proof or strong equilibrium outcomes are stable in the college admissions problem. But we show that both rematching proof and strong equilibrium outcomes with truncations at the match point are all stable in the college admissions problem. Further, all true stable matchings can be achieved in both rematching proof and strong equilibrium with truncations at the match point.

We show that any Nash equilibrium in truncations admits one and only one matching, stable or not. Therefore, the core at a Nash equilibrium in truncations must be small. But examples exist such that the set of stable matchings with respect to a Nash equilibrium may contain more than one matching. Nevertheless, each Nash equilibrium can only admit at most one true stable matching. If, indeed, there is a true stable matching at a Nash equilibrium, then the only possible equilibrium outcome will be the true stable matching, no matter how players manipulate their equilibrium strategies and how many other unstable matchings are there at the Nash equilibrium. Thus, we show that a necessary and sufficient condition for the stable matching rule to be implemented in a subset of Nash equilibria by a direct revelation game induced by a stable matching mechanism is that every Nash equilibrium profile in that subset admits one and only one true stable matching. *Journal of Economic Literature* Classification Numbers: C78, D71.

*This paper is a revision of the paper "Manipulation and Stability in a College Admissions Problem" circulated since 1994. I thank Rich McLean, Abraham Neyman, Sang-Chul Suh, Mark Satterthwaite, and Tetsuji Yamada for helpful discussions. I am grateful to the Kellogg G.S.M. at the Northwestern University for the hospitality for my visit. Any errors are mine.

1 Introduction

A great number of impossibility theorems have shown the fundamental difficulty to design an efficient mechanism under which players have the best interests to reveal their true preferences. These theorems imply that when players are confronted with a game induced by such a mechanism. players have incentive to manipulate their reported preferences. Since misreported information is different from the underlying truth. the objective of a designed mechanism may well be distorted by the misreported information. This calls up a question how the objective of a designed mechanism can be fulfilled with the misreported information. This paper studies this issue for a class of stable (core) matching mechanisms in the context of the college admissions problem.

In the college admissions problem Roth's impossibility theorems show that no stable matching mechanism exists that makes it a dominant strategy for all players to state their true preferences. Moreover. Roth showed that any individually rational matching can be supported by some Nash equilibrium of the game induced by a stable matching mechanism. Clearly. not all individually rational matchings are stable with respect to the true preferences. Thus. there are chances for a stable matching mechanism that generates a stable matching for each reported preference profile to generate unstable matchings with respect to the true preferences in Nash equilibrium.

The potential problem for a stable matching mechanism to generate unstable matchings with respect to the true preferences is in the contrast with the empirical findings in Roth (1984a. 1990b. 1991). which showed that the centralized matching mechanisms in the labor markets for new physicians that generate stable matchings with respect to each reported preference profile have succeeded to resolve the market failures. while matching mechanisms that generate unstable matchings with respect to the reported preferences have typically failed. The empirical evidences strongly support the hypothesis that a centralized stable matching mechanism may produce a stable matching with respect to not only the reported preferences but also the true preferences. Then. under what conditions does a stable matching mechanism always generate a true stable matching?

Our first focus is on the rematching proof and strong equilibrium. In the marriage problem. Ma (1995) showed that all rematching strong equilibrium outcomes are stable: Ma (1994) and Shin and Suh (1996) showed that all strong equilibrium outcomes are stable: Sönmez (1997) generalized these results to the $\mathcal{G}$-proof Nash equilibrium for the $\mathcal{G}$-core. But we find that both rematching proof and strong equilibrium outcomes are not always stable in the college admissions problem. unlike the marriage problem. Therefore. we consider a class of simple strategies called truncations by Roth and Vande Vate (1991) and Roth and Rothblum (1996). A truncation strategy for a student or a college is a preference ordering which is order-consistent with his or her true preference but has fewer acceptable colleges or students. This class of strategies exclude some complicated strategies like the changes in orders that may be profitable in manipulation. But Roth and Rothblum (1996) showed why it is plausible for players to consider this class of profitable strategies in truncations in manipulation, when players have very little information about the preferences of other players. But examples exist such that both rematching proof and strong equilibrium outcomes with truncations are not stable. Therefore, we use the truncations at the match point (e.g., deleting the (k+1)th and higher choices if a student is matched to his kth choice) in Roth and Peranson (1997b). We show that examples exist such that a Nash equilibrium outcome with truncations at the match point are not stable. But we show that both rematching proof and strong equilibrium outcomes with truncations at the match point are all stable.

Our result provides a support for the hypothesis illustrated above to some degree. Our result shows that some meaningful refinement of the Nash equilibrium exists such that all outcomes in the refined equilibrium notion are stable. On the other side, truncations at the match point are required in order to obtain a stable matching in both rematching proof and strong equilibrium. This may be objected on the ground that truncations at the match point are too strong, since players may need information about the outcome at the match point before reporting truncations at the match point. This objection may not be that sever. Roth and Peranson (1997b) conducted a number of experiments on truncations at the match point to see how they may affect outcomes in the National Resident Matching Program (NRMP). Their experiments were conducted for both the hospital and applicant proposing algorithms for the 1993, 1994, and 1995 matches; see Appendix B in Roth and Peranson (1997b) for detail. They found that "[i]n the majority of cases no change was produced when all ROLs [preferences] were truncated at the match point; and in no case were more than 3 applicants affected by such truncations" (Roth and Peranson (1997b), pp. 26) for both algorithms, among approximately 20,000 jobs filled each year. Their experiments suggest that players in the NRMP market may in fact report preferences (i.e., ROLs, Rank Order Lists) that are equivalent to the truncations at the match point. Our result shows that if, in addition, the reported preferences form a rematching proof or strong equilibrium, then we may confidently conclude that the outcome from the NRMP may indeed be a true stable matching.

Our second focus is why the core is so small in the Nash equilibrium. It is known that the set of stable matchings (the core) in the college admissions problem forms a lattice under the common interests of students due to the common interests of colleges. This implies that the stable matching preferred most by all students is the worst stable matching for all colleges and the stable matching preferred most by all colleges is the worst stable matching for all students. Further, the size of the lattice can grow exponentially as the sizes of the market grow. Since the set of stable matchings is potentially large, which stable matching should be selected? This issue is so important that the pre-existing hospital proposing algorithm implemented since 1951 in the NRMP market is recently replaced by the applicant proposing algorithm: see Roth (1996) and Roth and Peranson (1997a,b).

The empirical studies in Roth and Peranson (1997a,b) found that the set of stable matchings in the NRMP is in fact quite small. For the same set of reported preferences, the switch from the original hospital proposing algorithm to the new applicant proposing algorithm only affected few applicants (approximately 0.1%). This is surprising and it is in the sharp contrast with the theory, since the size of jobs filled in the NRMP market is quite large. Roth and Peranson (1997b) provided two insights why the core may be small. They argued that the small core may be due to the high correlation in preferences (e.g., two new physicians in the same field may have quite similar preferences over positions) and the fact that applicants and hospitals can only conduct a limited number of interviews. Here we provide one more evidence why the core may be "small" when students and colleges report their preferences strategically. We show that any Nash equilibrium in truncations contains one and only one matching (stable or not). Therefore, the core must be small in Nash equilibrium in truncations. Interestingly, there may exist more than one matching at a Nash equilibrium. Nonetheless, any Nash equilibrium can only admit at most one true stable matching. Further, if, indeed, a Nash equilibrium admits a true stable matching, then no unstable matching can be the equilibrium outcome. That is, the true stable matching will always be the outcome, no matter how the equilibrium profile is manipulated (not necessarily in truncations), as long as the equilibrium profile admits a true stable matching. Therefore, a necessary and sufficient condition for the stable matching rule to be implemented in a subset of Nash equilibrium by a direct revelation game induced by a stable matching mechanism is that every equilibrium profile in that subset admits one and only one true stable matching; see Theorem 11.

Our study is largely motivated by the significant empirical works in the labor markets made by Professor Alvin Roth and his co-authors. The first issue is related to the implementation of the stable matching rule with manipulation, a question articulated in Roth (1984b, 1990a) for the marriage problem; also see Roth and Sotomayor (1990) (henceforth, RS). This question has motivated the studies in Alcalde (1996), Kara and Sönmez (1996), Ma (1994,1995), Roth (1984b), Shin and Suh (1996), and Sönmez (1997). These papers study the marriage problem. Kara and Sönmez (1997) studied the college admissions problem and showed that the stable matching rule is essentially monotone. Therefore the stable matching rule is implementable in Nash equilibrium via a result in Danilov (1992). Their game form is quite different from the one studied in this paper. With our game form, the stable matching rule is not Nash implementable.

The remaining of the paper is organized as follows. Section 2 introduces the college admissions problem. Section 3 shows the main result about stable matchings. Section 4 shows the results about the small core. Section 5 concludes the paper.

2 The College Admissions Problem

We use some definitions from RS. The college admissions problem consists of two finite and disjoint sets, $S = \{S_1, \cdots, S_n\}$ of students and $C = \{C_1, \cdots, C_r\}$ of colleges, with each college $C_j \in C$ a quota $q_{C_j} \geq 1$ of enrollments. Each student $S_i \in S$ is enrolled in at most one college. Each student $S_i \in S$ has strict preferences P_{S_i} over colleges $C \cup \{S_i\}$ and each college $C_j \in C$ has strict preferences P_{C_j} over individual students $S \cup \{C_j\}$. Both P_{S_i} and P_{C_j} leave the possibility that a student may prefer not to be enrolled in some colleges and a college may prefer not to enroll some students. Let R_{S_i} and R_{C_j} denote the weak preferences associated with P_{S_i} and P_{C_j} respectively. Let Ω_{S_i} denote the set of all strict preferences for a student $S_i \in S$ and Ω_{C_j} denote the set of all strict preferences over individual students for a college C_j. Let $\Omega = \prod_{S_i \in S} \Omega_{S_i} \times \prod_{C_j \in C} \Omega_{C_j}$ denote the set of all preference profiles. The marriage problem is the college admissions problem with $q_{C_j} = 1$ for every college $C_j \in C$.

Define an unordered family of elements of any set X to be a collection of elements in which the order is immaterial. The set of unordered families of elements of X is denoted by $\underline{X}$.

Definition. A *matching* μ is such a function $\mu : S \cup C \rightarrow \underline{S \cup C}$ such that (a) $|\mu(S_i)| = 1$ for every $S_i \in S$ and $\mu(S_i) \in C$ whenever $\mu(S_i) \neq S_i$; (b) $|\mu(C_j)| = q_{C_j}$ for every $C_j \in C$, and if $| S \cap \mu(C_j) | < q_{C_j}$ then $\mu(C_j)$ is fulfilled to q_{C_j} by copies of C_j; (c) $\mu(S_i) = C_j$ if and only if $S_i \in \mu(C_j)$. Let $\mathcal{M}$ denote the set of all matchings.

Definition. Let $\mu, \lambda \in \mathcal{M}$ be two matchings. We say that a preference $\bar{P}_{C_j}$ for a college C_j over sets of students is *responsive* to its preference P_{C_j} over individual students if, whenever $\mu(C_j) = \lambda(C_j) \cup S_k \setminus \{\tau\}$ for $\tau \in \lambda(C_j)$ and $S_k \notin \lambda(C_j)$, then $\mu(C_j) \bar{P}_{C_j} \lambda(C_j)$ if and only if $S_k P_{C_j} \tau$.

Gale and Shapley (1962) originated the study of the college admissions problem. Roth (1985) reformulated the problem and introduced the responsive preferences. Responsive preferences uniquely determine preferences over individuals but not vice versa. Henceforth, we assume that colleges' preferences over groups of students are responsive, complete, and transitive. We always use $\bar{P}_{C_j}$ with a bar for college C_j's preferences over groups of students that are responsive and P_{C_j} without a bar for college C_j's preferences over individual students. No confusion will be made.

A pair of student S_i and college C_j blocks a matching μ if they are not matched under μ but student S_i prefers college C_j to his mate $\mu(S_i)$ and college C_j prefers student S_i to some member $\sigma \in \mu(C_j)$ in her class $\mu(C_j)$, i.e., $C_j P_{S_i} \mu(S_i)$ and $S_i P_{C_j} \sigma$ for some $\sigma \in \mu(C_j)$.

Definition. Given a profile $P \in \Omega$, a matching μ is (a) *individually rational* if $\mu(S_i) R_{S_i} S_i$ for all $S_i \in S$ and $\sigma R_{C_j} C_j$ for every $\sigma \in \mu(C_j)$ for all $C_j \in C$; (b) *pairwise stable* if it is not blocked by any pairs of student and college; (c) *stable* if it is both individually rational and pairwise stable. Let $\mathcal{S}(P)$ and $IR(P)$ denote the set of all stable matchings and the set of all individually rational matchings respectively with respect to a profile $P \in \Omega$.

Definition. A *matching mechanism* $\psi : \Omega \rightarrow \mathcal{M}$ is a map from profiles to matchings. A matching mechanism $\varphi : \Omega \rightarrow \mathcal{M}$ is *stable* if $\varphi(Q) \in \mathcal{S}(Q)$ for all $Q \in \Omega$. Let Φ denote the set of all stable matching mechanisms.

It follows from Lemma 5.6 in RS that $\mathcal{S}(P)$ is nonempty for any profile $P \in \Omega$. Therefore, the set of stable matching mechanisms Φ is nonempty. A stable matching mechanism $\varphi \in \Phi$ and an underlying true profile $P \in \Omega$ induce a normal form game $\Gamma(\varphi, P)$. The set $\Omega = \prod_{S_i \in S} \Omega_{S_i} \times \prod_{C_j \in C} \Omega_{C_j}$ is the set of strategies of the game $\Gamma(\varphi, P)$. We consider three equilibrium notions in pure strategies of this game.

Definition. A profile $Q \in \Omega$ is a *Nash equilibrium* (in pure strategies) of a game $\Gamma(\varphi, P)$ if

$$\varphi_{S_i}(Q_{-S_i}, Q_{S_i}) R_{S_i} \varphi_{S_i}(Q_{-S_i}, Q'_{S_i})$$

for all $S_i \in S$, $Q'_{S_i} \in \Omega_{S_i}$ and

$$\varphi_{C_j}(Q_{-C_j}, Q_{C_j}) \bar{R}_{C_j} \varphi_{C_j}(Q_{-C_j}, Q'_{C_j})$$

for all $C_j \in C$, $Q'_{C_j} \in \Omega_{C_j}$.

Definition. A *rematching proof equilibrium* Q of a game $\Gamma(\varphi, P)$ is a Nash equilibrium such that

$$\varphi_{S_i}(Q) R_{S_i} \varphi_{S_i}(Q_{-\{S_i,C_j\}}, Q'_{S_i}, Q'_{C_j}) \text{ and } \varphi_{C_j}(Q) \bar{R}_{C_j} \varphi_{C_j}(Q_{-\{S_i,C_j\}}, Q'_{S_i}, Q'_{C_j}),$$

for all $(S_i, C_j) \in S \times C$, all $(Q'_{S_i}, Q'_{C_j}) \in \Omega_{S_i} \times \Omega_{C_j}$.

Definition. A profile Q is a *strong equilibrium* of a game $\Gamma(\varphi, P)$ if it has the property that there exists no coalition $T \subset C \cup S$ and $Q'_T \in \Pi_{k \in T} \Omega_k$ such that

$$\varphi_k(Q_{-T}, Q'_T) P_k \varphi_k(Q) \text{ and } \varphi_l(Q_{-T}, Q'_T) \bar{P}_l \varphi_l(Q)$$

for all $k \in T \cap S$ and all $l \in T \cap C$.

Let $N(\varphi, P)$ ($N^{rp}(\varphi, P)$, $N^s(\varphi, P)$) denote the set of all Nash (rematching proof, strong) equilibria of the game $\Gamma(\varphi, P)$.

3 Stable Matchings

Since the true preferences are private information, a stable matching mechanism has to work with the reported preferences. Roth's impossibility theorems show that there does not exist any stable matching mechanism that makes it a dominant strategy for all students and colleges to report their true preferences. This implies that when students and colleges are confronted with a game $\Gamma(\varphi, P)$ induced by a stable matching mechanism φ, the reported preferences may well be different from the true preferences. Therefore, the stable matching produced by a stable matching mechanism with respect to a misreported preference profile may well be unstable with respect to the underlying true preference profile. Indeed, Roth showed that all individually rational matchings can be supported as Nash equilibrium outcomes for any stable matching mechanism: Given any $\varphi \in \Phi$, $IR(P) \subset \varphi(N(\varphi, P))$ for all $P \in \Omega$; see Theorem 5.18 in RS. In particular, $\mathcal{S}(P) \subset \varphi(N(\varphi, P))$ for every $P \in \Omega$. But many Nash equilibrium outcomes are not stable with respect to the true profile P.

In the marriage problem, Ma (1995) showed that all rematching proof equilibrium outcomes are stable: Given any $\varphi \in \Phi$, $\varphi(N^{rp}(\varphi, P)) \subset \mathcal{S}(P)$ for all $P \in \Omega$. Ma (1994) and Shin and Suh (1996) studied the strong equilibrium (due to Aumann) and showed that all strong equilibrium outcomes are stable: Given any $\varphi \in \Phi$, $\varphi(N^s(\varphi, P)) \subset \mathcal{S}(P)$ for all $P \in \Omega$. Sönmez (1997) introduced the $\mathcal{G}$-core and the $\mathcal{G}$-proof Nash equilibrium and generalized these results for the marriage problem. The

$\mathcal{G}$-core and $\mathcal{G}$-proof Nash equilibrium in Sönmez (1997) are defined for coalitions. Unfortunately, these results do not apply to the college admissions problem; as shown in the following example.

Example 1. Let $C = \{C_1, C_2\}$, $S = \{S_1, S_2\}$, $q_{C_1} = 2, q_{C_2} = 1$. Let

$$P_{C_1} = (S_1, S_2, C_1) \qquad P_{C_2} = (S_2, S_1, C_2)$$

$$P_{S_1} = (C_2, C_1, S_1) \qquad P_{S_2} = (C_1, C_2, S_2)$$

and let φ be any stable matching mechanism. Let

$$Q_{C_1} = (S_1, C_1, S_2).$$

Then

$$\varphi(P_{-C_1}, Q_{C_1}) = [(C_1; S_1, C_1), (C_2; S_2)].$$

We now show that (P_{-C_1}, Q_{C_1}) is a strong equilibrium. To prove this, it is sufficient to consider the coalition $\{C_1, S_1, S_2\}$ since C_2 cannot be better. To make C_1 better off, C_1 should be matched with S_1 and S_2. But this does not make student S_1 better off. To make S_1 better off, S_1 must be matched with C_2. But then it is impossible to make C_1 better off. Similarly, one can check the coalitions such as $\{C_1, S_2\}$ and $\{S_1, S_2\}$. This shows that (P_{-C_1}, Q_{C_1}) is a Nash, rematching proof, and strong equilibrium for the true profile P but $\varphi(P_{-C_1}, Q_{C_1})$ is not stable for P. □

This example shows that the college admissions problem is different from the marriage problem. Some additional conditions are needed for both rematching proof and strong equilibrium to generate stable matchings.

We follow Roth and Vande Vate (1991) and Roth and Rothblum (1996) to introduce a class of simple strategies called truncations. This class of strategies are introduced for the marriage problem in these papers. A college C_j is acceptable to a student S_i if $C_j R_{S_i} S_i$. A student S_i is acceptable for a college C_j if $S_i R_{C_j} C_j$. A truncation strategy Q_{S_i} (with respect to P_{S_i}) for a student S_i contains k $(0 \leq k)$ acceptable colleges such that the first k elements of Q_{S_i} are the first k elements, with the same order, in her true preference P_{S_i}, and the $(k+1)$th element in Q_{S_i} is S_i. The ordering after the $(k+1)$th element S_i in Q_{S_i} does not matter for our studies. Similarly, a truncation strategy Q_{C_j} for a college C_j (with respect to P_{C_j}) contains k $(0 \leq k)$ acceptable students such that the first k elements of Q_{C_j} are the first k elements, with the same order, in her true preference P_{C_j}, and the $(k+1)$th element in Q_{C_j} is C_j. Again the ranking order after C_j does not matter. For example, the equilibrium strategy Q_{C_1} in Example 1 is a truncation strategy for college C_1.

Truncation strategies exclude some other more complicated strategies such as the changes in orders that may be profitable for manipulation. Roth and Rothblum (1996) showed that more complicated and profitable strategies other than truncations do exist. But they also show that players need to know all preferences of the other players in order to exploit the benefits of such complicated strategies in manipulation. They convincingly showed why this class of simple strategies of truncations are plausible for the marriage problem in an environment with low information about the preferences of all other players. In the Sorority rush market Mongell and Roth (1989) found that truncation strategies are in fact used by players in practice: Players often truncate after their first choice.

A Nash equilibrium Q of a game $\Gamma(\varphi, P)$ is a Nash equilibrium in truncations if all equilibrium strategies are truncations. The equilibrium profile (P_{-C_1}, Q_{C_1}) in Example 1 is a Nash, rematching proof, and strong equilibrium in truncations. Example 1 shows that a Nash, rematching proof, and strong equilibrium outcome in truncations may be unstable. Therefore, some additional conditions are required to refine this class of equilibrium strategies. We introduce a class of strategies in truncations at the match point; see Roth and Peranson (1997b). We say that Q is a Nash equilibrium in truncations at the match point if the kth element in Q_{S_i} for student S_i is the college $\varphi_{S_i}(Q)$ or Q_{S_i} is any truncation strategy in the case that $\varphi_{S_i}(Q) = S_i$, and the kth element in Q_{C_j} for college C_j is the student that is the least preferred by the college C_j among all students in $\varphi_{C_j}(Q)$ or Q_{C_j} is any truncation strategy in the case that $\varphi_{C_j}(Q)$ does not contain any student. An equilibrium is an equilibrium in truncations at the match point if all equilibrium strategies are truncations at the match point. For example, the equilibrium strategies Q_{C_1}, P_{S_1}, and P_{S_2} are all truncation strategies at the match point in Example 1; The equilibrium strategy P_{C_2} is a truncation strategy but not a truncation strategy at the match point; Thus, the equilibrium profile (P_{-C_1}, Q_{C_1}) is not an equilibrium in truncations at the match point.

One may wonder whether truncations at the match point in Nash equilibrium are strong enough to generate stable matchings. The following example shows that the answer is negative.

Example 2 (Roth and Vande Vate (1991)). Let $S = (S_1, S_2)$, $C = (C_1, C_2)$, $q_{C_1} = q_{C_2} = 1$. Let, $i, j = 1, 2$, the true profile P is as follows:

$$P_{S_i} = (C_1,\ C_2,\ S_i)$$

$$P_{C_j} = (S_1,\ S_2,\ C_j).$$

Now consider the following Nash equilibrium Q:

$$Q_{S_i} = (C_1,\ S_i,\ C_2)$$

$$Q_{C_j} = (S_1,\ C_j,\ S_2).$$

Then the matching

$$\mu = [(S_1; C_1), (S_2; S_2), (C_2; C_2)]$$

is the unique stable matching for Q. Thus $\varphi(Q) = \mu$ for any stable matching mechanism φ. The profile Q is a Nash equilibrium in truncations at the match point. But μ is not stable for P. □

Proposition 3 below shows that all rematching proof equilibrium outcomes in truncations at the match point are stable. This is in the contrast to Examples 1 and 2 above. Since a strong equilibrium is also rematching proof, it follows that all strong equilibrium outcomes in truncations at the match point are stable as well.

Proposition 3. *Let $(\varphi, P) \in \Phi \times \Omega$ and Q be a rematching proof equilibrium in truncations at the match point. Then $\varphi(Q)$ is a stable matching in $\mathcal{S}(P)$.*

Proof. Let $(\varphi, P) \in \Phi \times \Omega$. Let Q be a rematching proof equilibrium in truncations at the match point. Suppose on the contrary that $\varphi(Q) \notin \mathcal{S}(P)$. Then $\exists (S_i, C_j) \in S \times C$ such that $C_j P_{S_i} \varphi_{S_i}(Q)$ and $S_i P_{C_j} \sigma$ for some $\sigma \in \varphi_{C_j}(Q)$. (Note that $\varphi(Q) \in IR(P)$ since Q is a truncation profile and $\varphi(Q) \in IR(Q)$). Let $\{S_{i_1}, \cdots, S_{i_q}\} = \varphi_{C_j}(Q) \cap S$ such that $S_{i_1} Q_{C_j} S_{i_2} Q_{C_j} \cdots Q_{C_j} S_{i_q}$. Note that $q \leq q_{C_j}$.

We discuss four cases.

Case a. $\varphi_{S_i}(Q) \neq S_i$ and $C_j \notin \varphi_{C_j}(Q)$.

Because $C_j P_{S_i} \varphi_{S_i}(Q)$ and Q_{S_i} is a truncation of P_{S_i} up to $\varphi_{S_i}(Q)$, it follows that $C_j Q_{S_i} \varphi_{S_i}(Q)$. Because $S_i P_{C_j} \sigma$ for some student $\sigma \in \varphi_{C_j}(Q)$ and Q_{C_j} is a truncation of P_{C_j} up to S_{i_q}, it follows that $S_i Q_{C_j} S_{i_q}$. This shows that (S_i, C_j) blocks $\varphi(Q)$ with respect to Q contradicting $\varphi(Q) \in \mathcal{S}(Q)$.

Case b. $\varphi_{S_i}(Q) = S_i$ and $C_j \notin \varphi_{C_j}(Q)$.

Since Q_{C_j} is a truncation of P_{C_j}, the assumption that $S_i P_{C_j} \sigma$ for some student $\sigma \in \varphi_{C_j}(Q)$ implies that $S_i Q_{C_j} S_{i_q} Q_{C_j} C_j$. Since $C_j P_{S_i} \varphi_{S_i}(Q)$ and Q_{S_i} is a truncation of P_{S_i}, it follows that $S_i Q_{S_i} C_j$. (Otherwise, $C_j Q_{S_i} S_i$ implies that (S_i, C_j) blocks $\varphi(Q)$ with respect to Q contradicting $\varphi(Q) \in \mathcal{S}(Q)$). Now let $Q'_{S_i} = (\cdots, C_j, S_i, \cdots)$ be a truncation of P_{S_i}, up to C_j. Then we show that

$$\varphi_{S_i}(Q_{-S_i}, Q'_{S_i}) \neq S_i.$$

Denote $Q' = (Q_{-S_i}, Q'_{S_i})$. Suppose on the contrary that $\varphi_{S_i}(Q') = S_i$. Then $\sigma Q_{C_j} S_i$ for all $\sigma \in \varphi_{C_j}(Q')$. Otherwise, $S_i Q_{C_j} \sigma$ for some $\sigma \in \varphi_{C_j}(Q')$ implies that (S_i, C_j) blocks $\varphi(Q')$ with respect to Q' contradicting $\varphi(Q') \in \mathcal{S}(Q')$. Therefore, $\varphi_{C_j}(Q') \subset S$. Thus there exists a student $\tau \in \varphi_{C_j}(Q')$ such that $\tau \notin \varphi_{C_j}(Q)$ and $\tau Q_{C_j} S_{i_q}$, since $\varphi_{C_j}(Q) \subset S$ and S_i is not in $\varphi_{C_j}(Q')$ but $S_i Q_{C_j} S_{i_q}$. This implies that $C_j Q_\tau \varphi_\tau(Q)$ since Q_τ is a truncation up to $\varphi_\tau(Q)$. Because $\tau Q_{C_j} S_{i_q}$, $C_j Q_\tau \varphi_\tau(Q)$ implies that (τ, C_j) blocks $\varphi(Q)$ with respect to Q contradicting $\varphi(Q) \in \mathcal{S}(Q)$. This shows that $\varphi_{S_i}(Q') \neq S_i$.

It follows that either $\varphi_{S_i}(Q') = C_j$ or $\varphi_{S_i}(Q') Q_{S_i} C_j$. Either case implies that

$$\varphi_{S_i}(Q') P_{S_i} \varphi_{S_i}(Q)$$

contradicting Q is a Nash equilibrium.

Case c. $\varphi_{S_i}(Q) = S_i$ and $C_j \in \varphi_{C_j}(Q)$.

Let $Q'_{S_i} = (C_j, S_i, \cdots)$ and $Q'_{C_j} = (S_{i_1}, S_{i_2}, \cdots, S_{i_q}, S_i, C_j, \cdots)$. Denote $\tilde{Q} = (Q_{-\{S_i, C_j\}}, Q'_{S_i}, Q'_{C_j})$. We show that

$$\varphi_{S_i}(\tilde{Q}) = C_j \text{ and } \varphi_{C_j}(\tilde{Q}) = \varphi_{C_j}(Q) \cup \{S_i\} \setminus \{C_j\}.$$

Note that $q < q_{C_j}$. If $\varphi_{S_i}(\tilde{Q}) = S_i$, then $\varphi_{C_j}(\tilde{Q}) \subset S$. (Otherwise, $C_j \in \varphi_{C_j}(\tilde{Q})$ implies that (S_i, C_j) blocks $\varphi(\tilde{Q})$ with respect to $\tilde{Q}$.) But this is impossible since $\varphi_{C_j}(\tilde{Q}) \subset \{S_{i_1}, \cdots, S_{i_q}\}$ and $q < q_{C_j}$. Therefore $\varphi_{S_i}(\tilde{Q}) = C_j$.

Let

$$\begin{aligned} \lambda(S_i) &= C_j \\ \lambda(C_j) &= \varphi_{C_j}(Q) \cup \{S_i\} \setminus \{C_j\} \end{aligned}$$

$$\lambda(S_{i'}) = \varphi_{S_{i'}}(Q), \forall S_{i'} \in S \setminus \{S_i\}$$

$$\lambda(C_{j'}) = \varphi_{C_{j'}}(Q), \forall C_{j'} \in C \setminus \{C_i\}$$

By the construction of $\tilde{Q}$ and λ, $\lambda \in S(\tilde{Q})$ since any blocking (S_k, C_l) of λ with respect to $\tilde{Q}$ is also a blocking pair of $\varphi(Q)$ with respect to Q. We need to show that $\varphi_{C_j}(\tilde{Q}) = \lambda(C_j)$. To show this, we show that $\mu(C_j) = \lambda(C_j)$ for every $\mu \in S(\tilde{Q})$. We consider two situations.

(a). $C_j \in \lambda(C_j)$. Theorem 5.13 in RS shows that any college that does not fill its quota at some stable matching is assigned the same set of students at every stable matching. Therefore it follows from Theorem 5.13 in RS that $\mu(C_j) = \lambda(C_j)$ for every $\mu \in S(\tilde{Q})$.

(b). $C_j \notin \lambda(C_j)$. This implies that $\lambda(C_j) = \{S_{i_1}, \cdots, S_{i_q}, S_i\}$ and $q = q_{C_j} - 1$. Suppose that there exists some $\mu \in S(\tilde{Q})$ such that $\mu(C_j) \neq \lambda(C_j)$. Then $C_j \in \mu(C_j)$. But Theorem 5.13 again shows that $\mu(C_j) = \lambda(C_j)$, a contradiction. Therefore both cases show that $\mu(C_j) = \lambda(C_j)$ for every $\mu \in S(\tilde{Q})$. Thus

$$\varphi_{C_j}(\tilde{Q}) = \lambda(C_j).$$

Now, since $\bar{P}_{C_j}$ is responsive and $S_i P_{C_j} C_j$, it follows that

$$\varphi_{S_i}(\tilde{Q}) P_{S_i} \varphi_{S_i}(Q)$$

$$\varphi_{C_j}(\tilde{Q}) \bar{P}_{C_j} \varphi_{C_j}(Q)$$

contradicting Q is rematching proof. Complete the proof of case (c).

Case d. $\varphi_{S_i}(Q) \neq S_i$ and $C_j \in \varphi_{C_j}(Q)$.

Let $Q'_{S_i} = (C_j, S_i, \cdots)$ and $Q'_{C_j} = (S_i, S_{i_1}, S_{i_2}, \cdots, S_{i_q}, C_j, \cdots)$. Denote $\tilde{Q} = (Q_{-\{S_i, C_j\}}, Q'_{S_i}, Q'_{C_j})$. We show that

$$\varphi_{S_i}(\tilde{Q}) = C_j \text{ and } \varphi_{C_j}(\tilde{Q}) = \varphi_{C_j}(Q) \cup \{S_i\} \setminus \{C_j\}.$$

First, suppose that $\varphi_{S_i}(\tilde{Q}) = S_i$. Then there exists $C_j \in \varphi_{C_j}(\tilde{Q})$, by the construction of Q'_{C_j}. But then this implies that $\varphi(\tilde{Q}) \notin S(\tilde{Q})$ since (S_i, C_j) blocks $\varphi(\tilde{Q})$ with respect to $\tilde{Q}$. Therefore $\varphi_{S_i}(\tilde{Q}) = C_j$.

Second, suppose that $\varphi_{C_j}(\tilde{Q}) \neq \varphi_{C_j}(Q) \cup \{S_i\} \setminus \{C_j\}$. Then there exist at least one $C_j \in \varphi_{C_j}(\tilde{Q})$ and at least one student $S_k \in \varphi_{C_j}(Q)$ such that $S_k \notin \varphi_{C_j}(\tilde{Q})$. Since $S_k \neq S_i$ and $S_k \in \varphi_{C_j}(Q)$, it follows that $C_j Q_{S_k} S_k$. Since $C_j \in \varphi_{C_j}(\tilde{Q})$, $\varphi(\tilde{Q}) \in S(\tilde{Q})$ and the fact that Q_{S_k} is a truncation of P_{S_k} up to $\varphi_{S_k}(Q) = C_j$, it follows that

$$\varphi_{S_k}(\tilde{Q}) Q_{S_k} C_j Q_{S_k} S_k.$$

Let $C_l = \varphi_{S_k}(\tilde{Q})$ and then $C_l Q_{S_k} C_j = \varphi_{S_k}(Q)$. Note that $C_l \neq C_j$ since $S_k \notin \varphi_{C_j}(\tilde{Q})$. It follows that $S_k Q_{C_l} C_l$ by individual rationality. Since $S_k \notin \varphi_{C_l}(Q)$ and $\varphi(Q) \in \mathcal{S}(Q)$, it follows that $\varphi_{C_l}(Q) \subset S$. (Otherwise, $C_l \in \varphi_{C_l}(Q)$ and $S_k Q_{C_l} C_l$. This implies that (S_k, C_l) blocks $\varphi(Q)$ with respect to Q contradicting $\varphi(Q) \in \mathcal{S}(Q)$.). Because $\varphi(Q) \in \mathcal{S}(Q)$, it follows that $\sigma Q_{C_l} S_k$ for every student $\sigma \in \varphi_{C_l}(Q)$. Therefore, $\sigma Q_{C_l} S_k Q_{C_l} C_l$ for every student $\sigma \in \varphi_{C_l}(Q)$. But Q_{C_l} is a truncation of P_{C_l} at the match point and there is no such a student S_k between the student σ and C_l, a contradiction.

Therefore, this shows that

$$\varphi_{S_i}(Q') = C_j \text{ and } \varphi_{C_j}(Q') = \varphi_{C_j}(Q) \cup \{S_i\} \setminus \{C_j\}.$$

Since $\bar{P}_{C_j}$ is responsive and $S_i P_{C_j} C_j$, we have that

$$\varphi_{S_i}(\tilde{Q}) P_{S_i} \varphi_{S_i}(Q)$$

$$\varphi_{C_j}(\tilde{Q}) \bar{P}_{C_j} \varphi_{C_j}(Q)$$

contradicting Q is rematching proof. This completes the proof of *Case d.* □

The next question is whether a rematching proof or strong equilibrium in truncations at the match point exists. The answer is yes. In fact our next result shows that every stable matching can be achieved in at least one strong equilibrium in truncations at the match point.

Let $N^{trp}(\varphi, P)$ and $N^{ts}(\varphi, P)$ denote the set of all rematching proof and strong equilibria in truncations at the match point respectively of the game $\Gamma(\varphi, P)$. As a corollary of Proposition 4, it follows that $\mathcal{S}(P) \subset \varphi(N^{trp}(\varphi, P))$.

Proposition 4. *For all $(\varphi, P) \in \Phi \times \Omega$, $\mathcal{S}(P) \subset \varphi(N^{ts}(\varphi, P))$.*

Proof. Let $\mu \in \mathcal{S}(P)$. Denote

$$\{S_{i_1}, \cdots, S_{i_{q_j}}\} = \mu(C_j) \cap S$$

such that $S_{i_1} P_{C_j} S_{i_2} P_{C_j} \cdots P_{C_j} S_{i_{q_j}}$.

Let Q be a truncation profile at the match point such that

$$Q_{S_i} = \begin{cases} (\overbrace{\cdots\cdots\cdots\cdots, \mu(S_i),}^{\text{truncation of } P_{S_i} \text{ up to } \mu(S_i)} \quad S_i, \cdots) & \text{if } \mu(S_i) \neq S_i \\ P_{S_i} & \text{otherwise} \end{cases}$$

$$Q_{C_j} = \begin{cases} (\overbrace{\cdots\cdots\cdots\cdots, S_{i_{q_j}},}^{\text{truncation of } P_{C_j} \text{ up to } S_{i_{q_j}}} \quad C_j, \cdots) & \text{if } \mu(C_j) \cap S \neq \emptyset \\ P_{C_j} & \text{otherwise} \end{cases}$$

Then $\varphi(Q) = \mu$ for all $\varphi \in \Phi$ since $\mathcal{S}(Q) = \{\mu\}$. We claim that the truncation profile Q at the match point is a strong equilibrium in $N^{ts}(\varphi, P)$ for all $\varphi \in \Phi$.

Suppose on the contrary that there exists some stable matching mechanism φ such that there exist a coalition $T \subset C \cup S$ and strategies $Q'_T \in \Pi_{k \in T}\Omega_k$ such that

$$\varphi_{S_i}(\tilde{Q}) \quad P_{S_i} \quad \varphi_{S_i}(Q), \forall S_i \in T \cap S \tag{1}$$

$$\varphi_{C_j}(\tilde{Q}) \quad \bar{P}_{C_j} \quad \varphi_{C_j}(Q), \forall C_j \in C \cap T \tag{2}$$

where $\tilde{Q} = (Q_{-T}, Q'_T)$.

Denote

$$T_1 = \{S_k \in S : \varphi_{S_k}(\tilde{Q})\, P_{S_k}\, \varphi_{S_k}(Q)\}$$

the set of all students who prefer $\varphi(\tilde{Q})$ to $\varphi(Q)$ and

$$T_2 = \{C_l \in C : \varphi_{C_l}(\tilde{Q})\, \bar{P}_{C_l} \varphi_{C_l}(Q)\}$$

the set of all colleges that prefer $\varphi(\tilde{Q})$ to $\varphi(Q)$. Clearly $T \subset T_1 \cup T_2$. We consider two cases below.

(a). $T_2 \neq \emptyset$. For any college C_j in T_2, by the responsiveness of $\bar{P}_{C_j}$ and (2), there exist a student $\sigma \in \varphi_{C_j}(\tilde{Q})$, $\sigma \notin \varphi_{C_j}(Q)$, and a member $\tau \in \varphi_{C_j}(Q)$ such that $\sigma P_{C_j} \tau$. If σ is in T_1, then $C_j P_\sigma \varphi_\sigma(Q)$ and then (σ, C_j) blocks μ with respect to P contradicting $\mu \in \mathcal{S}(P)$. Thus $\sigma \notin T_1$. Note that $\sigma \neq C_j$ since σ is a student. It follows that $\varphi_\sigma(Q) P_\sigma C_j$.

We obtain that

$$\varphi_\sigma(Q) P_\sigma C_j P_\sigma \sigma.$$

But this is impossible because there does not exist C_j between $\varphi_\sigma(Q)$ and σ by the construction of Q_σ.

(b). $T_2 = \emptyset$. Then $T_1 \neq \emptyset$ since $T \subset T_1$. Let $S_i \in T_1$ and let $C_l = \varphi_{S_i}(\tilde{Q})$. Clearly, $S_i \notin \varphi_{C_l}(Q)$ since $C_l P_{S_i} \varphi_{S_i}(Q)$. The assumption that $T_2 = \emptyset$ implies that $T \subset S$. By the construction of Q_{C_l}, it follows that

$$\begin{cases} S_i Q_{C_l} S_{i_{q_l}} & \text{if } \mu(C_l) \cap S \neq \emptyset \\ S_i Q_{C_l} C_l & \text{otherwise.} \end{cases}$$

Since Q_{C_l} is a truncation of P_{C_l}, we also have that

$$\begin{cases} S_i P_{C_l} S_{i_{q_l}} & \text{if } \mu(C_l) \cap S \neq \emptyset \\ S_i P_{C_l} C_l & \text{otherwise.} \end{cases}$$

Either case implies that (S_i, C_l) blocks μ with respect to P contradicting $\mu \in \mathcal{S}(P)$. □

We may summarize the results in Propositions 3 and 4 as follows.

Theorem 5 (Stable Matchings). *The stable matching correspondence $\mathcal{S} : \Omega \rightarrow \mathcal{M}$ can be implemented by any direct revelation mechanism $(\Pi_{S_i}\Omega_{S_i} \times \Pi_{C_j}\Omega_{C_j}, \varphi)$, where the outcome function φ is a stable matching mechanism in Φ, in both rematching proof and strong equilibrium in truncation strategies at the match point.*

Proof. Proposition 3 shows that $\varphi(N^{trp}(\varphi, P)) \subset \mathcal{S}(P)$. Proposition 4 shows that $\mathcal{S}(P) \subset \varphi(N^{ts}(\varphi, P))$. The proof is complete since a strong equilibrium is also rematching proof. □

Remark. Theorem 5 does show that there are some meaningful refinements of the Nash equilibrium that generate stable matchings even though the equilibrium profiles are manipulated. This provides a support for the hypothesis that a stable matching mechanism that generates a stable matching with respect to the report preference profiles may in fact generate stable matchings with respect to the true preferences. This in turn provides some important economic insights to the empirical evidences discovered by Roth (1984a, 1990b, 1991) in a large number of labor markets to some degree. But our result also reveals the difficulty in order to obtain true stable matchings from the misreported preferences. Both truncations at the match point and the rematching proof or strong equilibrium are needed to generate true stable matchings. The class of truncation strategies are simple and plausible. But the truncations at the match point may be objected on the ground that players need information to know the match point before reporting such strategies. Therefore, it is quite surprising to know from the experiments conducted in Roth and Peranson (1997b) for the NRMP markets in the 1993, 1994, and 1995 matches that the actual outcome of the hospital and applicant proposing algorithms is in fact equivalent to the outcome as if all reported strategies are truncations at the match point. Thus, even if the actual reportings may not be truncations at the match point, the reportings are equivalent to truncations at the match point in outcomes. Therefore, if the reporting strategies also form a rematching proof equilibrium, then it follows from Theorem 5 that the outcome can be a true stable matching. Hence, our results in Theorem 5 provide important insights to the empirical evidences.

4 The Small Core in Nash Equilibrium

In theory, the number of stable matchings in $\mathcal{S}(P)$ may grow exponentially as the sizes of the market grow; see Theorem 3.18 in RS. In contrast, Roth and Peranson (1997a,b) found that the set of stable matchings (the core) in the NRMP market is in fact quite small. For the same set of reported ranking order lists, the pre-existing NRMP algorithm and the new applicant proposing algorithm generate almost the same outcome. Quite few applicants (approximately 0.1%) are affected after the switch from the pre-existing NRMP algorithm to the applicant proposing algorithm. This calls up a question why the core is so small.

Let $\varphi(Q)$ be a stable matching with respect to a Nash equilibrium profile Q of the game $\Gamma(\varphi, P)$. Lemma 6 shows that no other stable matching μ in $\mathcal{S}(Q)$ exists such that some college C_j prefers $\mu(C_j)$ to $\varphi_{C_j}(Q)$ or some student S_i prefers $\mu(S_i)$ to $\varphi_{S_i}(Q)$.

Lemma 6. *Suppose $Q \in N(\varphi, P)$ is a Nash equilibrium, where $\varphi \in \Phi$. Then for any matching $\mu \in \mathcal{S}(Q)$,*

$$\varphi_{S_i}(Q)\, R_{S_i}\, \mu(S_i)$$

for every $S_i \in S$ and

$$\varphi_{C_j}(Q) \bar{R}_{C_j} \mu(C_j)$$

for every $C_j \in C$.

Proof. First, suppose on the contrary that $\mu(S_i)\, P_{S_i}\, \varphi_{S_i}(Q)$ for some $S_i \in S$. Then let $Q'_{S_i} = (\mu(S_i), S_i, \cdots)$ and note that $\mu \in \mathcal{S}(Q_{-S_i}, Q'_{S_i})$. Theorem 5.12 in RS shows that the set of students employed is the same at every stable matching. Therefore, it follows from Theorem 5.12 that

$$\mu(S_i) = \lambda(S_i), \forall \lambda \in \mathcal{S}(Q_{-S_i}, Q'_{S_i}).$$

Thus

$$\varphi_{S_i}(Q_{-S_i}, Q'_{S_i}) = \mu(S_i),$$

i.e.,

$$\varphi_{S_i}(Q_{-S_i}, Q'_{S_i})\, P_{S_i}\, \varphi_{S_i}(Q)$$

contradicting Q is a Nash equilibrium.

Now suppose on the contrary that there exists some $C_j \in C$ such that

$$\mu(C_j)\, \bar{P}_{C_j}\, \varphi_{C_j}(Q).$$

Then let $\{S_{i_1}, \cdots, S_{i_q}\} = S \cap \mu(C_j)$ such that $S_{i_1} Q_{C_j} S_{i_2} Q_{C_j} \cdots Q_{C_j} S_{i_q}$. Note that $q \leq q_{C_j}$. Let

$$Q'_{C_j} = (S_{i_1}, S_{i_2}, \cdots, S_{i_q}, C_j, \cdots).$$

We show that $\mu \in S(Q_{-C_j}, Q'_{C_j})$. Suppose this is not true. Then $\exists(\tilde{S}_i, \tilde{C}_j) \in S \times C$, with $\mu(\tilde{S}_i) \neq \tilde{C}_j$, such that $\tilde{C}_j Q_{\tilde{S}_i} \mu(\tilde{S}_i)$ and $\tilde{S}_i Q_{\tilde{C}_j} \sigma$ for some $\sigma \in \mu(\tilde{C}_j)$. Since $\tilde{C}_j \neq C_j$ and $\tilde{S}_i \notin \{S_{i_1}, \cdots, S_{i_q}\}$ by the construction of Q'_{C_j}, this implies that $(\tilde{S}_i, \tilde{C}_j)$ blocks μ with respect to Q contradicting $\mu \in S(Q)$.

We want to show that $\varphi_{C_i}(Q_{-C_j}, Q'_{C_j}) = \mu(C_j)$, which implies that

$$\varphi_{C_j}(Q_{-C_j}, Q'_{C_j})\, \bar{P}_{C_j}\, \varphi_{C_j}(Q)$$

contradicting Q is a Nash equilibrium.

We now show that $\lambda(C_j) = \mu(C_j)$ for all $\lambda \in S(Q_{-C_j}, Q'_{C_j})$.

(a) $C_j \in \mu(C_j)$. Now Theorem 5.13 in RS shows that any college that does not fill its quota at some stable matching is assigned precisely the same set of students at every stable matching. Therefore, it follows from Theorem 5.13 that for all $\lambda \in S(Q_{-C_j}, Q'_{C_j})$

$$\lambda(C_j) = \mu(C_j)$$

since $\mu \in S(Q_{-C_j}, Q'_{C_j})$ and $C_j \in \mu(C_j)$.

(b) $C_j \notin \mu(C_j)$. Now suppose $\mu(C_j) \neq \lambda(C_j)$ for some $\lambda \in S(Q_{-C_j}, Q'_{C_j})$. Then there must exist some $C_j \in \lambda(C_j)$ by the construction of Q'_{C_j}. But this implies, by Theorem 5.13 in RS again, that $\mu(C_j) = \lambda(C_j)$, a contradiction to the assumption. Therefore $\lambda(C_j) = \mu(C_j)$ for all $\lambda \in S(Q_{-C_j}, Q'_{C_j})$. This completes the proof. □

Is it possible for a Nash equilibrium to admit more than one matching? The following example provides a negative answer to this.

Example 7. Let $S = (S_1, S_2)$, $C = (C_1, C_2)$, $q_{C_1} = q_{C_2} = 1$. Suppose P is as follows:

$$P_{S_1} = (C_1,\ S_1, C_2) \qquad P_{S_2} = (C_2, S_2, C_1)$$
$$P_{C_1} = (S_1,\ C_1,\ S_2) \qquad P_{C_2} = (S_2, C_2, S_1)$$

The following Q is a Nash equilibrium of the game $\Gamma(\varphi, P)$, where φ is the student proposing algorithm.

$$Q_{S_1} = (C_1,\ C_2, S_1) \qquad Q_{S_2} = (C_2, C_1, S_2)$$
$$Q_{C_1} = (S_2,\ S_1,\ C_1) \qquad P_{C_2} = (S_1, S_2, C_2)$$

But $\mid \mathcal{S}(Q) \mid = 2$. □

Thus, the core at a Nash equilibrium Q does not always contain a unique matching. Thus, the core may not be small at a Nash equilibrium. But Theorem 8 shows that every Nash equilibrium in truncations contains one and only one matching, stable or not. Therefore, the core at a Nash equilibrium in truncations (not necessarily truncations at the match point) must be small. It contains a unique outcome. It should be noted that a Nash equilibrium in truncations may not be stable; see Example 2.

Theorem 8 (Small Core). Let $(\varphi, P) \in \Phi \times \Omega$ and $Q \in N(\varphi, P)$ be a Nash equilibrium in truncations. Then $\mathcal{S}(Q) = \{\varphi(Q)\}$.

Proof. Suppose there exists $\mu \in \mathcal{S}(Q)$ such that $\mu \neq \varphi(Q)$. Then, by Lemma 6, $\varphi_{S_i}(Q) P_{S_i} \mu(S_i)$ for all $S_i \in S$ with $\varphi_{S_i}(Q) \neq \mu(S_i)$. Since $Q \in \Omega$ is a truncation of P, we also obtain that $\varphi_{S_i}(Q) Q_{S_i} \mu(S_i)$ for all $S_i \in S$ with $\varphi_{S_i}(Q) \neq \mu(S_i)$ (all such students are matched with colleges).

Theorem 5.33 in RS shows that $\mathcal{S}(Q)$ forms a lattiee under the common preferences of colleges, $\bar{Q}_C$, and dual to the common preferences of students, Q_S. Therefore we have

$$\mu(C_j) \bar{Q}_{C_j} \varphi_{C_j}(Q), \tag{3}$$

for all C_j such that $\mu(C_j) \neq \varphi_{C_j}(Q)$, since students prefer $\varphi(Q)$ to μ in Q_S.

Theorem 5.13 in RS shows that any college that does not fill its quota at some stable matching is assigned the same set of students at every stable matching. For every C_j such that $\mu(C_j) \neq \varphi_{C_j}(Q)$, it follows from Theorem 5.13 that

$$\mu(C_j) \subset S \text{ and } \varphi_{C_j}(Q) \subset S. \tag{4}$$

It follows from (3), (4), and responsive preferences that college C_j prefers at least one student in $\mu(C_j)$ to a student in $\varphi_{C_j}(Q)$. Then Lemma 5.25 in RS shows that college C_j weakly prefers all students in $\mu(C_j)$ to all students in $\varphi_{C_j}(Q)$ in the related marriage problem. It follows from Lemma 5.25 in RS that

$$\mu(C_j)\bar{P}_{C_j}\varphi_{C_j}(Q)$$

for all C_j such that $\mu(C_j) \neq \varphi_{C_j}(Q)$, since Q_{C_j} is a truncation of P_{C_j}. This is a contradiction to Lemma 6. □

Given a stable matching mechanism $\varphi \in \Phi$ and a profile $P \in \Omega$, the matching $\varphi(Q)$ is stable with respect to the profile $Q \in \Omega$. But the matching $\varphi(Q)$ may or may not be stable for P. Of course, if the matching $\varphi(Q)$ is stable with respect to the profile P, then $\mathcal{S}(Q)$ contains at least one stable matching in $\mathcal{S}(P)$. But the set $\mathcal{S}(Q)$ may contain more than one element in $\mathcal{S}(P)$. The next result shows that if $\varphi(Q)$ is stable for P, then there exists no other matching in $\mathcal{S}(Q)$ that is also stable for P.

Lemma 9. *For every Nash equilibrium $Q \in N(\varphi, P)$ such that $\varphi(Q) \in \mathcal{S}(P)$, $\mathcal{S}(Q) \cap \mathcal{S}(P) = \{\varphi(Q)\}$.*

Proof. Suppose that $|\ \mathcal{S}(Q) \cap \mathcal{S}(P)\ | > 1$. Thus there exists $\mu \in \mathcal{S}(Q) \cap \mathcal{S}(P)$ such that $\varphi_k(Q) \neq \mu(k)$ for some $k \in S \cup C$. Theorem 5.26 in RS shows that a college C_j is indifferent (over groups of students) between μ and $\varphi(Q)$ only if $\mu(C_j) = \varphi_{C_j}(Q)$. It follows that either

$$\mu(C_j)\bar{P}_{C_j}\varphi_{C_j}(Q)$$

or

$$\varphi_{C_j}(Q)\bar{P}_{C_j}\mu(C_j)$$

for all $C_j \in C$ with $\varphi_{C_j}(Q) \neq \mu(C_j)$. Then Lemma 6 shows that

$$\varphi_{S_i}(Q)\ P_{S_i}\ \mu(S_i)$$

for all $S_i \in S$ with $\varphi_{S_i}(Q) \neq \mu(S_i)$ and

$$\varphi_{C_j}(Q)\bar{P}_{C_j}\mu(C_j)$$

for all $C_j \in C$ with $\varphi_{C_j}(Q) \neq \mu(C_j)$. But this contradicts Theorem 5.33 in RS which shows that if all students prefer $\varphi(Q)$ to μ, then all colleges prefer μ to $\varphi(Q)$. □

Example 7 shows that a Nash equilibrium may admit both a stable matching and a unstable matching. Our next result shows that any Nash equilibrium profile admits at most one true stable matching. If, indeed, a Nash equilibrium profile contains a true stable matching, then that matching will be achieved. Therefore, our result shows that there is no chance for the stable matching mechanism to achieve a unstable matching, as long as a Nash equilibrium admits a true stable matching, no matter how players manipulate their equilibrium strategies.

Lemma 10. *Let* $(\varphi, P) \in \Phi \times \Omega$ *and* $Q \in N(\varphi, P)$ *be an arbitrary Nash equilibrium. Then* $\mid \mathcal{S}(Q) \cap \mathcal{S}(P) \mid \leq 1$. *Further, if* $\mathcal{S}(Q) \cap \mathcal{S}(P) \neq \emptyset$, *then* $\varphi(Q) \in \mathcal{S}(P)$.

Proof. Let $\mu \in \mathcal{S}(Q)$ that is stable for P. If $\varphi(Q) \neq \mu$, then Lemma 6 shows that $\varphi_{S_i}(Q) R_{S_i} \mu(S_i)$ for all S_i and $\varphi_{C_j}(Q) \bar{R}_{C_j} \mu(C_j)$ for all C_j. Lemma 6 also shows that students and colleges strictly prefer $\varphi(Q)$ to μ if they are not the same. Suppose that $\varphi(Q)$ is not stable for P. Then μ is not in the core (in weak domination), contradicting Theorem 5.36 in RS. This shows that if $\mathcal{S}(Q) \cap \mathcal{S}(P) \neq \emptyset$, then $\varphi(Q)$ is stable for P.

Suppose that $\varphi(Q)$ is stable for P and there exists μ in $\mathcal{S}(Q)$ that is also stable for P. Then it follows from Lemma 6 that $\mathcal{S}(P)$ does not form a lattice under P_C dual to P_S, contradicting Corollary 5.32 in RS. □

We may conclude this section with the following theorem which provides a useful necessary and sufficient condition for the implementation of the stable matching correspondence in our context.

Theorem 11 (Necessary and Sufficient). *The stable matching correspondence* $\mathcal{S} : \Omega \rightarrow \mathcal{M}$ *is implementable by a direct revelation mechanism* $(\Pi_{S_i} \Omega_{S_i} \times \Pi_{C_j} \Omega_{C_j}, \varphi)$, *where the outcome function* φ *is a stable matching mechanism in* Φ, *in a subset* $\tilde{N}(\varphi, P) \subset N(\varphi, P)$ *of Nash equilibria if and only if* $\mid \mathcal{S}(Q) \cap \mathcal{S}(P) \mid = 1$ *for all* $Q \in \tilde{N}(\varphi, P)$.

Proof. Suppose that the stable matching correspondence is implementable in a subset $\tilde{N}(\varphi, P)$ of Nash equilibria by a game $\Gamma(\varphi, P)$ induced by a stable matching mechanism φ. Then it follows that $\varphi(Q) \in \mathcal{S}(P)$ for all $Q \in \tilde{N}(\varphi, P)$. Hence, it follows from Lemma 9 that $\mid \mathcal{S}(Q) \cap \mathcal{S}(P) \mid = 1$ for all $Q \in \tilde{N}(\varphi, P)$.

The sufficiency part follows from Lemma 10. □

5 Conclusions

Let us summarize what we may learn from this study about stable matchings in manipulated Nash equilibria. We show that every Nash equilibrium profile admits at most one true stable matching. If, indeed, a Nash equilibrium admits such a matching, then the true stable matching will always be achieved, even though a Nash equilibrium may admit some other unstable matchings; see Theorem 11. Moreover, any Nash equilibrium in truncations contains one and only one matching, stable or not; see Theorem 9. Since the set of stable matchings coincides with the core, the core at a Nash equilibrium in truncations must be "small".

Examples exist such that the Nash, rematching proof, and strong equilibrium outcomes in truncations are unstable; see Example 1. But we show that there are a large class of rematching proof and strong equilibria in truncations at the match point whose outcomes are all stable; see Proposition 3. Moreover, all stable matchings are supported in both rematching proof and strong equilibrium in truncations at the match point; see Proposition 4. Examples also exist such that Nash equilibrium outcomes in truncations at the match point are unstable; see Example 2. We hope that these results provide insights to the noted empirical findings in the literature in a great number of labor markets, e.g., Roth (1984a, 1990b, 1991) and Roth and Peranson (1997a,b).

References

[1] Alcalde, J.: "Implementation of Stable Solutions to Marriage Problem," *Journal of Economic Theory*, (1996).

[2] Danilov, V.: "Implementation via Nash Equilibrium," *Econometrica*, **60**, 43-56 (1992).

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

[4] Kara, T., Sönmez, T.: "Implementation of College Admissions Rules," *Economic Theory* (1997).

[5] Kara, T., Sönmez, T.: "Nash Implementation of Matching Rules," *Journal of Economic Theory*, 68, 425-439 (1996).

[6] Ma, J.: "Stable Matchings and Rematching-proof Equilibria in a Two-sided Matching Market," *Journal of Economic Theory* **66**, 352-369 (1995).

[7] Ma, J.: Stable Matchings and Rematching-proof Equilibria in a Two-sided Matching Market, Working paper #55, The Center for Rationality and Interactive Decision Theory, The Hebrew University of Jerusalem (1994).

[8] Mongell, S., Roth, A.: "Sorority Rush as a Two-sided Matching Mechanism," *American Economic Review*, 81, 441-464 (1991).

[9] Roth, A.E.: "The Evolution of the labor market for medical interns and residents: a case study in game theory, *Journal of Political Economy* **92**, 991-1016 (1984a).

[10] Roth, A.E.: " Misrepresentation and Stability in the Marriage Problem," *Journal of Economic Theory*, 34, 383-387 (1984b).

[11] Roth, A.E.: "The College Admissions Problem is not Equivalent to the Marriage Problem," *Journal of economic Theory*, 36, 277-88 (1985).

[12] Roth, A.E.: Two-sided Matching Markets: An Overview of Some Theory and Empirical Evidence, in "Game Theory and Applications" edited by T. Ichiishi, A. Neyman and Y. Tauman, Academic Press, New York, 1990a.

[13] Roth, A.E.: "New Physicians: A natural experiment in market organization," *Science*, **250**, 1524-1528 (1990b).

[14] Roth, A.E.: "A Natural Experiment in the Organization of Entry Level Labor Markets: regional markets for new physicians and surgeons in the U.K.," *American Economic Review*, **81**, 415-440 (1991).

[15] Roth A.E.: Interim Report #1: Evaluation of the Current NRMP Algorithm, and Preliminary Design of an Application-proposing Algorithm, Consultant's Report, 1996.

[16] Roth, A.E., Peranson, E.: "The Effects of the Change in the NRMP Matching Algorithm," *Journal of the American Medical Association*, 729-732 (1997a).

[17] Roth, A.E., Peranson, E.: "The Redesign of the Matching Market for American Physicians: Some Engineering Aspects of Economic Design," mimeo. University of Pittsburgh, (1997b).

[18] Roth, A.E., Sotomayor, M.: Two-sided Matching: a Study in Game-theoretic Modeling and Analysis, Cambridge University Press, Cambridge 1990.

[19] Roth, A.E., Vande Vate, J.H.: "Incentives in Two-sided Matching with Random Stable Mechanisms," *Economic Theory*, 1, 31-44 (1991).

[20] Shin, S., Suh, S-C.: "A Mechanism Implementing the Stable Rule in Marriage Problem," *Economics Letters*, 51, 185-189 (1996).

[21] Sönmez, T.: "Games of Manipulation in Marriage Problem," *Games and Economic Behavior* (1997).