

Ma, Jinpeng

Working Paper

Manipulation and Stability in the College Admissions Problem

Working Paper, No. 1997-01

Provided in Cooperation with:

Department of Economics, Rutgers University

Suggested Citation: Ma, Jinpeng (1997) : Manipulation and Stability in the College Admissions Problem, Working Paper, No. 1997-01, Rutgers University, Department of Economics, New Brunswick, NJ

This Version is available at:

<https://hdl.handle.net/10419/94315>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

MANIPULATION AND STABILITY IN THE COLLEGE ADMISSIONS PROBLEM*

Jinpeng Ma [†]

This version March 30, 1997

Implementation theory should, I believe, be largely driven by applications; and in principle each application should bring with it some assumption about how the agents in that specific situation will plausibly behave. *John Moore*

Abstract

Roth and Vande Vate (1991) studied the marriage problem and introduced the notion of truncation strategies and showed in an example that unstable matchings can arise at Nash equilibria in truncations. This paper studies the college admissions problem and shows that all rematching proof or strong equilibria in truncations produce stable matchings, even though the equilibrium profiles are manipulated, and all stable matchings can be achieved in rematching proof or strong equilibria in truncations. It is showed that a preference profile that is a rematching proof or strong equilibrium in truncations for one stable matching mechanism is also a rematching proof or strong equilibrium for all stable matching mechanisms. This result shows that there is no difference among all stable matching mechanisms in rematching proof or strong equilibria in truncations, which is in the contrast to the situation in which agents report their preferences in a straightforward manner.

Journal of Economic Literature Classification Numbers: C78, D71.

*I thank Rich McLean, Tetsuji Yamada and the conference participants in The Econometric Society Seventh World Congress (Tokyo, Japan) for their helpful comments on the first version of the paper.

[†]Economics, Rutgers University-Camden, NJ 08102. jinpeng@crab.rutgers.edu.

1 Introduction

In the college admissions problem (Gale and Shapley (1962) and Roth (1985)) there are two disjoint sets of students and colleges with each college a quota of enrollments. The marriage problem is the college admissions problem with each college precisely one enrollment. An outcome in the college admissions problem is a matching such that each student is enrolled in at most one college and each college enrolls students up to its quota. A matching is bilateral in nature in the sense that a student is matched to a college if and only if the college is matched to the student. A pair of a college and student blocks a matching if they are not matched under that matching but the student prefers to be matched with the college and the college prefers the student to one of its assignments under the matching. A matching is stable if it is individually rational for each student and college and is not blocked by any college and student pairs. This notion of (pairwise) stability has been extensively studied in the literature of the matching markets since Gale and Shapley (1962) originated the study of the kind. A well account of the literature has been provided by Roth and Sotomayor (1990) (henceforth, RS)

A stable matching mechanism assigns each profile a stable matching with respect to that profile. Therefore, a stable matching mechanism induces a direct revelation game and it has to work with reported profiles since the underlying true preference profile is not publicly known. When a mechanism is executed in practice, an incentive issue arises: Is it always the best for each participant to reveal his/her true preferences? We say that a mechanism is not manipulable if it does have this incentive property. But very few reasonable mechanisms have the nonmanipulable property. In the college admissions problem Roth's impossibility results show that a stable matching mechanism is manipulable by individual students or colleges; see Section 3. Consequently, as long as students and colleges are sophisticated, the reported profiles may be far from the true one. Therefore, the matching arrived at by a stable matching mechanism, though stable with respect to the reported preference profile, may be unstable with respect to the true preference profile. Indeed Roth showed that in the college admissions problem all individually rational matchings can arise in the Nash equilibria of the game induced by a stable matching mechanism; see Theorem 2. Individually rational matchings include not only the stable matchings but also the unstable ones. This motivates the question: under what conditions does a stable matching mechanism produce an outcome that is stable with respect to the underlying true preference profile?¹

This issue of stability under manipulation in the matching market literature was first addressed in Roth (1984b) for the marriage problem. Roth (1984b) studied the Gale and Shapley (1962) deferred proposal matching mechanism and showed that the Gale and Shapley mechanism produces

¹The question was first articulated in Roth (1984b, 1990). See open question 5 in RS and the question raised in Roth (1990). Our results here provide a partial answer to the question for the college admissions problem. His question has motivated many recent related works; see, e.g., Alcalde (1994), Kara and Sönmez (1993, 1994), Ma (1994b), Shin and Suh (1994), and Sönmez (1996).

stable matchings for the true preference profile in a class of undominated strategy equilibria, even though the equilibrium profiles are manipulated. A truncation strategy Q_a for an agent a contains k elements such that the k (≥ 0) elements of Q_a are the first k elements, with the same order, in her true preference P_a . Roth and Vande Vate (1991) (henceforth, RV) introduced this notion of truncation strategies for the marriage problem and showed that truncation strategies are a class of plausible strategies and the sets of truncation strategies can result in stable matchings at which everyone is matched. Ma (1995) proposed the notion of rematching proofness to extend Roth's (1984b) result to any stable mechanism in the marriage problem and showed that all rematching proof equilibria produce stable matchings. A rematching proof equilibrium is a Nash equilibrium and has the property that there does not exist a pair of college and student who can make profitably joint deviations, given the others stay put. The rematching proof equilibrium is an analogy to the strong equilibrium. It differs from the strong equilibrium in the aspect that it pays attention only to coalitions of size two, one from each side of the market.

Truncations relative to some other more sophisticated strategies has a great advantage in the information aspect in the sense that an agent can submit her truncation strategies regardless of the others' true preferences. Mongell and Roth (1990) found, in the American college sororities in which the "preferential bidding system" is used to match new members to sororities, that very high percentage of reported preferences was truncated to the first choice. But truncations alone do not always produce stable matchings, even at Nash equilibria. Indeed RV showed in an example that truncations may often produce unstable matchings in Nash equilibria; also see Mongell and Roth (1990). Their example is for the marriage problem but it also applies to the college admissions problem. They resolved this unstable issue associated with the Nash equilibria in truncations by devoting to the subgame perfect equilibrium of a dynamic random matching procedure (Roth and Vande Vate (1990)) in which agents have the opportunity to extend their truncation strategies in a random order manner.

The present paper is to address the same unstable issue by the combination of truncations in RV and the rematching proofness in Ma (1994b, 1995) to study the college admissions problem. Our main results show that all rematching proof equilibria in truncations produce stable matchings and all stable matchings can be achieved in rematching proof equilibria in truncations. All rematching proof equilibria in truncations are the same for all stable matching mechanisms. These results also apply to the strong equilibria in truncations. It is known that different stable matching mechanisms produce different stable matchings when agents report their preferences in a straightforward manner. But our results show that all stable matching mechanisms produce the same stable matching at a rematching proof or strong equilibrium in truncations. There is no difference among all stable matching mechanisms in the rematching proof or strong equilibria in truncations.

The remaining of the paper is organized as follows. Section 2 introduces the college admissions

problem. The model follows Gale and Shapley (1962), Roth (1985) and RS. Section 3 presents Roth's impossibility theorems and examples to show that the stable matching correspondence is not implementable in the Nash equilibrium. Section 4 provides a characterization of all Nash equilibria whose outcomes are stable matchings. This result is useful for showing Roth's impossibility theorems and examining which profiles are not Nash equilibria. Section 5 presents some results related to Nash equilibria in truncations. The main results are reported in Section 6. A remark in Section 7 on the undominated strategy equilibrium that does not always produce stable matchings 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, $S = \{S_1, \dots, S_n\}$ of students and $C = \{C_1, \dots, C_r\}$ 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 the set $C \cup \{S_i\}$ and each college $C_j \in C$ has strict preferences P_{C_j} over the set $S \cup \{C_j\}$. Hence P_{C_j} is college C_j 's preferences over individual students. 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. Denote Ω_{S_i} the set of all strict preferences for a student $S_i \in S$ and Ω_{C_j} the set of all strict preferences (over individual students) for a college C_j . Denote $\Omega = \prod_{S_i \in S} \Omega_{S_i} \times \prod_{C_j \in C} \Omega_{C_j}$. The marriage problem is the college admissions problem with $q_{C_j} = 1$ for every college $C_j \in C$.

Following RS, we define an unordered family of elements of any set X to be a collection of elements in which the order is immaterial. Note that an element in an unordered family may contain the same element in X more than once. 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 $S_i \in \mu(C_j)$. Denote \mathcal{M} the set of all matchings.

Definition. Let $\mu, \lambda \in \mathcal{M}$. 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 $\mu(C_j) = \lambda(C_j) \cup \{\sigma\} \setminus \{\tau\}$ for $\tau \in \mu(C_j)$ and $\sigma \notin \mu(C_j)$, then $\mu(C_j) \bar{P}_{C_j} \lambda(C_j)$ if and only if $\sigma P_{C_j} \tau$.

Roth (1985) first reformulated the college admissions problem in Gale and Shapley (1962) to

allow colleges to have preferences over groups of students as well as preferences over individual students. He was the first to propose the responsive preferences and found that the college admissions problem is not equivalent to the marriage problem in many substantial aspects. Henceforth we follow him and 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 and P_{C_j} without a bar for college C_j 's preferences over individual students in the context. No confusion will be made.

We say that a pair $(S_i, C_j) \in S \times C$, with $\mu(S_i) \neq C_j$, blocks a matching μ if $C_j P_{S_i} \mu(S_i)$ 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 $\neg S_i P_{S_i} \mu(S_i)$ for all $S_i \in S$ and $\neg C_j P_{C_j} \sigma$ 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. Denote $S(P)$ the set of all stable matchings with respect to $P \in \Omega$.

Denote $IR(P)$ the set of all individually rational matchings, $P \in \Omega$.

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 $\varphi(P)$ is stable for all $P \in \Omega$. Denote the set of all stable matching mechanisms by Φ .

It follows from Lemma 5.6 in RS and the existence of a stable matching in Gale and Sotomayor (1962) that $S(P)$ is nonempty for any profile of preferences $P \in \Omega$. Therefore, a stable matching mechanism is well defined. A stable matching mechanism $\varphi \in \Phi$ and an underlying true profile of preferences $P \in \Omega$ induce a normal form game $\Gamma(\varphi, P)$.

3 Nash Equilibrium

In the literature of mechanism design, the first question to ask is perhaps if it is always the best for every player to reveal his true preferences. While this nonmanipulable property is quite appealing, very few reasonable mechanisms own this property. In the college admissions problem Roth's impossibility results show that a stable matching mechanism is manipulable by individual students or colleges.

Theorem 1.1 (Roth). *There does not exist any stable matching mechanism that makes it a dominant strategy for all students and all colleges to report their true preferences.*

Theorem 1.2 (Roth). *There does not exist a stable matching mechanism that makes it a dominant strategy for all colleges to report their true preferences.*

Theorem 1.2 is from Roth (1985), which is in the sharp contrast to those in the marriage problem. In the marriage problem the Gale and Shapley man or woman proposal algorithm makes it a dominant strategy for every man or woman to report his or her true preference.

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

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

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

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

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

Denote $N(\varphi, P)$ the set of all Nash equilibria of the game $\Gamma(\varphi, P)$.

Roth also found that even if $S(P)$ contains a unique element, the profile P may not be a Nash equilibrium.

Example 1 (Roth). Let $C = \{C_1, C_2, C_3\}$ and $S = \{S_1, S_2, S_3, S_4\}$. Assume $q_{C_1} = 2$ and $q_{C_2} = q_{C_3} = 1$. The preferences of colleges are as follows: $P_{C_1} = (S_1, S_2, S_3, S_4, C_1)$, $P_{C_2} = (S_1, S_2, S_3, S_4, C_2)$ and $P_{C_3} = (S_3, S_1, S_2, S_4, C_3)$. The preferences of students are as follows: $P_{S_1} = (C_3, C_1, C_2, S_1)$, $P_{S_2} = (C_2, C_1, C_3, S_2)$, $P_{S_3} = (C_1, C_3, C_2, S_3)$ and $P_{S_4} = (C_1, C_2, C_3, S_4)$. The matching $x = [(C_1; S_3, S_4), (C_2; S_2), (C_3; S_1)]$ is the unique stable matching with respect to this profile P of preferences. Now suppose that college C_1 reports that $Q_{C_1} = (S_2, S_4, C_1, S_1, S_3)$. The unique stable matching with respect to the profile (Q_{C_1}, P_{-C_1}) is $y = [(C_1; S_2, S_4), (C_2; S_1), (C_3; S_3)]$. Because college C_1 prefers S_2 to S_3 , by responsiveness, college C_1 prefers the classes $\{S_2, S_4\}$ to $\{S_3, S_4\}$. Thus, college C_1 gets better off by unilaterally deviating. \square

A dominant strategy equilibrium is quite demanded. One may wonder whether all Nash equilibrium outcomes are stable. Unfortunately, this is not the case. The Nash equilibrium bites two little. Almost all meaningful matchings could be sustained by Nash equilibria.

Theorem 2 (Roth). For any $\varphi \in \Phi$, $IR(P) \subseteq \varphi(N(\varphi, P))$ for all profiles $P \in \Omega$.

Roth's results and example show that when a stable matching mechanism is implemented, students and colleges may have strong incentives to misreport their preferences. With the misreported preferences, the outcome achieved by a stable matching mechanism though stable with respect to the reported preferences may not be stable for the true preferences. Given this common knowledge, the obtained unstable matching under a stable matching mechanism may provide incentives for players not to commit to the achieved outcome because they may bypass the mechanism and do better on their own. Therefore the implemented mechanism may break down. This could happen even in the Nash equilibrium outcomes. But the empirical works found otherwise: Almost all stable matching mechanisms employed in the entry-level labor markets successfully resolved the market failures that seem associated with instability experienced when the markets are decentralized; see Roth (1984a, 1991). Therefore there is some gap between the empirical findings and the theoretical results. This motivates the question whether there is some meaningful refinement of the Nash equilibrium that is helpful to eliminate unwanted unstable outcomes. We attack this question in Section 6. First we would like to present several results about the Nash equilibrium and the truncations.

4 A Characterization of Nash Equilibria

In this section we provide a result that is useful to show Theorems 1.1 and 1.2. We first show a lemma that is useful for its own sake. Lemma 1 shows that there is no matching $\mu \in S(Q)$ at a Nash equilibrium profile Q such that a college C_j strictly prefers any new class $\mu(C_j)$ to her old class $\varphi_{C_j}(Q)$ with respect to any responsive preference \bar{P}_c .

Lemma 1. *Suppose $Q \in N(\varphi, P)$ is a Nash equilibrium. Then for any matching $\mu \in S(Q)$, $\varphi_{S_i}(Q) P_{S_i} \mu(S_i)$ for every $S_i \in S$ such that $\varphi_{S_i}(Q) \neq \mu(S_i)$ and $\neg \mu(C_j) \bar{P}_{C_j} \varphi_{C_j}(Q)$ for every $C_j \in C$ such that $\varphi_{C_j}(Q) \neq \mu(C_j)$.*

Proof. First, suppose on the contrary that $\mu(S_i) P_{S_i} \varphi_{S_i}(Q)$ for some $S_i \in S$ with $\varphi_{S_i}(Q) \neq \mu(S_i)$. Then let $Q'_{S_i} = (\mu(S_i), S_i, \dots)$ and note that $\{\mu\} \in S(Q_{-S_i}, Q'_{S_i})$. Theorem 5.12 in RS shows that the set of students employed is the same at every stable matching. Therefore, $\varphi_{S_i}(Q_{-S_i}, Q'_{S_i}) = \mu(S_i)$. It follows that $\varphi_{S_i}(Q_{-S_i}, Q'_{S_i}) P_{S_i} \varphi_{S_i}(Q_{-S_i}, Q_{S_i})$, which contradicts that Q is a Nash equilibrium.

Second, suppose on the contrary that $\mu(C_j) \bar{P}_{C_j} \varphi_{C_j}(Q)$ for some $C_j \in C$. Then let $Q'_{C_j} = (S_{i_1}, \dots, S_{i_q}, C_j, \dots)$, where $q = |S \cap \mu(C_j)| \leq q_{C_j}$, such that $S_{i_k} Q'_{C_j} S_{i_l}$ if and only if $S_{i_k} Q_{C_j} S_{i_l}$ for any $S_{i_k}, S_{i_l} \in \mu(C_j) \cap S$ and $C_j Q'_{C_j} S_i$ for all $S_i \in S \setminus \mu(C_j)$. Now we show that $\mu \in S(Q_{-C_j}, Q'_{C_j})$. Suppose not, then $\exists (S_i, \tilde{C}_j) \in S \times C$ with $\mu(S_i) \neq \tilde{C}_j$ such that $\tilde{C}_j Q_{S_i} \mu(S_i)$ and $S_i Q_{\tilde{C}_j} \sigma$ for some $\sigma \in \mu(\tilde{C}_j)$. Note that $\tilde{C}_j \neq C_j$ by the construction of Q'_{C_j} . This implies that $\mu \notin S(Q)$ because

(S_i, \bar{C}_j) blocks μ . Therefore, $\mu \in S(Q_{-C_j}, Q'_{C_j})$.

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 that if $C_j \in \mu(C_j)$, then $\lambda(C_j) = \mu(C_j)$ for all $\lambda \in S(Q_{-C_j}, Q'_{C_j})$ since $\mu \in S(Q_{-C_j}, Q'_{C_j})$.

Now suppose $\mu(C_j) \neq \lambda(C_j)$ for some $\lambda \in S(Q_{-C_j}, Q'_{C_j})$ and $\exists C_j \in \mu(C_j)$. Then there must exist some $C_j \in \lambda(C_j)$. Theorem 5.13 in RS shows that $\mu(C_j) = \lambda(C_j)$. This is a contradiction. Therefore $\varphi_{C_j}(Q_{-C_j}, Q'_{C_j}) = \mu(C_j)$. And then, $\varphi_{C_j}(Q_{-C_j}, Q'_{C_j}) \bar{P}_{C_j} \varphi_{C_j}(Q_{-C_j}, Q_{C_j})$. This contradicts that Q is a Nash equilibrium. \square

Given $\varphi \in \Phi$ and $P \in \Omega$, the outcome $\varphi(Q) \in S(Q)$ for $Q \in \Omega$. But $\varphi(Q)$ may or may not be in $S(P)$. Of course, if $\varphi(Q) \in S(P)$, then $S(Q)$ contains at least one element in $S(P)$. But, the set $S(Q)$ may contain more than one such outcome. The next result shows that whenever the outcome $\varphi(Q)$ is in $S(P)$, then $\varphi(Q)$ is the unique outcome in both $S(P)$ and $S(Q)$, providing the profile Q is a Nash equilibrium.

Theorem 3. *For every $Q \in N(\varphi, P)$ such that $\varphi(Q) \in S(P)$, $S(Q) \cap S(P) = \{\varphi(Q)\}$.*

Proof. Suppose, by the way of contradiction, that $|S(Q) \cap S(P)| > 1$. Thus there exists $\mu \in S(Q) \cap S(P)$ 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 1 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 $S(P)$ is a lattice under the common preferences of colleges, \bar{P}_C , and due to the common preferences of students, P_S . \square

Clearly Theorem 3 is not true for an arbitrary profile Q . It is often difficult in the college admissions problem to determine whether a profile is a Nash equilibrium. The contribution of Theorem 3 helps eliminate a large class of profiles that are not in the equilibrium. Especially, we can use Theorem 3 to show Theorem 2.1. For examples, suppose that there are two stable matchings in $S(P)$. Then Theorem 3 shows that such a profile P is not an equilibrium profile for any game $\Gamma(\varphi, P)$ induced by any $\varphi \in \Phi$. This shows that there does not exist any game $\Gamma(\varphi, P)$ such that the profile P is a dominant strategy equilibrium.

5 Truncations

We saw from Roth's result in Theorem 2 that the set of Nash equilibrium outcomes includes both stable and unstable matchings. We now use the notion of truncations which is introduced in RV for the marriage problem and obtain several results for the college admissions problem that are quite similar to theirs in the marriage problem. Truncation strategies alone may produce stable matchings under certain conditions.

We first introduce the truncations in RS. A truncation strategy Q_{S_i} (with respect to P_{S_i}) for a student S_i contains k ($0 \leq k \leq r$) elements such that the k elements of Q_{S_i} are the first k elements, with the same order, in her true preference P_{S_i} . Similarly, a truncation strategy Q_{C_j} for a college C_j (with respect to P_{C_j}) contains k ($0 \leq k \leq n$) elements such that the k elements of Q_{C_j} are the first k elements, with the same order, in her true preference P_{C_j} . We first show that all stable matchings can be achieved in the strong equilibrium in truncations. Therefore, a strong Nash equilibrium in truncations always exists. This result is somehow sharpening Theorem 3 (i) in RV.

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

$$\varphi_k(Q_{-T}, Q'_T) P_k \varphi_k(Q)$$

for all $k \in T \cap S$ and

$$\varphi_k(Q_{-T}, Q'_T) \bar{P}_k \varphi_k(Q)$$

for all $k \in T \cap C$. Denote the set of all strong equilibria in truncations of the game $\Gamma(\varphi, P)$ by $N^{ts}(\varphi, P)$.

One additional notation is helpful. Given a matching $\mu \in \mathcal{M}$ and $P \in \Omega$, define $t_j^\mu \in \mu(C_j) \cap S$ such that $\tau P_{C_j} t_j^\mu$ for all $\tau \in \mu(C_j) \cap S \setminus \{t_j^\mu\}$. Given a preference profile $P \in \Omega$ and a matching μ , we say a truncation strategy profile $Q^\mu \in \Omega$ *exact*, with respect to P and μ , if (a) the strategy $Q_{S_i}^\mu \in \Omega_{S_i}$ for S_i ranks the colleges in the same order as in P_{S_i} , up to $\mu(S_i)$, who is then immediately followed by S_i , whenever $\mu(S_i) \neq S_i$, and $Q_{S_i}^\mu = Q_{S_i}$ if $\mu(S_i) = S_i$, for all $S_i \in S$; (b) the strategy $Q_{C_j}^\mu \in \Omega_{C_j}$ for C_j ranks all students in the same order as in P_{C_j} , up to the student t_j^μ , who is then immediately followed by C_j , whenever $\mu(C_j) \cap S \neq \emptyset$, and $Q_{C_j}^\mu = Q_{C_j}$ if $\mu(C_j) \cap S = \emptyset$, for all $C_j \in C$. For example, suppose $P_{S_i} = (C_1, C_2, C_3, C_4, S_i, \dots)$ and $\mu(S_i) = C_2$. Then $Q_{S_i}^\mu = (C_1, C_2, S_i, \dots) \in \Omega_{S_i}$ is an exact truncation for S_i . $Q_{S_i} = (C_1, C_2, C_3, S_i, \dots) \in \Omega_{S_i}$ and $Q'_{S_i} = (C_1, S_i, \dots) \in \Omega_{S_i}$ are both truncation strategies for S_i but they are not exact truncations. Similarly, suppose college C_j 's true preference is $P_{C_j} = (S_1, S_2, S_3, S_4, S_5, C_j, \dots) \in \Omega_{C_j}$ and his quota $q_{C_j} = 3$. Given a matching μ such that $\mu(C_j) = \{S_1, S_3, C_j\}$, one exact truncation strategy for C_j is $Q_{C_j}^\mu = (S_1, S_2, S_3, C_j, \dots) \in \Omega_{C_j}$. Both $Q_{C_j} = (S_1, S_2, C_j, \dots) \in \Omega_{C_j}$ and

$Q'_{C_j} = (S_1, S_2, S_3, S_4, C_j, \dots)$ are truncations but they are not exact truncations. Hence if a student S_i is matched with a college, then the ranking orders and the elements in her exact truncations before S_i are uniquely determined by P_{S_i} and her matched college. If a college C_j is matched with at least one student under a matching, the ranking orders and the elements in its exact truncations before C_j are uniquely determined by P_{C_j} and the least preferred student in its assignment under that matching. If a student is not matched with any college or a college is not matched with at least one student under a matching, any truncation for the student or college is exact.

Theorem 4. *For all profiles $P \in \Omega$ and all $\varphi \in \Phi$, $S(P) \subseteq \varphi(N^{ts}(\varphi, P))$.*

Proof. It follows from Lemma 5.6 in RS that $S(P) \neq \emptyset$ for any profile $P \in \Omega$. Given $\mu \in S(P)$, let Q^μ be an exact strategy profile in truncations. Then $\varphi(Q^\mu) = \mu$ for all stable matching mechanisms $\varphi \in \Phi$ since $S(Q^\mu) = \{\mu\}$. We claim that this is a strong equilibrium in truncations of $\Gamma(\varphi, P)$ for all $\varphi \in \Phi$.

Suppose that this is not true for some stable matching mechanism φ . Then by the definition of the strong equilibrium, there exists a coalition $T \subseteq C \cup S$, with strategies $R_T \in \prod_{k \in T} \Omega_k$, such that, for all $S_i \in T \cap S$,

$$\varphi_{S_i}(\tilde{Q}) P_{S_i} \varphi_{S_i}(Q^\mu), \quad (1)$$

and for all $C_j \in T \cap C$,

$$\varphi_{C_j}(\tilde{Q}) \bar{P}_{C_j} \varphi_{C_j}(Q^\mu), \quad (2)$$

where $\tilde{Q} = (Q^\mu_{-T}, R_T)$.

Define $T_1 = \{S_k \in S : \varphi_{S_k}(\tilde{Q}) P_{S_k} \varphi_{S_k}(Q^\mu)\}$ and $T_2 = \{C_l \in C : \varphi_{C_l}(\tilde{Q}) \bar{P}_{C_l} \varphi_{C_l}(Q^\mu)\}$. Thus, $T \subseteq T_1 \cup T_2$. We consider two cases.

(a). $T_2 = \emptyset$. Then $T_1 \neq \emptyset$ since $T \subseteq T_1$. Let $C_l = \varphi_{S_i}(\tilde{Q})$. It follows that $S_i \notin \varphi_{C_l}(Q^\mu) = \mu(C_l)$ since $C_l P_{S_i} \varphi_{S_i}(Q^\mu) = \mu(S_i)$. Because $T_2 = \emptyset$, by the definition of $Q^\mu_{C_l}$ that is a truncation of P_{C_l} , it follows that $S_i Q^\mu_{C_l} t_l^\mu$ whenever $\mu(C_l) \cap S \neq \emptyset$ or $S_i Q^\mu_{C_l} C_l$ if $\mu(C_l) \cap S = \emptyset$. Since Q_{C_l} is a truncation of P_{C_l} , we also have that $S_i P_{C_l} t_l^\mu$ whenever $\mu(C_l) \cap S \neq \emptyset$ or $S_i P_{C_l} C_l$ if $\mu(C_l) \cap S = \emptyset$. Hence (S_i, C_l) blocks μ with respect to P . This is a contradiction.

(b). $T_2 \neq \emptyset$. For any college C_j in T_2 , by the responsiveness of \bar{P}_{C_j} and (2), there exists a student $\sigma \in \varphi_{C_j}(\tilde{Q})$ and $\tau \in \varphi_{C_j}(Q^\mu)$ with $\sigma \notin \varphi_{C_j}(Q^\mu)$ such that $\sigma P_{C_j} \tau$. Note that σ is not in T_1 . Otherwise, $C_j P_\sigma \varphi_\sigma(Q^\mu)$ implies that (σ, C_j) blocks μ with respect to P . This means that $\neg C_j P_\sigma \varphi_\sigma(Q^\mu)$. Now by the individual rationality, it follows that $\varphi_\sigma(Q^\mu) P_\sigma C_j P_\sigma \sigma$. But this is impossible because there does not exist C_j between $\varphi_\sigma(Q^\mu)$ and σ . \square

Given a profile $P \in \Omega$, Theorem 4 shows that there always exist a strong equilibrium in truncations of $\Gamma(\varphi, P)$ for all $\varphi \in \Phi$. Theorem 5 below shows that the set of stable matchings with

respect to a Nash equilibrium in truncations always contains a unique element.

Theorem 5. Let $Q \in N(\varphi, P)$, $(\varphi, P) \in \Phi \times \Omega$, in truncations. Then $S(Q) = \{\varphi(Q)\}$.

Proof. Suppose on the contrary that $\mu \in S(Q)$. Then by Lemma 1, $\varphi_{S_i}(Q)P_{S_i}\mu(S_i)$ for all $S_i \in S$ with $\varphi_{S_i}(Q) \neq \mu(S_i)$ and $\neg\mu(C_j)\bar{P}_{C_j}\varphi_{C_j}(Q)$ for all $C_j \in C$ with $\mu(C_j) \neq \varphi_{C_j}(Q)$. 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 $\varphi_{S_i}(Q) \neq \mu(S_i)$ and $\neg\mu(C_j)\bar{Q}_{C_j}\varphi_{C_j}(Q)$ for all $C_j \in C$ with $\mu(C_j) \neq \varphi_{C_j}(Q)$. But this is a contraction to Theorem 5.33 in RS which shows that $S(Q)$ is a lattice under the common preferences of colleges, Q_C , and dual to the common preferences of students, Q_S . \square

Denote $N^t(\varphi, P)$ the set of all Nash equilibria in truncations of the game $\Gamma(\varphi, P)$, $(\varphi, P) \in \Phi \times \Omega$.

Corollary 1. For any $P \in \Omega$, $N^t(\varphi_1, P) = N^t(\varphi_2, P)$ for any $\varphi_1, \varphi_2 \in \Phi$.

Theorem 3(ii) in RV shows that no set of truncation strategies can result in an unstable matching at which everyone is matched. Their result in the college admissions problem has the following form.

Theorem 6. Let $Q \in \Omega$ be any truncation preference profile and $\mu \in S(Q)$ such that $\mu(S_i) \neq S_i$ for all $S_i \in S$ and $\mu(C_j) \cap S \neq \emptyset$ for all $C_j \in C$. Then $\mu \in S(P)$.

Proof. It follows, from the truncations and $\mu \in IR(Q)$, that $\mu \in IR(P)$. Now suppose on the contrary that $\mu \notin S(P)$, then $\exists(S_i, C_j) \in S \times C$ such that $C_j P_{S_i} \mu(S_i)$ and $S_i P_{C_j} \sigma$ for some $\sigma \in \mu(C_j)$. Recall that $t_j^\mu \in \mu(C_j) \cap S$ such that $\tau Q_{C_j} t_j^\mu$ for all $\tau \in \mu(C_j) \cap S \setminus \{t_j^\mu\}$. τ may not exist if t_j^μ is the only element in $\mu(C_j)$ but t_j^μ always exists since $\mu(C_j) \cap S \neq \emptyset$. Because $C_j P_{S_i} \mu(S_i) P_{S_i} S_i$ and Q_{S_i} is a truncation of P_{S_i} , it follows that $C_j Q_{S_i} \mu(S_i) Q_{S_i} S_i$. Because $S_i P_{C_j} \sigma$ for some $\sigma \in \mu(C_j)$ and Q_{C_j} is a truncation of P_{C_j} , it follows that $S_i Q_{C_j} t_j^\mu$. This shows that (S_i, C_j) blocks μ with respect to Q , a contradiction to $\mu \in S(Q)$. \square

Truncations in Theorem 6 are crucial. In Example 1 every student is matched with a college and every college is matched with at least one student under the matching y . But the matching y is not stable with respect to P . This is because the profile (Q_{C_1}, P_{-C_1}) is not a truncation profile. Nevertheless truncations do not always produce stable matchings, even at a Nash equilibrium.

Example 2 (RV). Let $n = r = 2$, $P_{S_i} = (C_1, C_2, S_i)$ for $i = 1, 2$ and $P_{C_j} = (S_1, S_2, C_j)$ for $j = 1, 2$. Each college has quota of 1. Construct a truncated profile Q as follows: $Q_{S_i} = (C_1, S_i, C_2)$

for $i = 1, 2$ and $Q_{C_j} = (S_1, C_j, S_2)$ for $j = 1, 2$. Now the matching $\mu = [(S_1; C_1), (S_2; S_2), (C_2; C_2)]$ is not stable for the true profile P but it is a unique stable matching for the profile Q . Thus, $\varphi(Q) = \mu$ for all stable mechanisms. To see that Q is a Nash equilibrium of the game $\Gamma(\varphi, P)$, it is sufficient to observe that C_2 can not do better by unilaterally listing S_2 and S_2 can not do better by unilaterally listing C_2 . Further, this equilibrium is an undominated strategy equilibrium (Lemma 2 in RV). \square

6 Rematching Proof Equilibrium

A reported profile is a rematching proof equilibrium if it is a Nash equilibrium and there are no pairs of a college and student who have joint deviations to make both better off, given the others' reports fixed. Therefore, the notion of rematching proof equilibria excludes all possibilities of strategically joint playing by a pair of a college and student. The rematching proof equilibrium may be considered as bilateral rationality, in contrast to the unilateral rationality of the Nash equilibrium. Since the college admissions problem of the study is bilateral in nature, the proposed rematching proof equilibrium notion captures this bilateral feature of the market.

Definition. A *rematching proof equilibrium* Q of a game $\Gamma(\varphi, P)$ is a Nash equilibrium and has the property that if

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

then

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

for all $(S_i, C_j) \in S \times C$, all $(Q'_{S_i}, Q'_{C_j}) \in \Omega_{S_i} \times \Omega_{C_j}$. Denote the set of all rematching proof equilibria in truncations of the game $\Gamma(\varphi, P)$ by $N^{trp}(\varphi, P)$.

The intuition why the rematching proof equilibria are helpful in producing stable matchings may be understood by the marriage problem. Assume that $q_{C_j} = 1$ for all $C_j \in C$. Suppose $Q \in N^{rp}(\varphi, P)$ is a rematching proof equilibrium and $\varphi(Q)$ is not stable. That is, there is a pair of student S_i and college C_j that blocks $\varphi(Q)$, i.e., $S_i P_{C_j} \varphi_{C_j}(Q)$ and $C_j P_{S_i} \varphi_{S_i}(Q)$. Now construct two strategies $Q'_{S_i} = (C_j, S_i, \dots)$ and $Q'_{C_j} = (S_i, C_j, \dots)$ and consider the outcome of a stable matching mechanism at the new profile $(Q_{-\{S_i, C_j\}}, Q'_{S_i}, Q'_{C_j})$. It is easy to see that S_i and C_j will be matched with each other under a stable matching mechanism. This disputes the assumption that Q is rematching proof.

While the above intuition is instructive but it is somehow misleading for the college admissions problem. For example, y is not stable in Example 1 because C_1 and S_3 blocks it. One may hope

that both the student S_3 and the college C_1 will get better off by constructing such strategies $Q'_{S_3} = (C_1, S_3, \dots) \in \Omega_{S_3}$ and $Q'_{C_1} = (S_2, S_3, C_1, \dots) \in \Omega_{C_1}$, the same idea used in the marriage problem. But this is not the case because the unique stable outcome at the new profile $(P_{-\{C_1, S_3\}}, Q'_{C_1}, Q'_{S_3})$ is $z = [(C_1; S_3, C_1), (C_2; S_2), (C_3, S_1), (S_4, S_4)]$, which matches college C_1 with $\{S_3, C_1\}$ rather than $\{S_2, S_3\}$. Because C_1 's preferences over group of students are responsive and transitive, C_1 prefers $\{S_2, S_4\}$ to $\{S_2, C_1\}$ to $\{S_3, C_1\}$. Hence C_1 gets worse off in comparison with the matching y in which C_1 is matched with $\{S_2, S_4\}$. The same idea that works for the marriage problem does not apply to the college admissions problem. This is why the proof of Theorem 7 becomes more sophisticated in the college admissions problem relative to the marriage problem.

Given a college admissions problem one can construct a related marriage problem by making as many copies as its quota for each college, with each copy of a college one enrollment and the same preferences over individual students. A matching is stable in this related marriage problem if and only if it is stable for the college admissions problem (Lemma 5.6 in RS). But this technique is not so useful to discuss the strategic aspects of a stable matching mechanism because it requires every copy of the same college in the related marriage problem must have the same strategies and the same underlying true preferences. A Nash equilibrium in the related marriage problem in which every copy is considered as an independent individual is not identical to a Nash equilibrium in the college admissions problem, and vice versa.

Lemma 2. *Every exact truncation strategy profile that is also rematching proof produces a stable matching with respect to P .*

Proof. Let $(\varphi, P) \in \Phi \times \Omega$. Let $Q \in N^{trp}(\varphi, P)$ be an exact rematching proof equilibrium in truncations. Suppose on the contrary that $\varphi(Q) \notin S(P)$. 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 and $\varphi(Q) \in IR(Q)$). We discuss four cases.

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

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 $\sigma \in \varphi_{C_j}(Q)$ and Q_{C_j} is a truncation of P_{C_j} , it follows that $S_i Q_{C_j} t_j^{\varphi(Q)}$ ($t_j^{\varphi(Q)}$ exists since $\varphi_{C_j}(Q) \subseteq S$). This shows that (S_i, C_j) blocks $\varphi(Q)$ with respect to Q , a contradiction to $\varphi(Q) \in S(Q)$.

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

Since Q_{C_j} is a truncation of P_{C_j} , the assumption that $S_i P_{C_j} \sigma$ for some $\sigma \in \varphi_{C_j}(Q)$ implies that $S_i Q_{C_j} t_j^{\varphi(Q)} Q_{C_j} C_j$. 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 . Now let $Q'_{S_i} = (\dots, C_j, S_i, \dots)$ be a truncation of P_{S_i} , up to C_j . Then we show that $\varphi_{S_i}(Q_{-S_i}, Q'_{S_i}) \neq S_i$.

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

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

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

contradicting the fact that Q is a Nash equilibrium.

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

Construct a strategy $Q'_{S_i} = (\dots, C_j, S_i, \dots)$ for the student S_i by shifting C_j right to the front of S_i in Q_{S_i} and the rest ranking order in Q'_{S_i} is exactly the same as in Q_{S_i} . Construct a strategy Q'_{C_j} for the college C_j such that $Q'_{C_j} = Q_{C_j}$ if $S_i Q_{C_j} C_j$; otherwise, $Q'_{C_j} = (\dots, S_i, C_j, \dots)$ by shifting S_i right to the front of C_j and the rest ranking order in Q'_{C_j} is exactly the same as in Q_{C_j} . 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, if $\varphi_{S_i}(\tilde{Q}) = S_i$, then $\varphi_{C_j}(\tilde{Q}) \subseteq S$. Otherwise, $\exists C_j \in \varphi_{C_j}(\tilde{Q})$ implies that (S_i, C_j) blocks $\varphi(\tilde{Q})$ with respect to \tilde{Q} . But if $\varphi_{C_j}(\tilde{Q}) \subseteq S$, then there exists $\tau \in \varphi_{C_j}(\tilde{Q})$ such that $\tau \notin \varphi_{C_j}(Q)$. Because $\tau \neq S_i$, it follows that $C_j Q_{\tau} \varphi_{\tau}(Q)$ since Q_{τ} is a truncation. By the construction of Q'_{C_j} , we have that $\tau Q_{C_j} C_j$. Therefore (τ, C_j) blocks $\varphi(Q)$ with respect to Q . Therefore, $\varphi_{S_i}(\tilde{Q}) \neq S_i$.

Second, suppose that $\varphi_{S_i}(\tilde{Q}) \neq C_j$. Then $\varphi(\tilde{Q}) \in S(Q)$ by the constructions of Q'_{S_i} and Q'_{C_j} . But Theorem 5.12 in RS shows that the set of students enrolled is the same at every stable matching. Therefore, the fact that $\varphi_{S_i}(\tilde{Q}) \neq S_i$ and $\varphi(\tilde{Q}) \in S(Q)$ implies that $\varphi_{S_i}(Q) \neq S_i$. This is a contradiction to our assumption in *Case c*. Therefore, $\varphi_{S_i}(\tilde{Q}) = C_j$.

Finally, $\varphi_{S_i}(\tilde{Q}) = C_j$ implies $\varphi_{C_j}(\tilde{Q}) = \varphi_{C_j}(Q) \cup \{S_i\} \setminus \{C_j\}$. This is because $S(Q) = \{\varphi(Q)\}$ and any blocking pair (τ, C_j) of $\varphi(\tilde{Q})$ will be a blocking pair of $\varphi(Q)$ with respect to Q , by the constructions of Q'_{S_i} and Q'_{C_j} .

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

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

and

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

which is a contradiction to the rematching proofness of Q . This completes the proof of *Case c*.

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

Let $S_{k_1}, \dots, S_{k_q} \in \varphi_{C_j}(Q) \cap S$ such that $S_{k_1}Q_{C_j} \cdots Q_{C_j}S_{k_q}$. Note that $q < q_{C_j}$ and $t_j^{\varphi(Q)} = S_{k_q}$. If $t_j^{\varphi(Q)}$ does not exist, let $Q'_{S_i} = (C_j, S_i, \dots)$ and $Q'_{C_j} = (S_i, C_j, \dots)$. Clearly, $\varphi_{S_i}(Q_{-\{S_i, C_j\}}, Q'_{S_i}, Q'_{C_j}) = C_j$. This is a contradiction to the rematching proofness of Q .

Assume that $t_j^{\varphi(Q)}$ exists. We want to construct a strategy Q'_{S_i} for S_i and a strategy Q'_{C_j} for C_j to crowd out C_j .

Let $Q'_{S_i} = (C_j, S_i, \dots)$ and $Q'_{C_j} = (S_i, S_{k_1}, \dots, S_{k_{(q-1)}}, S_{k_q}, C_j, \dots)$. We show that $\varphi_{S_i}(\tilde{Q}) = C_j$ and $\varphi_{C_j}(\tilde{Q}) = \{S_i\} \cup \varphi_{C_j}(Q) \setminus \{C_j\}$, where $\tilde{Q} = (Q_{-\{S_i, C_j\}}, Q'_{S_i}, Q'_{C_j})$.

First, suppose that $\varphi_{S_i}(\tilde{Q}) = S_i$. Then there exists $C_j \in \varphi(\tilde{Q})$. This follows from the construction of Q'_{C_j} . But then this implies that $\varphi(\tilde{Q}) \notin S(\tilde{Q})$ since (S_i, C_j) blocks $\varphi(\tilde{Q})$. This shows that $\varphi_{S_i}(\tilde{Q}) = C_j$.

Second, suppose that $\varphi_{C_j}(\tilde{Q}) \neq \{S_i\} \cup \varphi_{C_j}(Q) \setminus \{C_j\}$. 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 $S_k \neq S_i$, it follows that $C_j Q_{S_k} S_k$ since $\varphi_{S_k}(Q) = C_j$. Because $\exists C_j \in \varphi_{C_j}(\tilde{Q})$ and $\varphi(\tilde{Q}) \in S(\tilde{Q})$, it follows that $\varphi_{S_k}(\tilde{Q}) Q_{S_k} C_j$ by the fact that Q_{S_k} is a truncation of P_{S_k} up to C_j . Hence, we have that $\varphi_{S_k}(\tilde{Q}) Q_{S_k} C_j Q_{S_k} S_k$.

Let $C_l = \varphi_{S_k}(\tilde{Q})$. Note that $C_l \neq C_j$. It follows that $S_k Q_{C_l} C_l$, by individual rationality. Since $S_k \notin \varphi_{C_l}(Q)$, it follows from the stability of $\varphi(Q)$ that $\varphi_{C_l}(Q) \subseteq S$. From the stability of $\varphi(Q)$ again, $\sigma Q_{C_l} S_k$ 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 an exact truncation of P_{C_l} . There is no such a student S_k between σ and C_l . This is a contradiction.

Therefore, it is shown that $\varphi_{S_i}(Q') = C_j$ and $\varphi_{C_j}(Q') = \{S_i\} \cup \varphi_{C_j}(Q) \setminus \{C_j\}$. Since \bar{P}_{C_j} is responsive and $S_i P_{C_j} C_j$, we have that

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

and

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

which contradicts the assumption that Q is rematching proof. This completes the proof of *Case d*.

Four cases together complete the proof of the lemma since $(\varphi, P) \in \Phi \times \Omega$ is arbitrarily chosen and $S(Q) = \{\varphi(Q)\}$ according to Theorem 5. \square

We now drop the exact condition in Lemma 2 to obtain the main result in this paper.

Theorem 7. *Every rematching proof equilibrium in truncations produces a stable matching with respect to P .*

Proof. Let $(\varphi, P) \in \Phi \times \Omega$, and $Q \in N^{trp}(\varphi, P)$ be a rematching proof equilibrium in truncations. Denote $Q^* = Q^{\varphi(Q)}$ an exact preference profile with respect to P and $\varphi(Q)$. Theorem 5 shows that $S(Q) = \{\varphi(Q)\}$. By the construction of Q^* , $\varphi(Q) \in S(Q^*)$ and it follows that $S(Q^*) = \{\varphi(Q)\}$. Then it follows from Lemma 2 that $\varphi(Q^*)$ is in $S(P)$ if we can show that Q^* is rematching proof.

Step 1. Q^* is a Nash equilibrium.

Suppose on the contrary that Q^* is not a Nash equilibrium. Then there exists $S_i \in S$, with $Q'_{S_i} \in \Omega_{S_i}$, such that

$$\varphi_{S_i}(Q^*_{-S_i}, Q'_{S_i}) P_{S_i} \varphi_{S_i}(Q^*) = \varphi_{S_i}(Q)$$

or there exists $C_j \in C$, with $Q'_{C_j} \in \Omega_{C_j}$, such that

$$\varphi_{C_j}(Q^*_{-C_j}, Q'_{C_j}) \bar{P}_{C_j} \varphi_{C_j}(Q^*) = \varphi_{C_j}(Q).$$

Consider first that it is S_i who deviates. Then denote $C_l = \varphi_{S_i}(Q^*_{-S_i}, Q'_{S_i})$. It follows that $S_i Q^*_{C_l} \sigma$ for some $\sigma \in \varphi_{C_l}(Q^*)$. We consider two cases: (i) $\varphi_{S_i}(Q^*) \neq S_i$ and (ii) $\varphi_{S_i}(Q^*) = S_i$. In case (i), we have that $C_l Q^*_{S_i} \varphi_k(Q^*)$ and then (k, j) blocks $\varphi(Q^*)$ under Q^* . (Δ): In case (ii), let $\hat{Q}_{S_i} = (C_l, S_i, \dots)$. For any $\mu \in S(Q_{-S_i}, \hat{Q}_{S_i})$, $\mu(S_i) = S_i$ implies that (S_i, C_l) blocks μ under $(Q_{-S_i}, \hat{Q}_{S_i})$. Therefore $\mu(S_i) = C_l$ for every $\mu \in S(Q_{-S_i}, \hat{Q}_{S_i})$. But this is a contradiction to the fact that Q is a Nash equilibrium.

Second consider that it is C_j that deviates. Then $C_j Q^*_{\tau} \varphi_{\tau}(Q)$ for every $\tau \in \varphi_{C_j}(Q^*_{-C_j}, Q'_{C_j})$ and $\tau \notin \varphi_{C_j}(Q)$. Consider two cases: (i) $\varphi_{C_j}(Q^*) \cap S \neq \emptyset$ and (ii) $\varphi_{C_j}(Q^*) \cap S = \emptyset$. In case (i) there exists at least one $\tau \in \varphi_{C_j}(Q^*_{-C_j}, Q'_{C_j})$ and $\tau \notin \varphi_{C_j}(Q)$ such that $\tau Q^*_{C_j} t_j^{\varphi(Q)}$, and then $\tau Q_{C_j} t_j^{\varphi(Q)}$. But then (τ, C_j) blocks $\varphi(Q)$ under Q . In case (ii), choose a student τ such that $C_j P_{\tau} \varphi_{\tau}(Q)$, since there exists τ such that $C_j Q^*_{\tau} \varphi_{\tau}(Q)$. Let $Q'_{\tau} = (C_j, \tau, \dots) \in \Omega_{\tau}$ and $Q'_{C_j} = (\tau, C_j, \dots) \in \Omega_{C_j}$. Clearly, $\mu(\tau) = C_j$ for every $\mu \in S(Q_{-\{\tau, C_j\}}, Q'_{\tau}, Q'_{C_j})$. But this implies that Q is not rematching proof. This is a contradiction.

Step 2. Q^* is rematching proof.

Suppose on the contrary that Q^* is not rematching proof. Then there exists $(S_i, C_j) \in S \times C$, with $(Q'_{S_i}, Q'_{C_j}) \in \Omega_{S_i} \times \Omega_{C_j}$, such that

$$\varphi_{S_i}(\tilde{Q}^*) \ P_{S_i} \ \varphi_{S_i}(Q^*) = \varphi_{S_i}(Q) \quad (3)$$

$$\varphi_{C_j}(\tilde{Q}^*) \ \bar{P}_{C_j} \ \varphi_{C_j}(Q^*) = \varphi_{C_j}(Q) \quad (4)$$

where $\tilde{Q}^* = (Q^*_{-\{S_i, C_j\}}, Q'_{S_i}, Q'_{C_j})$. We consider three cases.

Case a. $\varphi_{C_j}(Q) \subseteq S$.

By responsiveness of \bar{P}_{C_j} and (4), $\exists \sigma \in \varphi_{C_j}(\tilde{Q}^*)$ and $\tau \in \varphi_{C_j}(Q^*)$ such that $\sigma \notin \varphi_{C_j}(Q^*)$ and $\sigma P_{C_j} \tau$. It follows from the construction of $Q^*_{C_j}$ that $\sigma Q_{C_j} \tau$. If $S_i = \sigma$, then $C_j P_{S_i} \varphi_{S_i}(Q^*) = \varphi_{S_i}(Q)$. We consider two cases: (i) $\varphi_{S_i}(Q^*) \neq S_i$ and (ii) $\varphi_{S_i}(Q^*) = S_i$. In case (i) we have $C_j Q_{S_i} \varphi_{S_i}(Q)$ and then (S_i, C_j) blocks $\varphi(Q)$ with respect to Q , a contradiction. In case (ii) let $\hat{Q}_{S_i} = (C_j, \dots) \in \Omega_{S_i}$ and then $\mu(S_i) = C_j$ for every $\mu \in S(Q^*_{-S_i}, \hat{Q}_{S_i})$ (see Δ), which contradicts to the fact that Q^* is a Nash equilibrium. This shows that $\sigma \neq S_i$.

But if $\sigma \neq S_i$, then either (i) $C_j Q_{\sigma} \varphi_{\sigma}(Q^*)$ if $\varphi_{\sigma}(Q) \neq \sigma$ or (ii) $C_j Q_{\sigma} \sigma$ if $\varphi_{\sigma}(Q) = \sigma$. (i) follows from (3), the assumption $\varphi_{\sigma}(Q) \neq \sigma$ and the construction of Q^*_{σ} . Either case implies that $C_j Q_{\sigma} \varphi_{\sigma}(Q)$. This implies that (σ, C_j) blocks $\varphi(Q)$ with respect to Q .

Case b. $\varphi_{C_j}(Q^*) \cap S \neq \emptyset$ and $\exists C_j \in \varphi_{C_j}(Q^*)$.

First suppose $C_j = \varphi_{S_i}(\tilde{Q}^*)$. We discuss two subcases, (i) $\varphi_{S_i}(Q^*) \neq S_i$ and (ii) $\varphi_{S_i}(Q^*) = S_i$.

In case (i), $C_j P_{S_i} \varphi_{S_i}(Q)$ implies that $C_j Q^*_{S_i} \varphi_{S_i}(Q)$. And then (S_i, C_j) blocks $\varphi(Q^*)$ with respect to Q^* , since $C_j \in \varphi_{C_j}(Q^*)$ and $S_i Q^*_{C_j} C_j$. In case (ii), let $\hat{Q}_{S_i} = (C_j, S_i, \dots)$ and denote $\hat{Q} = (Q_{-S_i}, \hat{Q}_{S_i})$. Again, $\mu(S_i) = C_j$ for every $\mu \in S(\hat{Q})$; see Δ . But this implies that Q is not a Nash equilibrium, a contradiction.

Second, suppose $C_l = \varphi_{S_i}(\tilde{Q}^*)$ and then $C_l \neq C_j$. It follows that $S_i Q^*_{C_l} \tau$ for some $\tau \in \varphi_{C_l}(Q^*)$. Let $\hat{Q}_{S_i} = (C_l, S_i, \dots)$ and consider $\varphi_{S_i}(\hat{Q}^*)$, where $\hat{Q}^* = (Q^*_{-S_i}, \hat{Q}_{S_i})$. If $\varphi_{S_i}(\hat{Q}^*) = C_l$, then Q^* is not a Nash equilibrium. Therefore, $\varphi_{S_i}(\hat{Q}^*) = S_i$. This implies that $\varphi_{C_l}(\hat{Q}^*) \subseteq S$ and $\nu Q^*_{C_l} S_i$ for all students $\nu \in \varphi_{C_l}(\hat{Q}^*)$. Moreover, $\nu Q^*_{C_l} \tau$ for some $\tau \in \varphi_{C_l}(Q^*)$ since $S_i Q^*_{C_l} \tau$. If $C_l Q^*_{\nu} \varphi_{\nu}(Q^*)$, then (ν, C_l) blocks $\varphi(Q^*)$ with respect to Q^* . Therefore, we have that $\varphi_{\nu}(Q^*) Q^*_{\nu} C_l$. By individual rationality and the construction of Q^*_{ν} , we also have that $C_l Q^*_{\nu} \nu$ since $\nu \neq S_i$. But this is impossible because there does not exist any C_l between $\varphi_{\nu}(Q^*)$ and ν in Q^*_{ν} , no matter whether $\varphi_{\nu}(Q^*) = \nu$ or $\varphi_{\nu}(Q^*) \neq \nu$.

Case c. $\varphi_{C_j}(Q) \cap S = \emptyset$.

First, suppose that $C_j = \varphi_{S_i}(\hat{Q})$. We have that $S_i P_{C_j} C_j$ by individual rationality. Let $\hat{Q}_{S_i} = (C_j, S_i, \dots)$ and $\hat{Q}_{C_j} = (S_i, C_j, \dots)$. Denote $\hat{Q} = (Q_{-\{S_i, C_j\}}, \hat{Q}_{S_i}, \hat{Q}_{C_j})$. For any stable matching $\mu \in S(\hat{Q})$, we have $\mu(S_i) = C_j$ and $S_i \in \mu(C_j)$. Therefore, by the responsiveness, we have,

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

and

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

which contradicts to the assumption that Q is rematching proof.

Second, suppose $C_l = \varphi_{S_i}(\hat{Q})$ and $C_l \neq C_j$. Then one can repeat the proof of the last paragraph in *Case b* above. This completes the proof. \square .

We conclude this section by two useful corollaries from Theorems 4 and 7, and Corollary 1. They follow from the fact that a strong equilibrium is rematching proof.

Corollary 2. *A rematching proof equilibrium in truncations always exists. The sets of all rematching proof equilibria in truncations coincide for all stable matching mechanisms.*

Corollary 3. *All strong equilibria in truncations produce stable matchings with respect to the true preference profile P . The sets of all strong equilibria in truncations coincide for all stable matching mechanisms.*

These results show that there is no difference among all stable matching mechanisms in rematching proof or strong equilibria in truncations. This is in the sharp contrast to the situation in which agents report their preferences in a straightforward manner. It is known that the student proposal algorithm (Gale and Shapley (1962), RS) produces a stable matching that is optimal for all students but it is the least preferred stable matching for the college, and the college proposal algorithm produces a stable matching that is optimal for all colleges but it is a matching that is the least preferred by all students. But at a rematching proof or strong equilibrium in truncations one and only stable matching exists and this matching is stable not only for the (mis)reported profile but also for the true profile. All stable matching mechanisms, either the college proposal algorithm or the student proposal algorithm, produce precisely the same stable matching. Which stable matching a stable matching mechanism will produce in a rematching proof or strong equilibrium in truncations completely depends on agents' reported equilibrium strategies not on the specific feature of the stable matching mechanism. These results may provide some instructive

guidances in the design of a stable matching mechanism in practice.

7 Conclusions

When we consider a possible refinement of Nash equilibrium in pure strategies, there are limited choices of the existing equilibrium refinements. One possible direction is to see if the undominated strategy equilibrium may also have the properties as the truncated rematching proof or strong one does. One important example in this direction is the work in Roth (1984b) for the marriage problem and for the Gale and Shapley mechanism. In the general framework, the answer to this direction of search is negative. A truncated equilibrium manipulation in RV for the marriage problem is undominated and an outcome of such an equilibrium may produce unstable matchings, as shown in Example 2. Since the marriage problem is a special situation of the college admissions problem, the search in this direction will be negative either. Therefore, some additional conditions on this class of strategies are necessary for them to produce stable outcomes. Or we have to give up stable mechanisms. Indeed, Alcalde (1994) considered an unstable mechanism which fully implements the stable matching correspondence in undominated strategy equilibria for the marriage problem. Whether this is a prospective direction of investigation for the college admissions problem remains unclear.

A remaining question is whether any Nash equilibrium whose outcome is stable must be rematching proof. The other open question is whether Theorem 7 holds for more complicated preferences beyond the responsiveness. For example, does it hold when colleges' preferences over classes are substitutable²? These two questions are left for the future study.

²See RS for the definition of substitutable preferences (without money) and Kelso and Crawford (1982) for the definition of the gross substitutes condition (with money).

References

- [1] Alcalde, J.: Implementation of Stable Solutions to Marriage Problem, Universitat Autònoma de Barcelona. (1994).
- [2] Aumann, R.J.: Acceptable Points in General Cooperation n-Persons Games, in “Contributions to the Theory of Games IV”, Princeton University Press, Princeton (1959).
- [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, mimeo, University of Rochester. Forthcoming in *Economic Theory* (1994).
- [5] —, —: “Nash Implementation of Matching Rules,” mimeo, University of Rochester. Forthcoming in *Journal of Economic Theory* (1993).
- [6] Kelso, A.S., Jr., Crawford, V.P.: “Job Matching, Coalition Formation, and Gross Substitutions,” *Econometrica*, 50, 1483-1504 (1982).
- [7] Ma, J.: “Strategy Proofness and the Strict Core in a Market with Indivisibilities,” *International Journal of Game Theory*, 23, 75-83 (1984a).
- [8] —: 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 (1994b).
- [9] — : “Stable Matchings and Rematching-proof Equilibria in a Two-sided Matching Market,” *Journal of Economic Theory* 66, 352-369 (1995).
- [10] Mongell, S., Roth, A.: “Sorority Rush as a Two-sided Matching Mechanism,” *American Economic Review*, 81, 441-464 (1991).
- [11] Moore, J.: Implementation in Environments with Complete Information, in “Advances in Economic Theory” edited by J.J. Laffont, Cambridge University Press, New York 1992.
- [12] Roth, A.E.: “The Economics of Matching: Stability and Incentives,” *Mathematics of Operation Research*, 7, 617-628 (1982).
- [13] —: “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).
- [14] —: “ Misrepresentation and Stability in the Marriage Problem,” *Journal of Economic Theory*, 34, 383-387 (1984b).

- [15] —: “The College Admissions Problem is not Equivalent to the Marriage Problem,” *Journal of economic Theory*, 36, 277-88 (1985).
- [16] —: 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 1990.
- [17] —: “A Natural Experiment in the Organization of Entry-Level Labor Markets: regional markets for new physicians and surgeons in the United Kingdom,” *American Economic Review*, 81, 415-440 (1991).
- [18] —, Sotomayor, M.: “The College Admissions Problem Revisited,” *Econometrica*, 57, 559-70 (1989).
- [19] —, —: Two-sided Matching: a Study in Game-theoretic Modeling and Analysis, Cambridge University Press, Cambridge 1990.
- [20] —, Vande Vate, J.H.: “Incentives in Two-sided Matching with Random Stable Mechanisms,” *Economic Theory*, 1, 31-44 (1991).
- [21] —, —: “Random Paths to Stability in Two-sided Matching,” *Econometrica* **58**, 1475-80 (1990).
- [22] —, Xing, X. (1994a): “Jumping the Gun: Imperfections and Institutions Related to the Timing of market Transaction,” *American Economic Review*, 84, 992-1044 (1994a).
- [23] —, —: Turnaround Time and Bottlenecks in Markets Clearing: A Comparative Analysis of Labor Market Institutions, mimeo, University of Pittsburgh (1994b).
- [24] Shin, S., Suh, S-C.: A Mechanism Implementing the Stable Rule in Marriage Problem, University of Windsor. Forthcoming in *Economics Letters* (1994).
- [25] Sönmez, T.: Games of Manipulation in Marriage Problem, mimeo, University of Michigan (1996).
- [26] Zhou, L.: “Stable Matchings and Equilibrium Outcomes of the Gale-Shapley Algorithm for the Marriage Problem,” *Economics Letters*, 36, 25-29 (1991).