Games and Economic Behavior **33**, 206–230 (2000)
doi:10.1006/game.1999.0779, available online at http://www.idealibrary.com on IDEAL®

On the Existence of Stable Roommate Matchings

Kim-Sau Chung*

Department of Economics, Northwestern University, 2003 Sheridan Road, Evanston, Illinois 60208-2600
E-mail: sau@nwu.edu

Received November 20, 1997

This paper identifies a condition called "no odd rings" that is sufficient for the existence of stable roommate matchings in the weak preferences case. It shows that the process of allowing randomly chosen blocking pairs to match converges to a stable roommate matching with probability one as long as there are no odd rings. This random-paths-to-stability result generalizes that of Roth and Vande Vate (1990, *Econometrica* **58**, 1475–1480) and may not hold if there are odd rings. The "no odd rings" condition can also be used to prove a number of other sufficient conditions that are more economically interpretable. *Journal of Economic Literature* Classification Numbers: C78, D71.

1. INTRODUCTION

The marriage problem can be viewed as a special case of the roommates problem with specific restrictions on preferences. Gale and Shapley (1962) showed that stable matchings may not exist in the roommates problem, but always exist in the marriage problem. This raises the question of why the marriage problem stands out as a special case of the roommates problem that always admits stable matchings.

In the special case where agents' preferences are *strict*, this puzzle has been fully solved by Tan (1991), who identified a necessary and sufficient condition, stated in the form of preference restriction, for the existence of stable roommate matchings. However, in the general case where agents'

* This is the first chapter of my Ph.D. thesis at University of Wisconsin-Madison. I thank my advisor Steven Durlauf for his encouragement. I also thank Yeon-Koo Che, Brian Krauth, Larry Samuelson, an associate editor, three anonymous referees, and seminar participants at Universities of Pittsburgh, Toronto, Western Ontario, and Wisconsin-Madison for comments. I am especially grateful to Hernán Abeledo for pointing me to the references of Tan (1991) and Abeledo and Rothblum (1994), and to Alvin Roth for pointing out a mistake in my earlier proof of Lemma 1.

0899-8256/00 $35.00

preferences are allowed to be *weak*, no existence theorems have yet been provided in the literature. This paper identifies a condition called "no odd rings" that is sufficient for the existence of stable roommate matchings in the weak preferences case. Since the condition is always satisfied by the marriage problem, it explains why the marriage problem always admits stable matchings. This result complements that of Tan, whose necessary and sufficient condition can only explain why the marriage problem always admits stable matchings in the strict preferences case.

The "no odd rings" condition is important not only in the sense that it guarantees the existence of stable roommate matchings, but also in the sense that it guarantees the convergence of the Roth–Vande Vate (1990) process to a stable roommate matching with probability one. The Roth–Vande Vate process is a decentralized decision making process of allowing randomly chosen blocking pairs to match, and can be thought of as an approximation to the real life dynamics in a decentralized "matching economy." Whether the Roth–Vande Vate process can converge to a stable roommate matching is crucial to whether stability itself is a good solution concept. If there are odd rings, even when stable roommate matchings exist, the Roth–Vande Vate process may not converge to any one of them.

The "no odd rings" condition is quite abstract, and may not be interpretable in economic terms. However, it is easy to use the "no odd rings" condition to prove the sufficiency of a number of more economically interpretable conditions. We shall give eight examples in this paper. Each of these examples represents one preference restriction that implies the absence of odd rings. Whether these preference restrictions are useful to applied matching theorists will depend on the particular "matching economies" under study.

The structure of this paper is as follows. Section 2 introduces the roommates and the marriage problems. Section 3 states the "no odd rings" condition and relates it to previous existence theorems in the strict preferences case. Section 4 proves the random-paths-to-stability result, which in turn also provides a proof for the sufficiency of the "no odd rings" condition. Section 5 uses the "no odd rings" condition to prove a number of other sufficient conditions that are more economically interpretable. Section 6 relates the roommates problem to other matching problems.

2. THE ROOMMATES AND THE MARRIAGE PROBLEMS

Let I denote the set of agents, with a, b, c, d, e, w, x, y, and z typical agents in I. Every agent has a preference ordering over I. As usual, we assume the preference orderings are complete, reflexive, and transitive. If x weakly prefers y to z, we write $y \succcurlyeq_x z$. If x strictly prefers y to z, we write $y \succ_x z$. If x is indifferent between y and z, we write $y \sim_x z$. If x weakly prefers y to himself (i.e., $y \succcurlyeq_x x$), x weakly prefers having y as roommate to living alone. If x strictly prefers living alone to y (i.e., $x \succ_x y$), y is "unacceptable" to x as a roommate. The collection of all agents' preferences, $\succcurlyeq$ (i.e., $(\succcurlyeq_x)_{x \in I}$), is called the "preference profile."

A "matching" is an assignment of roommates. We consider assignment rules such that every agent is assigned at most one roommate. If an agent is not assigned any roommate, he lives alone, and we say that he is "single."

Definition 1. A "matching" μ is a one-to-one mapping from I onto itself such that for all $\{x, y\} \subset I$, $\mu(x) = y$ if and only if $\mu(y) = x$. If $\mu(x) = x$, x is "single" under matching μ.

A matching satisfies "individual rationality" (hereafter IR) if no agent is assigned a roommate not acceptable to him. Two agents $\{x, y\}$ "block" a matching μ if they are not assigned as roommates of each other under μ, but both strictly prefer each other to the roommates assigned to them under μ. Intuitively, when this happens, x and y will both have (strictly positive) incentive to disobey matching μ and move in with each other instead.

Definition 2. $\{x, y\} \subset I$, $x \neq y$, "block" μ if $y \succ_x \mu(x)$ and $x \succ_y \mu(y)$.

We are interested in whether there exists any matching such that no agent is assigned an unacceptable roommate, and no pair of agents can block the matching. If such a matching exists, we say that it is "stable."

An instance of "the roommates problem" is completely defined by the collection of agents I and the preference profile $\succcurlyeq$. A related, but much more fully studied, class of problems in the literature is "the marriage problem." In "the marriage problem," agents can be divided into two types, usually called men and women. An agent of one type can only be matched to an agent of another type, or can remain single. However, if we consider only stable matchings, this restriction on matching rule has the same effect on the set of stable matchings as does a restriction on preferences such that every agent finds all other agents of the same sex unacceptable.

Definition 3. An instance of the roommates problem is an instance of the marriage problem if I can be partitioned into I_1 and I_2, both being non-empty, such that for any $x, y \in I_1$ (respectively, I_2) such that $x \neq y$, $x \succ_x y$ and $y \succ_y x$.

Gale and Shapley (1962) showed that stable matchings may not exist in the roommates problem. An example is as follows.

EXAMPLE 1. Consider $I = \{x, y, z\}$, and the following preferences:

$$y \succ_x z \succ_x x,$$
$$z \succ_y x \succ_y y,$$
$$x \succ_z y \succ_z z.$$

There are no stable matchings for this preference profile. To see this, suppose x and y are roommates. Then z can propose to move in with y and kick away x. The same is true if we start with any other pair of roommates.

In contrast to the possible non-existence of stable roommate matchings, Gale and Shapley (1962) proved that stable matchings always exist in the marriage problem.[1] This raises the question of why the marriage problem stands out as a special case of the roommates problem that always admits stable matchings, and what more general preference restrictions can guarantee the existence of stable roommate matchings. The next section provides an answer to these two questions.

3. THE EXISTENCE OF STABLE ROOMMATE MATCHINGS

Before we introduce the "no odd rings" condition, we need one more definition.

DEFINITION 4. A "ring" is an ordered subset of agents $(x_1, x_2, \ldots, x_k)$, $k \geq 3$, such that (subscript modulo k)

$$x_{i+1} \succ_{x_i} x_{i-1} \succcurlyeq_{x_i} x_i, \qquad \text{for odd } i, 1 \leq i \leq k,$$
$$x_{i+1} \succcurlyeq_{x_i} x_{i-1} \succcurlyeq_{x_i} x_i, \qquad \text{for even } i, 1 < i \leq k.$$

DEFINITION 5. An "odd ring" is a ring $(x_1, x_2, \ldots, x_k)$ such that k is odd.

The sufficiency of the "no odd rings" condition can now be formally stated as the following theorem. The proof of the theorem follows immediately from Lemma 1 in Section 4.

[1]Gale and Shapley only considered strict preferences. But it is well known that their existence result still holds in the weak preferences case provided their definition of "stability" is generalized as in the text. Notice that there are other ways to generalize their definition to the weak preferences case. For example, $\{x, y\} \subset I$, $x \neq y$, can be said to block μ as long as $y \succ_x \mu(x)$ and $x \succcurlyeq_y \mu(y)$ (i.e., we can relax the requirement that both strictly prefer each other to their assigned roommates). But then the existence result of Gale and Shapley will no longer hold in the weak preferences case.

THEOREM 1. *If the preference profile has no odd rings, there exist stable roommate matchings.*

The existence of stable matchings in the marriage problem follows immediately from the above theorem, as all possible rings have an even number of agents by construction.

COROLLARY 1 (Gale and Shapley). *The marriage problem always has stable matchings.*

Proof. Suppose an odd ring $(x_1, x_2, \ldots, x_k)$ exists. W.l.o.g. assume x_1 is a man. Then x_2 is a woman, x_3 is a man, ..., x_k is a man (because k is odd), and x_1 is a woman, a contradiction. Q.E.D.

The converse of Theorem 1 is not true, and hence the "no odd rings" condition is not necessary for the existence of stable roommate matchings. This can be seen from the following example.

EXAMPLE 2. Consider $I = \{w, x, y, z\}$, and the following preferences:

$$x \succ_w w \succ_w \cdots,$$

$$w \succ_x y \succ_x z \succ_x x,$$

$$z \succ_y x \succ_y y \succ_y w,$$

$$x \succ_z y \succ_z z \succ_z w.$$

An odd ring exists: (x, y, z). But yet there exists a stable roommate matching μ such that $\mu(w, y) = (x, z)$.[2]

A large part of the existing literature on matching theory deals with strict preferences. Hence, in the rest of this section, we shall relate Theorem 1 to previous existence theorems in the strict preferences case. Notice that when preferences are strict, there is no need to distinguish an x_i with odd i from an x_i with even i in Definition 4. To avoid confusion, we shall call a ring adapted to the strict preferences case a "strict ring," and a strict ring with an odd number of agents a "strict odd ring." The following corollary follows trivially from Theorem 1.

COROLLARY 2. *If the preference profile is strict and has no strict odd rings, there exist stable roommate matchings.*

Corollary 2 is not new to the literature. It can also be derived from Tan's (1991) necessary and sufficient condition for the existence of stable roommate matchings, or from Abeledo and Rothblum's (1994) characterization of the extreme points of the fractional stable matching polytope.

[2] $\mu(w, y) = (x, z)$ is the shorthand for $\mu(w) = x$ and $\mu(y) = z$. Throughout this paper, we shall adopt the same convention.

In the special case of strict preferences, Tan (1991) proved that every instance of the roommates problem has at least one "stable partition." A stable partition is a partition of the agent set, $I = \cup A$, such that (1) every partition set, $A = \{x_1, \ldots, x_k\}$, is a singleton, is a pair of mutually acceptable agents, or corresponds to a strict ring, $(x_1, \ldots, x_k)$, and (2) for any $x_i \in A$ and $y_j \in A'$ (without ruling out $A = A'$) such that $y_j \neq x_{i+1}$, $y_j \succ_{x_i} x_{i-1} \Longrightarrow y_{j-1} \succ_{y_j} x_i$. Tan's necessary and sufficient condition states that, when agents' preferences are strict, stable roommate matchings exist if and only if every partition set in every stable partition either is a singleton or has an even number of agents. This result implies Corollary 2, because if the strict preference profile has no strict odd rings, every partition set in every stable partition will either be a singleton or have an even number of agents.

Abeledo and Rothblum (1994) found that, in the special case of strict preferences, every instance of the roommates problem corresponds to a nonempty set called the "fractional stable matching polytope." They proved that every integral extreme point of this polytope corresponds to a stable roommate matching, and every non-integral extreme point corresponds to a strict odd ring in the preference profile. This result also implies Corollary 2, because if the strict preference profile has no strict odd rings, the fractional stable matching polytope will have at least one integral extreme point (because it is non-empty), and hence there will be at least one stable roommate matching.

4. RANDOM PATHS TO STABLE MATCHINGS

This section serves two interrelated purposes. The first purpose is to prove that, when there are no odd rings, a decentralized decision making process known as the Roth–Vande Vate process will converge to a stable roommate matching with probability one. This lends support to the focus on stable roommate matchings as a solution concept in the roommates problem. Since if there are no stable roommate matchings, it would not have made sense to say that some process converges to a stable roommate matching at the very beginning, our random-paths-to-stability result also provides a proof to the sufficiency of the "no odd rings" condition for the existence of stable roommate matchings—this is the second purpose of this section.

One of the justifications for using stability as a solution concept in matching theory is that many decentralized decision making processes are likely to converge to a stable matching. A formal version of this statement was first provided and proved by Roth and Vande Vate (1990) in the context of the marriage problem. They proved that, in the special case of strict preferences, the decentralized decision making process of allowing randomly chosen blocking pairs to match will converge to a stable marriage with probability one. It remains unanswered whether this random-paths-to-stability result still holds in the context of the roommates problem with weak preferences. Two difficulties immediately surface. The first is that stable roommate matchings may not exist, and hence there is nowhere the Roth–Vande Vate process can converge to. The second is that even when stable roommate matchings exist, the Roth–Vande Vate process may be trapped in an endless cycle and never arrive at a stable roommate matching. This second difficulty can be illustrated with a very simple example.

EXAMPLE 3. Consider $I = \{w, x, y, z\}$, and the following preferences:

$$x \sim_w w \succ_w \cdots,$$

$$w \succ_x y \succ_x z \succ_x x,$$

$$z \succ_y x \succ_y y \succ_y w,$$

$$x \succ_z y \succ_z z \succ_z w.$$

In the above example, there exists a unique stable matching ν, where $\nu(w, y) = (x, z)$. However, starting from the unstable IR matching μ, where $\forall a \in \{w, x, y, z\}$, $\mu(a) = a$, the Roth–Vande Vate process will never arrive at ν.

However, it can be proved that whenever the preference profile has no odd rings, the Roth–Vande Vate process will still converge to a stable roommate matching with probability one. Since preference profiles in the marriage problem never contain odd rings (see Section 3), our result nests that of Roth and Vande Vate as a special case. Before we prove this result, we need one more definition.

DEFINITION 6. Suppose $\{x, y\} \subset I$ block a matching μ. A new matching ν is obtained from μ by "satisfying the blocking pair" $\{x, y\}$ if (1) $y = \nu(x)$ and $x = \nu(y)$, (2) $x \neq z = \mu(x)$ implies $\nu(z) = z$, and $y \neq z = \mu(y)$ implies $\nu(z) = z$, and (3) $\forall z \in I \setminus \{x, y, \mu(x), \mu(y)\}$, $\nu(z) = \mu(z)$.

LEMMA 1. *If the preference profile has no odd rings, then for any initial IR matching μ, there exists a finite sequence of matchings $(\mu_1, \ldots, \mu_k)$, such that $\mu_1 = \mu$, μ_k is stable, and for each $i = 1, \ldots, k-1$, there is a blocking pair for μ_i such that μ_{i+1} is obtained from μ_i by satisfying that blocking pair.*

Proof. For any initial IR matching μ, select a subset $S \subset I$ of agents such that there are no blocking pairs for μ contained in S, and μ does not match any agent in S to any agent not in S. (For example, S could be a pair of agents matched under μ or a single agent.) W.l.o.g. assume that there exists a blocking pair $\{x_1, x\}$ for μ such that $x_1 \in I \setminus S$ and $x \in S$. We shall prove that there exists a finite sequence of matchings, with the property described in the Lemma, which ends with a new IR matching μ' such that there are no blocking pairs for μ' contained in $S' \equiv S \cup \{x_1\}$, and μ' does not match any agent in S' to any agent not in S'. Once we have proved this, we can then repeat this process until $S' = I$. The corresponding μ' will then be a stable matching defined on I, and the proof will be done.

Define $\mu_1 \equiv \mu$. Construct μ_2 by satisfying the blocking pair $\{x_1, x_2\}$, with $x_2 \in S$, most preferred by x_1. If there is an agent $x_3 = \mu_1(x_2)$, he will be left single under this new matching. Define $\mu_3 \equiv \mu_2$.[3] There may now be a blocking pair $\{x_3, x_4\}$, with $x_4 \in S'$, for μ_3. If so, construct μ_4 by satisfying the blocking pair $\{x_3, x_4\}$, with $x_4 \in S'$, most preferred by x_3. If there is an agent $x_5 = \mu_3(x_4)$, he will be left single under this new matching. Define $\mu_5 \equiv \mu_4 \ldots$ The process continues in this way within S', and generates a sequence of affected agents $(x_1, x_2, x_3, \ldots)$ and a sequence of corresponding matchings $(\mu_1, \mu_2, \mu_3, \ldots)$. We make three observations immediately. First, for any $n \geq 1$, μ_n is IR. Second, for any odd $n \geq 3$, x_n is left single under matchings μ_n $(= \mu_{n-1})$ because her original roommate, $\mu_{n-2}(x_n)$ $(= x_{n-1})$, breaks up with her and moves in with x_{n-2}. Third, for any even $n \geq 2$, $x_{n-1} \succ_{x_n} x_{n+1}$ because $\{x_{n-1}, x_n\}$ is a blocking pair for μ_{n-1}.

We shall label the agent x_n as a man if n is even, and as a woman if n is odd. This labeling may be inconsistent if some agent appears more than once in the sequence, and appears sometimes at an even location and sometimes at an odd location. We say that the agent changes sex if such a situation arises. We shall prove that no agent changes sex throughout the sequence.

Suppose not. Let time j be the first time we ever see any agent change sex. Therefore the first agent who changes sex is x_j, and this agent has appeared at least once in the sequence before time j. Let the last time this agent appears in the sequence before time j be time i. Therefore x_i is the same agent as x_j, and this agent does not appear in the sequence between time i and j. Moreover, by the fact that this agent changes sex, $j - i$ is odd. It can be shown that i is odd and j is even.[4]

[3]The reason we sometimes want to give two names to a single matching is to simplify the proof.

[4]Suppose i is even and j is odd. Then $\{x_{i-1}, x_i\}$ is a blocking pair for μ_{i-1} by construction. Moreover, $x_{i-1} = \mu_k(x_i)$ from $k = i$ to $k = j - 2$ (because x_i does not appear in the sequence between time i and j). At time $j - 2$, $\{x_{j-2}, x_{j-1}\}$ form a blocking pair for μ_{j-2}, resulting in matching μ_j $(= \mu_{j-1})$ under which x_j is left single. Therefore x_{j-1} and x_j were originally roommates under μ_{j-2}. We hence conclude that $x_{i-1} = \mu_i(x_i) = \mu_{j-2}(x_i) = \mu_{j-2}(x_j) = x_{j-1}$. But then, since $(j - 1) - (i - 1) = j - i$ is odd, x_{j-1} (instead of x_j) is the first agent who changes sex, a contradiction.

Now the chain $(x_{i+1}, \dots, x_j)$ has an odd number of agents. If no agents appear in the chain more than once, the chain will then be an odd ring. To see this, first notice that $x_{j-1} \succ_{x_j} x_{i+1}$, because x_j breaks away from $\mu_{j-1}(x_j)$ $(=\mu_{i+1}(x_j) = \mu_{i+1}(x_i) = x_{i+1})$ and moves in with x_{j-1}. Second, recall that for any even n such that $i+1 \le n \le j-2$, $x_{n-1} \succ_{x_n} x_{n+1}$. Therefore it suffices to verify that for any odd n such that $i+2 \le n \le j-1$, $x_{n-1} \succcurlyeq_{x_n} x_{n+1}$. This is true for $n = j-1$, otherwise $\{x_i, x_{j-1}\}$ would have been a more preferable (from x_i's point of view) blocking pair for μ_i than $\{x_i, x_{i+1}\}$ is. For any odd n such that $i+2 \le n \le j-3$, since x_{n+1} strictly prefers x_n to x_{n+2}, and strictly prefers x_{n+2} to any of his previous roommates,[5] we have $x_{n-1} \succcurlyeq_{x_n} x_{n+1}$ both if it is the first time x_n appears in the sequence (otherwise $\{x_n, x_{n+1}\}$ would have blocked μ) and if not (otherwise x_n would have proposed to x_{n+1} instead of x_{n-1} the last time when she proposed).

Since we assume no odd rings exist, there are some agents who appear more than once in the chain $(x_{i+1}, \dots, x_j)$. Our goal is to remove some chunks from the chain $(x_{i+1}, \dots, x_j)$ so that the remaining part forms an odd ring. To do this, we move from x_j "backward" until we hit the first agent who appears more than once in the chain. Suppose x_m is this agent's first appearance in the chain and x_n is his last one, where $i < m < n < j$. We shall remove the subchain $(x_{m+1}, \dots, x_n)$ from the chain $(x_{i+1}, \dots, x_j)$. Since x_n has not changed sex, the subchain $(x_{m+1}, \dots, x_n)$ has an even number of agents. Hence the remaining part of the chain still has an odd number of agents. By successively removing subchains in similar manner, the remaining part of the chain will eventually contain no agents who appear more than once, and yet has an odd number of agents.[6]

We shall prove that the remaining part of the chain always forms an odd ring. Still denote a typical removed subchain by $(x_{m+1}, \dots, x_n)$. First observe that both m and n are odd.[7] It means x_{n+1} strictly prefers x_m $(=x_n)$

[5]It is because x_{n+1} has always been a man, and men always receive proposals from women and switch to better roommates.

[6]For example, if the original chain is

$$(x_4, x_5, x_6, x_7, x_8, x_9, x_6, x_5, x_{12}, x_{13}, x_{14}, x_{15}, x_{12}, x_5, x_{18}, x_{19}, x_{14}, x_{13}, x_{22}, x_{23}, x_3),$$

then $i = 3$ and $j = 24$, and x_3 is the first agent who changes sex. In the first round we shall remove the subchain $(x_{14}, x_{15}, x_{12}, x_5, x_{18}, x_{19}, x_{14}, x_{13})$. In the second round we shall remove the subchain $(x_6, x_7, x_8, x_9, x_6, x_5)$. Therefore the remaining part of the chain is $(x_4, x_5, x_{12}, x_{13}, x_{22}, x_{23}, x_3)$ and has an odd number of agents.

[7]W.l.o.g. assume $(x_{m+1}, \dots, x_n)$ is the first subchain to be removed. Suppose n is even. Then x_{n+1} $(= \mu_{n-1}(x_n))$, instead of x_n, would have been the first agent who appears more than once in the chain when we move from x_j "backward." To see this, recall that x_n appears more than once in the chain. W.l.o.g. assume x_m to be x_n's only other appearance in the chain. Then m is even (because x_n does not change sex), and $x_{m-1} = \mu_k(x_m)$ from $k = m$ to $k = n-1$. Therefore we have $x_{m-1} = \mu_{n-1}(x_m) = \mu_{n-1}(x_n) = x_{n+1}$.

to any of his previous roommates. We then have $x_{m+1} \succcurlyeq_{x_m} x_{n+1}$, otherwise $\{x_m, x_{n+1}\}$ would have been a more preferable (from x_m's point of view) blocking pair for μ_m than $\{x_m, x_{m+1}\}$ is. Combining this with $x_{m-1} \succcurlyeq_{x_m} x_{m+1}$ (which can be proved using the same arguments in the paragraph two paragraphs above), we have $x_{m-1} \succcurlyeq_{x_m} x_{n+1}$. The rest of this part of the proof resembles the paragraph two paragraphs above.

We have proved that if any agent changes sex along the sequence, there will exist an odd ring. Since we assume no odd rings exist, no agent changes sex throughout the sequence. Suppose the sequence never terminates. Then there is at least one agent who appears in the sequence as a man infinitely often. But every time he appears in the sequence, he switches to a strictly more preferable roommate, which is impossible as there are only finitely many agents in S'. Hence the sequence terminates after a finite number of steps. Denote the terminal sequence by $(x_1, x_2, \ldots, x_n)$ and the terminal matching by μ_n. Apparently μ_n is IR, and does not match any agent in S' to any agent not in S'. It remains to prove that there are no blocking pairs for μ_n contained in S'.

So far we have partitioned S' into three groups of agents: men, women, and those who do not appear in $(x_1, x_2, \ldots, x_n)$. Notice that for any $x \in S$ not in $(x_1, x_2, \ldots, x_n)$, $\mu_n(x) = \mu(x)$. Suppose $\{y, z\} \subset S'$ block μ_n. Then at least one of them appears in $(x_1, x_2, \ldots, x_n)$, otherwise they would have been a blocking pair for μ. Suppose y is a woman but z is not. Then y would have proposed to z instead of $\mu_n(y)$ the last time when she proposed, a contradiction. Suppose both y and z are women, then one of them would have proposed to the other instead of her roommate under μ_n the last time when she proposed, a contradiction. Therefore neither y nor z can be a woman. Suppose y is a man. Then $z \succ_y \mu_n(y) \succcurlyeq_y \mu(y)$. Now, no matter whether z appears in $(x_1, x_2, \ldots, x_n)$ or not, $\{y, z\}$ would have been a blocking pair against μ, a contradiction.[8]

Therefore there are no blocking pairs for μ_n contained in S'. We can now set $\mu' \equiv \mu_n$, and the proof is completed. Q.E.D.

Our random-paths-to-stability result now follows immediately from the standard Markov-chain argument.

THEOREM 2. *If the preference profile has no odd rings, then for any initial IR matching μ, the decentralized decision making process of allowing randomly chosen blocking pairs to match will converge to a stable roommate matching with probability one.*

[8]Recall that z is not a woman, and hence $x_1 \neq z \in S$.

5. OTHER APPLICATIONS

We shall further demonstrate the usefulness of Theorem 1 in this section by proving a number of other sufficient conditions that are more economically interpretable. But before we move on, we shall first introduce an important concept called "tie-breaking versions."

DEFINITION 7. For any two preference profiles $\succcurlyeq$ and $\succcurlyeq'$, $\succcurlyeq'$ is a "tie-breaking version" of $\succcurlyeq$ if

$$\forall x, y, z \in I, \qquad x \succ_y z \Longrightarrow x \succ'_y z.$$

Suppose μ is a stable roommate matching for the preference profile $\succcurlyeq'$. If we construct another preference profile $\succcurlyeq$ by changing some of the strict preference relations in $\succcurlyeq'$ to indifference (i.e., from $y \succ'_x z$ to $y \sim_x z$, for some x, y, and z), clearly μ will still be a stable roommate matching for the new preference profile $\succcurlyeq$, as no new blocking pair will arise from these changes. However, the original preference profile $\succcurlyeq'$ is simply a tie-breaking version of the new preference profile $\succcurlyeq$. Hence we have the following useful fact: *a preference profile admits stable roommate matchings if one of its tie-breaking versions admits stable roommate matchings*.[9] In the subsequent proofs of various sufficient conditions for the existence of stable roommate matchings, we shall either directly prove that the preference profile is odd-ring-free or we shall make use of the above fact and prove that there exist some tie-breaking versions of the original preference profile that are odd-ring-free.

Our first application of Theorem 1 is to reprove a sufficient condition identified by Bartholdi and Trick (1986). Bartholdi and Trick discussed a restriction on the preferences such that agents' traits can be represented by points in a metric space, and every agent strictly prefers agents who have traits similar to his to those who have more different traits, and strictly prefers having a roommate to not. They showed that this restriction implies the existence of stable roommate matchings. Their result is an immediate consequence of Theorem 1, because their preference restriction rules out all kinds of rings.

THEOREM 3 (Bartholdi and Trick). *If agents can be represented by points in a metric space, and every agent strictly prefers agents closer to him to those farther away, and strictly prefers having a roommate to not, there exist stable roommate matchings.*

[9]The converse of this observation is also true, but we will not make use of that in this paper.

Proof. Let $d(i, j)$ denote the distance between agents i and j in the metric space. Suppose a ring $(x_1, x_2, \ldots, x_k)$ exists. Then, by the definition of a ring, we have:

$$d(x_k, x_1) > d(x_1, x_2) \geq d(x_2, x_3) > \cdots \geq d(x_k, x_1),$$

where the last inequality is strict if k is odd. Since $k \geq 3$, this is impossible. Q.E.D.

Our second application deals with "unisexual marriage" within the framework of Becker (1973). In his now classic paper, Becker analyzed marriage within a "household production" framework. People marry because some commodities (one of Becker's examples is "own childs") can be more efficiently produced by a man and a woman living together. Extending Becker's framework to the unisexual case, a household production function is a mapping Z from $I \times I$ to the real line, such that $Z(x, y)$ $(=Z(y, x))$ is the output of the household $\{x, y\}$, and $Z(x, x)$ is the output of agent x if he remains single. Becker considered two ways to divide the household output between the two agents within the household. We shall consider the one which Becker called "complete rigidity in the division of output." Agents are assumed to maximize their consumption of household output when they decide on whether or not to marry and to whom they should marry.

In the case of "complete rigidity in the division of output," every agent x is associated with a strictly increasing function f_x, such that x's share of household output is $f_x(Z(x, \cdot))$. In general, for any $\{x, y\} \subset I$, $f_x(Z(x, y)) + f_y(Z(x, y))$ can be larger (because some household output is consumed nonexclusively) or smaller (because the division of output may incur some costs) than $Z(x, y)$. We shall call the roommates problem induced from such a household output division rule the Beckerian unisexual marriage market. An immediate consequence of Theorem 1 is that the Beckerian unisexual marriage market always has stable matchings because it rules out all kinds of rings.

THEOREM 4. *The Beckerian unisexual marriage market always has stable matchings.*

Proof. Suppose a ring $(x_1, x_2, \ldots, x_k)$ exists. Then, by the definition of a ring, we have $f_{x_1}(Z(x_1, x_k)) < f_{x_1}(Z(x_1, x_2))$, $f_{x_2}(Z(x_1, x_2)) \leq f_{x_2}(Z(x_2, x_3)), \ldots, f_{x_k}(Z(x_{k-1}, x_k)) \leq f_{x_k}(Z(x_1, x_k))$, where the last inequality is strict if k is odd. But, since the f_x's are strictly increasing, we have:

$$Z(x_1, x_k) < Z(x_1, x_2) \leq Z(x_2, x_3) < \cdots \leq Z(x_1, x_k),$$

where the last inequality is strict if k is odd. Since $k \geq 3$, this is impossible. Q.E.D.

The study of preference restrictions is by no means new to social choice theorists. It is known that if agents' preference profiles are unrestricted, any social welfare function which satisfies a small set of reasonable axioms cannot always be transitive. Economists are especially concerned about the simple majority decision rule (hereafter SMDR), which is the only pairwise group decision rule that simultaneously satisfies the characteristics of "decisiveness," "anonymity," "neutrality," and "positive responsiveness" (May, 1952, 1953). Under what kind of preference restrictions will SMDR be always transitive?

In a contribution to the social choice theory, Inada (1964) identified four sufficient conditions for this. The first condition says that SMDR will always be transitive if the preference file is Dichotomous. We adapt the definition of a Dichotomous preference profile to the roommates problem as follows.

DEFINITION 8. A preference profile is "Dichotomous" if every agent classifies all agents into two groups in such a way that he is indifferent among agents in each group.

Notice that the classification of agents may differ from agent to agent. It can be proved that Dichotomous preference profiles rule out odd rings and hence admit stable matchings.

THEOREM 5. *If the preference profile is Dichotomous, there exist stable roommate matchings.*

Proof. Let the bijection $l : I \longrightarrow \{1, 2, \ldots, |I|\}$ be an enumeration of the agents in I, where $|I|$ is the total number of agents in I. Consider a tie-breaking version of the original preference profile such that every agent w finds all agents in his own group unacceptable, and ranks the agents in the other group in such a way that $x \succ_w y$ if and only if $l(x) < l(y)$. Suppose an odd ring $(x_1, x_2, \ldots, x_k)$ exists in this tie-breaking version. Then we have $l(x_k) > l(x_2) > l(x_4) > \cdots > l(x_{k-1}) > l(x_1) \geq l(x_3) \geq \cdots \geq l(x_{k-2}) \geq l(x_k)$. Since $k \geq 3$, this is impossible. Q.E.D.

The second and the third conditions Inada (1964) identified say that SMDR will always be transitive if the preference profile is either Single-Peaked or Single-Caved and the number of agents is odd.[10] Both of these preference restrictions require that the alternatives be lined up in a single line. This idea can be formalized as follows.

[10]Single-Peakedness was actually first identified by Arrow (1951). Bartholdi and Trick (1986) also proved that a slightly different version of Single-Peakedness implies the existence of stable roommate matchings.

DEFINITION 9. The relation $\gg$ on I is a "strong ordering" relation if (1) $\forall x \in I$, not $x \gg x$, (2) $\forall x \neq y \in I$, either $x \gg y$ or $y \gg x$, and (3) $\forall x, y, z \in I$, $x \gg y$ and $y \gg z$ together imply $x \gg z$.

Let $B(x, y, z)$ mean "y is between x and z" with respect to a fixed strong ordering $\gg$.

DEFINITION 10. If $\gg$ is a strong ordering relation on I, define $B(x, y, z)$ to mean that either $x \gg y$ and $y \gg z$, or $z \gg y$ and $y \gg x$.

If we line up the agents in I according to the strong ordering $\gg$, and plot a typical agent's utility against his possible partners, the graph of the utility function will have a single peak in the Single-Peaked case, and a single cave in the Single-Caved case. We adapt the definitions of Single-Peakedness and Single-Cavedness to the roommates problem as follows.

DEFINITION 11. A preference profile is "Single-Peaked" if there exists a strong ordering $\gg$ such that for all $w \in I$ and for all ordered triple (x, y, z), where $\{x, y, z\} \subset I$, $x \succcurlyeq_w y$ and $B(x, y, z)$ together imply $y \succ_w z$.

DEFINITION 12. A preference profile is "Single-Caved" if there exists a strong ordering $\gg$ such that for all $w \in I$ and for all ordered triple (x, y, z), where $\{x, y, z\} \subset I$, $y \succcurlyeq_w z$ and $B(x, y, z)$ together imply $x \succ_w y$.

To understand how a Single-Peaked preference profile will arise, imagine the case that the agents can be ranked according to a one-dimensional scale called "intelligence." There are two types of agents. An "envious" agent prefers to live with a roommate no smarter than he is. An "elitist" agent prefers to live with a roommate no dumber than he is. If, furthermore, every agent has an ideal choice of roommate, and prefers a roommate whose "smartness" is closer to his ideal to one whose smartness is farther away, then their preferences can be described by a Single-Peaked preference profile. It can be proved that Single-Peaked preference profiles rule out odd rings and hence admit stable roommate matchings.

THEOREM 6. *If the preference profile is Single-Peaked, there exist stable roommate matchings.*

Proof. Suppose an odd ring $(x_1, x_2, \ldots, x_k)$ exists. Since $x_2 \succ_{x_1} x_k \succcurlyeq_{x_1} x_1$, it cannot be the case that $B(x_k, x_1, x_2)$. W.l.o.g. assume that $x_1 \gg x_k$ and $x_1 \gg x_2$. Similarly, $x_3 \succcurlyeq_{x_2} x_1 \succcurlyeq_{x_2} x_2$ and $x_1 \gg x_2$ together imply $x_3 \gg x_2$, $x_4 \succ_{x_3} x_2 \succcurlyeq_{x_3} x_3$ and $x_3 \gg x_2$ together imply $x_3 \gg x_4$, and so on. By repeatedly applying this logic, we have $x_5 \gg x_4, \ldots$, and eventually $x_k \gg x_1$ (remember that k is odd), which contradicts the fact that $\gg$ is a strong ordering. Q.E.D.

It can be proved that Single-Caved preference profiles also admit stable roommate matchings. But since the proof does not make use of Theorem 1, we refer interested readers to the Appendix, where we shall also discuss some further relations between stable roommate matchings and the voting paradox.

The fourth condition Inada (1964) identified says that SMDR will always be transitive if the preference profile is Two-Group Separable and the number of agents is odd. We adapt the definition of Two-Group Separability to the roommates problem as follows.

DEFINITION 13. A preference profile is "Two-Group Separable" if every subset $J \subset I$, with $|J| \geq 3$, can be partitioned into two non-empty groups A and B, such that every agent either strictly prefers any agent in A to any agent in B, or strictly prefers any agent in B to any agent in A.

For example, if the partition of $\{x, y, z\}$ is $\{\{x\}, \{y, z\}\}$, there will not be any agent w whose preference ordering is $y \succcurlyeq_w x \succcurlyeq_w z$. To understand how a Two-Group Separable preference profile will arise, imagine a group of economists seeking collaborators. A matching theorist who contemplates writing a paper on implementation of matching rules may have the following preference: among all potential collaborators, he prefers any theorist to any econometrician; among all theorists, he prefers any micro theorist to any macro theorist; among all micro theorists, he prefers any theorist who knows something about either implementation theory or matching theory to any who does not; among all theorists who know something about either implementation theory or matching theory, he prefers any implementation theorist to any matching theorist... On the other hand, a macro theorist who contemplates conducting an empirical test on his model may go down exactly the same binary tree when looking for an appropriate collaborator, but nevertheless have a different preference: among all potential collaborators, he prefers any econometrician to any theorist; among all econometricians, he prefers any time-series econometrician to any micro econometrician; among all time-series econometricians, he prefers any Bayesian to any frequentist... The preferences of these economists can then be described by a Two-Group Separable preference profile. It can be proved that Two-Group Separable preference profiles rule out odd rings and hence admit stable roommate matchings.

THEOREM 7. *If the preference profile is Two-Group Separable, there exist stable roommate matchings.*

Proof. Suppose an odd ring $(x_1, x_2, \ldots, x_k)$ exists. Since $k \geq 3$, the subset $\{x_1, x_2, \ldots, x_k\}$ can be partitioned into non-empty groups A and B with the required property. W.l.o.g. assume $x_1 \in A$. Since $x_2 \succ_{x_1} x_k \succcurlyeq_{x_1} x_1$, $x_2 \in A$ implies $x_k \in A$, and hence $x_{k-1} \in A$ (because $x_1 \succ_{x_k} x_{k-1} \succcurlyeq_{x_k} x_k$), and hence $x_{k-2} \in A$ (because $x_k \succcurlyeq_{x_{k-1}} x_{k-2} \succcurlyeq_{x_{k-1}} x_{k-1}$), ..., and hence $x_3 \in A$ (because $x_5 \succcurlyeq_{x_4} x_3 \succcurlyeq_{x_4} x_4$), and hence all agents are in A, and hence B is empty. Therefore $x_2 \notin A$ and $x_2 \in B$. Since $x_3 \succcurlyeq_{x_2} x_1 \succcurlyeq_{x_2} x_2$, we have $x_3 \in A$, and hence $x_4 \in B$ (because $x_4 \succ_{x_3} x_2 \succcurlyeq_{x_3} x_3$), ..., and hence $x_{k-1} \in B$ (because $x_{k-1} \succ_{x_{k-2}} x_{k-3} \succcurlyeq_{x_{k-2}} x_{k-2}$ and $k-1$ is even), and hence $x_k \in A$ (because $x_k \succcurlyeq_{x_{k-1}} x_{k-2} \succcurlyeq_{x_{k-1}} x_{k-1}$ and k is odd). Again, we arrive at a contradiction because $x_1 \succ_{x_k} x_{k-1} \succcurlyeq_{x_k} x_k$. Q.E.D.

In a remarkable paper, Inada (1969) identified three more sufficient conditions for transitive SMDR. He also proved that these three new sufficient conditions, together with the four identified five years earlier (India, 1964), cover all possible sufficient conditions of a certain type (sufficient conditions of other types are relatively trivial and were ignored by Inada). The preference restriction problem regarding transitive SMDR is hence to a large extent resolved.

The first of the three new conditions identified by Inada (1969) says that SMDR will always be transitive if the preference profile is Echoic. We adapt the definition of Echoic preference profile to the roommates problem as follows.

DEFINITION 14. A preference profile is "Echoic" if for any ordered triple (x, y, z), where $\{x, y, z\} \subset I$, whenever there exists an agent $w \in I$ with preference $x \succ_w y \succ_w z$, every other agent in I weakly prefers x to z (i.e., $\forall w' \in I$, $x \succcurlyeq_{w'} z$).

Echoic preference profile captures the idea that agents are subject to similar socialization processes when forming their individual preferences, so that if someone ranks x far higher than z, the same socialization process must also have moulded other agents to rank x as least as high as z. Imagine a group of boys who wish to divide up into pairs of roommates. Tidiness is the single most important concern every boy has in choosing his roommate. Everyone's tidiness is publicly observed, and hence everyone agrees on the same tidiness-ranking of all boys. However, each boy also has some other minor concerns, and hence his preference over potential roommates is the tidiness-ranking plus some small perturbations. If these perturbations are small enough, the boys' preferences can then be described by an Echoic preference profile. It can be proved that Echoic preference profiles rule out all kinds of rings and hence admit stable roommate matchings.

THEOREM 8. *If the preference profile is Echoic, there exist stable roommate matchings.*

Proof. Suppose a ring $(x_1, x_2, \ldots, x_k)$ exists. W.l.o.g. assume that $\forall 1 \leq i \leq k$, $x_{i-1} \succ_{x_i} x_i$ (subscript modulo k).[11] By the definition of a ring, we have $x_2 \succ_{x_1} x_k \succ_{x_1} x_1$. By the Echoic assumption, we have $x_2 \succcurlyeq_{x_2} x_1$, which contradicts our earlier assumption that $x_1 \succ_{x_2} x_2$. Q.E.D.

Another new condition identified by Inada (1969) says that SMDR will always be transitive if the preference profile is Antagonistic. We adapt the definition of Antagonistic preference profile to the roommates problem as follows.

DEFINITION 15. A preference profile is "Antagonistic" if for any ordered triple (x, y, z), where $\{x, y, z\} \subset I$, whenever there exists an agent $w \in I$ with preference $x \succ_w y \succ_w z$, every other agent $w' \in I$ will have preference either $x \succ_{w'} y \succ_{w'} z$, or $z \succ_{w'} y \succ_{w'} x$, or $x \sim_{w'} z$.

Antagonistic preference profile is a concept just opposite to Echoic preference profile, but we do not have a natural interpretation of it (nor did Inada). It can be proved that Antagonistic preference profiles rule out all kinds of rings and hence admit stable roommate matchings.

THEOREM 9. *If the preference profile is Antagonistic, there exist stable roommate matchings.*

Proof. Suppose a ring $(x_1, x_2, \ldots, x_k)$ exists. W.l.o.g. assume that $\forall 1 \leq i \leq k$, $x_{i-1} \succ_{x_i} x_i$ (subscript modulo k). Starting with $x_2 \succ_{x_1} x_k \succ_{x_1} x_1$, by the Antagonistic assumption, we have $x_3 \succcurlyeq_{x_2} x_1 \succ_{x_2} x_k \succ_{x_2} x_2$. Apply the Antagonistic assumption again, we have $x_4 \succ_{x_3} x_2 \succ_{x_3} x_k \succ_{x_3} x_3$. Iterating this process, we have $x_5 \succcurlyeq_{x_4} x_3 \succ_{x_4} x_k \succ_{x_4} x_4, \ldots$, and eventually, $x_k \succcurlyeq_{x_{k-1}} x_{k-2} \succ_{x_{k-1}} x_k \succ_{x_{k-1}} x_{k-1}$. Since $k \geq 3$, this is impossible. Q.E.D.

Notice that in the above proof, we have not made use of the full implication of the Antagonistic assumption. In particular, whenever there exists an agent w with preference $x \succ_w y \succ_w z$, we only make use of the part of the assumption that $z \succ_{w'} x$ implies $z \succ_{w'} y \succ_{w'} x$, but not the part that $x \succ_{w'} z$ implies $x \succ_{w'} y \succ_{w'} z$. So our result will still hold even if the Antagonistic assumption is relaxed to the Extremal Restriction as studied by Sen and Pattanaik (1969) in the context of social decision function. We adapt the definition of Extremal Restriction to the roommates problem as follows.

DEFINITION 16. A preference profile satisfies "Extremal Restriction" if for any ordered triple (x, y, z), where $\{x, y, z\} \subset I$, whenever there exists

[11] Suppose no odd rings satisfy this hypothesis. Then we can construct a tie-breaking version $\succcurlyeq'$ of the original preference profile $\succcurlyeq$ as follows: for any $x, y \in I$, $y \sim_x x \Longrightarrow x \succ'_x y$. Notice that all the odd rings in $\succcurlyeq$ disappear in $\succcurlyeq'$, and no new odd rings emerge in the due course. Therefore $\succcurlyeq'$ is odd-ring-free, and hence we are done.

an agent $w \in I$ with preference $x \succ_w y \succ_w z$, any other agent $w' \in I$ will have preference $z \succ_{w'} x$ if and only if his preference is $z \succ_{w'} y \succ_{w'} x$.

The proof of the following corollary is then exactly the same as the proof of Theorem 9.

COROLLARY 3. *If the preference profile satisfies Extremal Restriction, there exist stable roommate matchings.*

The last new condition identified by Inada (1969) says that SMDR will always be transitive if the preference profile has Taboos and the number of agents is odd. We adapt the definition of Taboo preference profile to the roommates problem as follows.

DEFINITION 17. A preference profile has "Taboos" if for any triple $\{x, y, z\} \subset I$, there is an ordered pair, say (x, y), such that everyone weakly prefers x to y; moreover, none of the agents have preference $x \sim y \sim z$.

The name of this preference restriction comes from the fact that "$y \succ x$" is essentially a Taboo. It can be proved that Taboo preference profiles rule out odd rings and hence admit stable roommate matchings.

THEOREM 10. *If the preference profile has Taboos, there exist stable roommate matchings.*

Proof. Suppose an odd ring $(x_1, x_2, \ldots, x_k)$ exists. W.l.o.g. assume that $\forall 1 \leq i \leq k$, $x_{i-1} \succ_{x_i} x_i$ (subscript modulo k). Consider any triple of the form $\{x_{i-1}, x_i, x_{i+1}\}$, where i is odd and $1 \leq i \leq k$ (subscript modulo k). Since $x_i \succcurlyeq_{x_{i-1}} x_{i-2} \succ_{x_{i-1}} x_{i-1}$, $x_{i+1} \succ_{x_i} x_{i-1} \succ_{x_i} x_i$, and $x_i \succ_{x_{i+1}} x_{i+1}$, the only possible Taboo is $x_{i-1} \succ x_{i+1}$. Pool together these Taboos for all odd is with $1 \leq i \leq k$, we have $\forall w \in I$, $x_1 \succcurlyeq_w x_{k-1} \succcurlyeq_w x_{k-3} \succcurlyeq_w \cdots \succcurlyeq_w x_4 \succcurlyeq_w x_2 \succcurlyeq_w x_k$. This implies $x_1 \succcurlyeq_{x_1} x_k$, which contradicts our earlier assumption that $x_k \succ_{x_1} x_1$. Q.E.D.

Notice that in the above proof, we have not made use of the full implication of the Taboo assumption. In particular, we have not used the assumption that "$x \sim y \sim z$" is forbidden for every triple $\{x, y, z\}$. So our result will still hold even if the Taboo assumption is relaxed to the Limited Agreement assumption as studied by Sen and Pattanaik (1969) in the context of social decision function. We adapt the definition of Limited Agreement to the roommates problem as follows.

DEFINITION 18. A preference profile satisfies "Limited Agreement" if for any triple $\{x, y, z\} \subset I$, there is an ordered pair, say (x, y), such that everyone weakly prefers x to y.

The relationship between Taboos and Limited Agreement is self-evident: by imposing some Taboos on the community, the agents achieve some Limited Agreement. Both of these preference restrictions try to capture the idea of socialization in the process of preference formation, as does the Echoic preference profile assumption.

The proof of the following corollary is exactly the same as the proof of Theorem 10.

COROLLARY 4. *If the preference profile satisfies Limited Agreement, there exist stable roommate matchings.*

6. CONCLUDING REMARKS

This paper has identified a sufficient condition for the existence of stable roommate matchings. The condition is always satisfied by the marriage problem, and hence explains why the marriage problem stands out as a special case of the roommates problem that always admits stable matchings. The marriage problem rules out odd rings—objects that we have proved to be the only possible source of instability—by construction.

However, this is not the only way to look at the question of why the marriage problem always admits stable matchings. The marriage problem can be viewed not only as a special case of the roommates problem, but as a special case of two-sided matching as well.[12] Two-sided matching problems are characterized such that agents are partitioned into two types, and each agent can choose only agents of the other type as partners. A two-sided matching problem which nests the marriage problem as a special case is known as the college admission problem, where each student can enroll only at one college but each college may have more than one student. The college admission problem always admits stable matchings. Since the marriage problem can be viewed as a special case of the college admission problem where each college can only have one student, it also explains why the marriage problem always admits stable matchings. This explanation is different from the one provided in this paper, as the roommates problem and the college admission problem are different problems, with the marriage problem their only point of contact. It will be interesting for future research to construct a unified framework, which nests both the roommates problem and the college admission problem as its special cases, and characterize the class of domains that admits stable matchings.

It is also worth mentioning that Sönmez (1996) provided a unified framework, called the "generalized matching problem," which nests both the

[12]For masterful surveys of two-sided matching, see Roth and Sotomayor (1990, 1992).

roommates problem and the housing market (Shapley and Scarf (1974); see also Roth (1982)) as its special cases. The difference between the roommates problem and the housing market is that feasible matchings in the roommates problem are restricted to be of order two (i.e., for any feasible matching μ, $\forall x \in I$, $\mu^2(x) = x$), while this restriction is lifted in the housing market. In the roommates problem, the set of stable roommate matchings coincides with the core defined by strong domination. In contrast to the fact that the core in the roommates problem may be empty, the core in the housing market is always non-empty. It will also be interesting for future research to better characterize the condition for non-empty core in the generalized matching problem.

APPENDIX: STABLE ROOMMATE MATCHINGS AND THE VOTING PARADOX

In this appendix, we shall first prove that the last Inada condition (1964), namely Single-Cavedness, also implies the existence of stable roommate matchings. This result, together with the fact that each of the other six Inada conditions implies the existence of stable roommate matchings, raises the question of whether there exists any deep connection between stable roommate matchings and transitive SMDR. We shall discuss the possible connection between these two problems in the rest of this appendix.

THEOREM 11. *If the preference profile is Single-Caved, there exist stable roommate matchings.*

Proof. We prove this theorem by providing an algorithm to construct a stable roommate matching.

We shall say that $x \in I' \subset I$ is the "maximal agent" within I' if for all $w \in I'$ such that $w \neq x$, $x \gg w$. Similarly we can define the "minimal agent" within I'. We shall say that x is an "extremal agent" within I' if he is either the maximal or the minimal agent within I'. Now consider the two extremal agents within I, say x and y. Suppose $x \succ_x y$. Then x is single in any stable matching as all other agents in I are unacceptable to him, and we can leave x alone and start the procedure again with I replaced by $I' \equiv I \setminus \{x\}$. So w.l.o.g. we may assume $y \succcurlyeq_x x$ and $x \succcurlyeq_y y$. We shall then match x with y, and it is obvious that they are each other's most preferred roommates.

Now replace I with $I' \equiv I \setminus \{x, y\}$ and observe that the original preference profile as restricted to I' is still Single-Caved with respect to the same strong ordering $\gg$. So we can start the above procedure again with I' and match its two extremal agents, say x' and y', with each other. By iterating the above procedure until all agents are either matched or left alone at some stage, we will arrive at a matching. It is easy to verify that it is also a stable roommate matching. Q.E.D.

Since each of the seven Inada conditions is sufficient for transitive SMDR, and are jointly necessary for transitive SMDR *in some sense*,[13] a natural question is whether transitive SMDR is necessary and sufficient for the existence of stable roommate matchings. In any case, isn't the simplest example of non-existence of stable roommate matchings (i.e., Example 1) also the simplest example of the voting paradox?

The answer to this question is negative. The following two counterexamples show that transitive SMDR is neither necessary nor sufficient for the existence of stable roommate matchings.

EXAMPLE 4. Consider $I = \{x, y, z\}$, and the following preferences:

$$x \succ_x y \succ_x z,$$

$$y \succ_y z \succ_y x,$$

$$z \succ_z x \succ_z y.$$

The same voting paradox arises as in Example 1. But the matching μ, where $\forall a \in \{x, y, z\}$, $\mu(a) = a$, is a stable roommate matching.

EXAMPLE 5. Consider $I = \{a, b, c, d, e\}$, and the following preferences:

$$b \succ_a e \succ_a d \succ_a c \succ_a a,$$

$$e \succ_b d \succ_b c \succ_b a \succ_b b,$$

$$e \succ_c d \succ_c b \succ_c c \succ_c a,$$

$$e \succ_d c \succ_d d \succ_d b \succ_d a,$$

$$a \succ_e d \succ_e e \succ_e c \succ_e b.$$

In this example, SMDR will generate a transitive social preference $e \succ d \succ c \succ b \succ a$, yet no stable roommate matchings exist. To see this, observe that in any stable roommate matching, at least one of the agents is single. If a (respectively, b and e) is single, he can propose to e (respectively, a and d). If c is single, he can propose to b, as b cannot have been matched with either e or d in any stable roommate matching. If d is single, he can propose to c, as c cannot have been matched with e in any stable roommate matching.

However, probing for a connection between stable roommate matchings and transitive SMDR "over agents" may be off the point. In any case, the majority voting over partners does not solve a given instance of the roommates problem.[14] A more promising direction of research may be to

[13]See Inada (1969) for the exact sense in which the Inada conditions are jointly necessary.

[14]I thank a referee for pointing this out to me.

look for a possible connection between stable roommate matchings and majority voting "over matchings."

To meaningfully talk about majority voting over matchings, we can extend agents' preferences over possible roommates to preferences over possible matchings in the *selfish* manner; i.e., any agent x weakly (strictly) prefers μ to μ' if he weakly (strictly) prefers $\mu(x)$ to $\mu'(x)$. This is the standard way to extend preferences in the literature on implementation of matching rules.[15] Given these extended preferences, it is now meaningful to ask whether the agents can use SMDR to collectively decide on the choice of matching. Notice that, although the preferences over agents and the preferences over matchings are intimately related, it may happen that SMDR over agents is transitive while SMDR over matchings is not. The reader can check that this is actually what happens in Example 5.

Let $\mathcal{S}$ be the set of stable matchings, and $\mathcal{I}$ is the set of IR matchings. By definition, we have $\mathcal{S} \subset \mathcal{I}$. For any two matchings μ and μ', define $I(\mu P \mu') \equiv \{x \in I \mid \mu(x) \succ_x \mu'(x)\}$. For any two matchings μ and μ', if $|I(\mu P \mu')| > |I(\mu' P \mu)|$, we say that μ "Condorcet-beats" μ', or simply μ "beats" μ'. Let $\mathcal{C}$ be the set of Condorcet winners among all matchings; i.e., for any two matchings μ and μ', $\mu \in \mathcal{C} \Longrightarrow |I(\mu P \mu')| \geq |I(\mu' P \mu)|$. Let $\mathcal{C}_{\mathcal{I}}$ be the set of Condorcet winners among IR matchings; i.e., for any two matchings $\mu, \mu' \in \mathcal{I}$, $\mu \in \mathcal{C}_{\mathcal{I}} \Longrightarrow |I(\mu P \mu')| \geq |I(\mu' P \mu)|$. There is no straightforward inclusion relation between $\mathcal{C}$ and $\mathcal{C}_{\mathcal{I}}$, although both are supersets of $\mathcal{C} \cap \mathcal{I}$.[16] In Example 5, both $\mathcal{C}$ and $\mathcal{C}_{\mathcal{I}}$ are empty.[17]

The main result in this appendix is that, in the special case of strict preferences over agents, stable roommate matchings exist only if there exist Condorcet winners in majority voting over matchings.

THEOREM 12. *If preferences over agents are strict, then* $\mathcal{C} = \emptyset \Longrightarrow \mathcal{S} = \emptyset$.

Proof. It suffices to prove that $\mathcal{S} \subset \mathcal{C}$. Let $\mu \in \mathcal{S}$, and μ' is an arbitrary matching. For any $a \in I$ such that $\mu'(a) = a$, $\mu(a) \succcurlyeq_a a = \mu'(a)$, and hence $a \notin I(\mu' P \mu)$. For any $\{a, b\} \subset I$ such that $a = \mu'(b) \neq b$, either

[15]See, among other contributions to this literature, Alcalde and Barberá (1994), Kara and Sönmez (1996), Sönmez (1997), and Chung (1998).

[16]In particular, a Condorcet winner among IR matchings, $\mu \in \mathcal{C}_{\mathcal{I}}$, may not beat every non-IR matching, and hence may not be a Cordorcet winner (i.e., $\mu \notin \mathcal{C}$). Similarly, a Condorcet winner, $\mu \in \mathcal{C}$, may not be IR, and hence may not be a Condorcet winner among IR matchings (i.e., $\mu \notin \mathcal{C}_{\mathcal{I}}$).

[17]In Example 5, there are totally 26 possible matchings, of them 11 are IR. It seems to us that there are no alternatives other than to check that each of the 26 matchings is beaten by some other matchings, and that each of the 11 IR matchings is beaten by some other IR matchings.

$a = \mu(b)$—and hence both a and b are indifferent between μ and μ'—or $a \neq \mu(b)$. In the latter case, by the stability of μ, at least one of $\{a, b\}$, say a, has preference $\mu(a) \succ_a b$. Therefore $|I(\mu' P \mu)|$ cannot be bigger than $|I(\mu P \mu')|$. Q.E.D.

COROLLARY 5. *If preferences over agents are strict, then* $\mathcal{C}_{\mathcal{I}} = \emptyset \Longrightarrow \mathcal{S} = \emptyset$.

Proof. Since $\mathcal{S} \subset \mathcal{I}$, together with Theorem 12, we have $\mathcal{S} \subset \mathcal{C} \cap \mathcal{I} \subset \mathcal{C}_{\mathcal{I}}$. Q.E.D.

The converses of both Theorem 12 and Corollary 5 are not true even in the special case of strict preferences over agents. This can be seen from the following example.

EXAMPLE 6. Consider $I = \{a, b, c, d, e, f\}$, and the following preferences:

$$b \succ_a c \succ_a d \succ_a a \succ_a \cdots,$$

$$c \succ_b a \succ_b e \succ_b b \succ_b \cdots,$$

$$a \succ_c b \succ_c f \succ_c c \succ_c \cdots,$$

$$a \succ_d d \succ_d \cdots,$$

$$b \succ_e e \succ_e \cdots,$$

$$c \succ_f f \succ_f \cdots.$$

We can see that $\mathcal{S} = \emptyset$ because (a, b, c) is a strict odd ring that bites. But the matching μ, where $\mu(a, b, c) = (d, e, f)$, belongs to both $\mathcal{C}$ and $\mathcal{C}_{\mathcal{I}}$.

Both Theorem 12 and Corollary 5 fail to generalize to the case of weak preferences. This can be seen from the following example.

EXAMPLE 7. Consider $I = \{a, b, c, d, e, f, g, h, i\}$, and the following preferences:

$$b \succ_a d \succ_a c \succ_a a \sim_a x, \quad \forall x \in \{e, f, g, h, i\},$$

$$c \succ_b e \succ_b a \succ_b b \sim_b x, \quad \forall x \in \{d, f, g, h, i\},$$

$$a \succ_c f \succ_c b \succ_c c \sim_c x, \quad \forall x \in \{d, e, g, h, i\},$$

$$a \sim_d g \succ_d d \sim_d x, \quad \forall x \in \{b, c, e, f, h, i\},$$

$$b \sim_e h \succ_e e \sim_e x, \quad \forall x \in \{a, c, d, f, g, i\},$$

$$c \sim_f i \succ_f f \sim_f x, \quad \forall x \in \{a, b, d, e, g, h\},$$

$$d \succ_g g \sim_g x, \quad \forall x \in \{a, b, c, e, f, h, i\},$$

$$e \succ_h h \sim_h x, \quad \forall x \in \{a, b, c, d, f, g, i\},$$

$$f \succ_i i \sim_i x, \quad \forall x \in \{a, b, c, d, e, g, h\}.$$

There are seven stable matchings:[18]

$$\mu = \{\{a, d\}, \{b, e\}, \{c, f\}, \{g\}, \{h\}, \{i\}\},$$

$$\mu_g = \{\{a, d\}, \{b, e\}, \{c, f\}, \{g\}, \{h, i\}\},$$

$$\mu_h = \{\{a, d\}, \{b, e\}, \{c, f\}, \{g, i\}, \{h\}\},$$

$$\mu_i = \{\{a, d\}, \{b, e\}, \{c, f\}, \{g, h\}, \{i\}\},$$

$$\mu_a = \{\{a, d\}, \{b, c\}, \{g\}, \{e, h\}, \{f, i\}\},$$

$$\mu_b = \{\{a, c\}, \{b, e\}, \{d, g\}, \{h\}, \{f, i\}\},$$

$$\mu_c = \{\{a, b\}, \{c, f\}, \{d, g\}, \{e, h\}, \{i\}\}.$$

However, $\mathscr{C} = \mathscr{C}_{\mathscr{I}}$ (because all possible matchings are IR) and both are empty. To see this, first observe that $\forall i \in \{a, b, c\}$ and $\forall j \in \{g, h, i\}$, μ_i beats μ_j and μ. Moreover,

$$I(\mu_a P \mu_b) = \{a, b, h\}, \quad I(\mu_b P \mu_a) = \{c, g\},$$

$$I(\mu_b P \mu_c) = \{b, c, i\}, \quad I(\mu_c P \mu_b) = \{a, h\},$$

$$I(\mu_c P \mu_a) = \{c, a, g\}, \quad I(\mu_a P \mu_c) = \{b, i\};$$

i.e., μ_a beats μ_b beats μ_c beats μ_a, and hence none of the stable matchings are Condorcet winners.

Suppose $\mathscr{C} \neq \varnothing$ and $\nu \in \mathscr{C}$. Suppose $\nu(d) \sim_d d$. Then ν will be beaten by matching ν' which coincides with ν except that $\nu'(d) = g$, a contradiction. Therefore $\nu(d) \succ_d d$. Similarly, $\nu(e) \succ_e e$ and $\nu(f) \succ_f f$.

Suppose $\nu(d, e, f) = (g, h, i)$. Then there exists at least one agent in $\{a, b, c\}$, say a, such that $\nu(a) \sim_a a$. But then ν will be beaten by matching ν' which coincides with ν except that $\nu'(a) = c$, a contradiction.

Suppose $\nu(d, e, f) = (a, h, i)$. Then ν either equals μ_a or can be beaten by μ_a, a contradiction. Similarly, it cannot be the case that $\nu(d, e, f) = (g, b, i)$ or $\nu(d, e, f) = (g, h, c)$.

Suppose $\nu(d, e, f)) = (a, b, i)$. Then ν can be beaten by μ_a. Similarly, it cannot be the case that $\nu(d, e, f) = (a, h, c)$ or $\nu(d, e, f) = (g, b, c)$. Since we have exhausted all the possible matchings, we conclude that $\mathscr{C} = \mathscr{C}_{\mathscr{I}} = \varnothing$.

[18] $\mu = \{\{x, y\}, \ldots\}$ means $\mu(x) = y \ldots$, etc.

REFERENCES

Abeledo, H. G., and Rothblum, U. G. (1994). “Stable Matchings and Linear Inequalities,” *Discrete Appl. Math.* **54**, 1–27.

Alcalde, J., and Barberá, S. (1994). “Top Dominance and the Possibility of Strategy-Proof Stable Solutions to Matching Problems,” *Econom. Theory* **4**, 417–435.

Arrow, K. J. (1951). *Social Choice and Individual Values*, Cowles Foundation Monograph 12, New Haven: Yale University Press.

Bartholdi, J. J., III, and Trick, M. A. (1986). “Stable Matching with Preferences Derived from a Psychological Model,” *Oper. Res. Lett.* **5**, 165–169.

Becker, G. S. (1973). “A Theory of Marriage: Part I,” *J. Polit. Econ.* **81**, 813–846.

Chung, K. S. (1998). “Implementing Stable Matchings under Incomplete Information,” mimeo, University of Wisconsin.

Gale, D., and Shapley, L. (1962). “College Admissions and the Stability of Marriage,” *Amer. Math. Monthly* **69**, 9–15.

Inada, K. I. (1964). “A Note on the Simple Majority Decision Rule,” *Econometrica* **32**, 525–531.

Inada, K. I. (1969). “The Simple Majority Decision Rule,” *Econometrica* **37**, 490–506.

Kara, T. and Sönmez, T. (1996). “Nash Implementation of Matching Rules,” *J. Econom. Theory* **68**, 425–439.

May, K. O. (1952). “A Set of Independent Necessary and Sufficient Conditions for Simple Majority Decision,” *Econometrica* **20**, 680–684.

May, K. O. (1953). “A Note on the Complete Independence of the Conditions for Simple Majority Decision,” *Econometrica* **21**, 172–173.

Roth, A. E. (1982). “Incentive Compatibility in a Market with Indivisible Goods,” *Econ. Lett.* **9**, 127–132.

Roth, A. E., and Sotomayor, M. (1990). *Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis*, Econometric Society Monograph 18, Cambridge: Cambridge Univ. Press.

Roth, A. E., and Sotomayor, M. (1992). “Two-Sided Matching,” in *Handbook of Game Theory*, (R. J. Aumann and S. Hart, Eds.), Vol. 1, pp. 485–541. Amsterdam: Elsevier Science.

Roth, A. E., and Vande Vate, J. H. (1990). “Random Paths to Stability in Two-Sided Matching,” *Econometrica* **58**, 1475–1480.

Sen, A. K., and Pattanaik, P. K. (1969). “Necessary and Sufficient Conditions for Rational Choice under Majority Decision,” *J. Econom. Theory* **1**, 178–202.

Shapley, L., and Scarf, H. (1974). “On Cores and Indivisibility,” *J. Math. Econ.* **1**, 23–37.

Sönmez, T. (1996). “Implementation in Generalized Matching Problems,” *J. Math. Econ.* **26**, 429–439.

Sönmez, T. (1997). “Games of Manipulation in Marriage Problems,” *Games Econ. Behav.* **20**, 169–176.

Tan, J. J. M. (1991). “A Necessary and Sufficient Condition for the Existence of a Complete Stable Matching,” *J. Algorithms* **12**, 154–178.