



ISSN 2282-6483

Alma Mater Studiorum - Università di Bologna
DEPARTMENT OF ECONOMICS

**A simple mechanism for the
roommate problem**

Bora Evcı

Quaderni - Working Paper DSE N°975



A SIMPLE MECHANISM FOR THE ROOMMATE PROBLEM*

Bora Evci[†]

October 27, 2014

Abstract

[Gale and Shapley \(1962\)](#) proposed that there is a similar game to the marriage problem called "the roommate problem". And, they showed that unlike the marriage problem, the roommate problem may have unstable solutions. In other words, the stability theorem fails for the roommate problem.

In this paper, we propose a new mechanism for the roommate problem. The mechanism is successful in determining the reason of instability in our game scenario. And, we show that our mechanism implements the full set of stable matchings in the existence of stability, and it ends up with Pareto Optimal matching in the instance of instability.

JEL CODES: C78, D71, D78.

Keywords: Roommate Problem, Stable Matching, Implementation.

*I would like to thank Jean Lainé for the initial discussion at the beginning of the project and encouraging me to work on the Roommate Problem. I am grateful to Vincenzo Denicolò for his great comments and support. All errors are my own.

[†]University of Bologna, Italy. Email: bora.evci3@unibo.it.

1 Introduction

In their seminal paper, [Gale and Shapley \(1962\)](#) described the well known roommate problem. In the problem, there are even number of college boys and dormitory rooms for the pairs of boys. Each boy has a preference ordering over the other boys. The objective is to allocate the boys to these rooms in pairs. Gale and Shapley showed that the stability theorem does not hold for the roommate problem by giving a counter-example in their paper.

[Knuth \(1976\)](#) showed that multiple solutions could exist for the roommate problem, like in the marriage problem. In his 12 famous questions, which he raised during the lectures at the University of Montréal in 1976, he asked for an efficient algorithm to find a stable solution for the roommate problem.

[Irving \(1985\)](#) proposed an algorithm for the roommate problem. The algorithm has two phases. The first one is similar to Gale and Shapley's algorithm. Instead of simultaneously, the proposals are made sequentially. At end of the first phase, a deletion process takes place. Basically, the unachievable agents for the proposers and the worse agents for the deciders than the current mates are deleted from the preference lists of all agents. Whether the original preference profile has no stable solution or has one or more stable solutions could be determined from this reduced form of the preference profile. If there exists a preference list without any agent, then it means that for the original preference profile there is no stable solution. If all the lists contain only one agent, then the profile has a unique stable solution. If some preference lists contain more than one agent, the algorithm proceeds with the second phase which involves further deletions. The second phase ends up with a unique stable matching, in the existence of multi stability, since the algorithm breaks the cycles among the agents. Thus, his algorithm tells whether a given profile has a stable solution or not and if there exists some, the algorithm finds one.

Tan (1991) proposed a criterion which is a necessary and sufficient condition for the existence of stability; namely, as he calls, the non-existence of odd party. He showed that the existence of the so-called odd party is the reason of instability for the roommate problem. He also proposed an algorithm which is a modified version of Irving (1985). The first phase is again a propose-reject process. After the deletion of impossibility agents at the end of the first phase, if every person has zero or one entry on the lists, then this leads to a unique stable matching. If there exists a person having more than one agent in his list, the algorithm proceeds to the second phase. The second phase continues with the elimination of the cycles. If the elimination of a cycle makes some lists empty, it indicates an odd party, which signals an instability solution for the original preference profile.

Tan and Hsueh (1995) proposed a new algorithm to the problem. In their paper, the preference orderings are allowed to expand. In their model, they make the analysis of a new comer to the game or a leave of an agent on the existence of, as Tan (1991) calls, the stable partitions and how to find one (stable matching). When there is a new comer, he proposes to the agents according to his preference orderings. They describe the proposal sequence, the positions of the new comer(s) (one by one in an order) among the old agents and the number of new members to maintain the stability of the preference profile. Stability is maintained as long as there is no new "odd parties" if the initial profile does not have any or all the odd parties are eliminated if starting preference profile has some.

Cechlarova and Fleiner (2005) proposed a model for the roommate problem with parallel edges. That is two agents are matched with different issues at the same time. They show an equivalence with this model to the original roommate problem and they proposed an extension of Irving (1985) to seek stability. The

main difference is that after the algorithm eliminates the rotations in the second phase, it returns back to the first phase. So, it has more complexity than the original algorithm of [Irving \(1985\)](#).

In this paper, we propose a simple mechanism for the roommate problem for strict preferences. We simply extend the mechanism proposed in [Evcı \(2014\)](#). The problem here is that his mechanism is designed for the two-sided matching games. But, the roommate problem has a one-sided framework. Hence, firstly we convert the model of the roommate problem into a two-sided game by using auxiliary functions, and then we apply the mechanism by [Evcı \(2014\)](#) to this modified market. We show that the mechanism finds a/the stable matching in the existence of stability and ends up with Pareto optimal matching in the absence of stability. And, we also describe the method to the multi stability case by using our mechanism. In other words, we show how to fully implement the set of stable matchings for any preference profile in the existence of stability.

There is a literature on the domain restrictions to obtain stability for any preference profile and also a literature analyzing the restrictions on the colleges for the two-sided matching markets. But, these are out of the scope of this paper.

The paper is organized as follows. Section 2 introduces the preliminaries. In Section 3, we present the mechanism of [Evcı](#) and our refinement with its stability analysis. Section 4 concludes.

2 Basics and Examples

"The Roommate Problem" is one of the most interesting examples of matching theory. The problem was proposed firstly by [Gale and Shapley \(1962\)](#). In the roommate problem, we have two finite sets; there are $2n$ college boys and n dormitory rooms. Each boy has a preference ordering over the other $(2n - 1)$ boys. The objective is to allocate these boys to the rooms in pairs.

Let $B = \{b_1, \dots, b_k\}$ be a non-empty sets of agents (e.g. college boys). Each agent has a strict preference ordering R over the other agents of the set; for example R_{b_i} is the preference ordering of $b_i \in B$ over $B/\{b_i\}$. For any $b_i, b_j, b_k \in B$, $b_j R_{b_i} b_k$ means b_i prefers b_j over b_k . A *Preference Profile* $\mathbb{R} = R^B = (R_i)_{i \in B}$ is a set of preference orderings, one for each agent in the model. Let \mathfrak{R} be the set of all preference profiles.

$r_{b_j}(b_i)$ is the rank of agent $b_i \in B$ in the preference ordering of agent $b_j \in B$. For example, $r_{b_j}(b_i) = k$ means that b_i is the k^{th} best roommate for b_j .

A (one-sided) matching $\mu : B \rightarrow B$ is an injection. For any $b_i, b_j \in B$, $\mu(b_i) = b_j$ means that b_j is the match of b_i and vice versa. Here we explicitly assume that no agent remains single in the matching; that is $\nexists b_i \in B$ such that $\mu(b_i) = b_i$. Π^B is the set of all matchings among the agents of B .

Let $\mu_x, \mu_y \in \Pi^B$ be two matchings and $b_i \in B$. We can rank matchings, from the point view of agent b_i , according to how b_i ranks the agents he is matched with. If $\mu_x(b_i) R_{b_i} \mu_y(b_i)$, then we say that for agent b_i , μ_x *Pareto Dominates* μ_y . If $\mu_x(b_i) = \mu_y(b_i)$, then b_i is *indifferent* between μ_x and μ_y and we denote this by $\mu_x I_{b_i} \mu_y$. (In this model, we work under strict preferences).

For any preference profile $\mathbb{R} = R^B = (R_i)_{i \in B}$ and a matching μ , for any $b_i, b_j \in B$, $(b_i, b_j) \notin \mu$ is called a *blocking pair*, if $b_j R_{b_i} \mu(b_i)$ and $b_i R_{b_j} \mu(b_j)$. If there is no blocking pair for μ , then we say μ is *stable*; otherwise, it is *unstable*.

A *Matching Mechanism* γ is a procedure to select a matching from every preference profile. Formally

$$\gamma : \mathfrak{R} \longrightarrow \Pi^B.$$

In their paper, [Gale and Shapley \(1962\)](#) give a counter example which shows that the stability theorem, which holds for the marriage problem, fails for the roommate problem. They say "...consider boys α, β, γ and δ , where α ranks β first, β ranks γ first, γ ranks α first, and α, β and γ all rank δ last. Then

regardless of δ 's preferences there can be no stable pairing, for whoever has to room with δ will want to move out, and one of the other two will be willing to take him in...". We shall demonstrate their example with the following preference profile \mathbb{R}_1 ,

$$\mathbb{R}_1 = \begin{array}{cccc} \underline{a} & \underline{b} & \underline{c} & \underline{d} \\ b & c & a & c \\ c & a & b & b \\ d & d & d & a \end{array}$$

where $N = \{a, b, c, d\}$ be the set of boys and $\Pi^N = \{\mu_1, \mu_2, \mu_3\}$ be the set of all possible matchings, where

$$\begin{aligned} \mu_1 &= \{(a, b), (c, d)\}, \\ \mu_2 &= \{(a, c), (b, d)\}, \\ \mu_3 &= \{(a, d), (b, c)\}. \end{aligned}$$

None of these matchings is stable for \mathbb{R}_1 ; μ_1 is blocked by (b, c) , μ_2 is blocked by (a, b) and μ_3 is blocked by (a, c) . So, in this one-sided game, we observe unstable solutions as well as the stable ones.

3 The Mechanism

In this section, we present our mechanism which is basically a refinement of the mechanism by [Evcı \(2014\)](#).

3.1 The Dynamic Mechanism by Evcı(2014)

[Evcı \(2014\)](#) proposed a dynamic mechanism for the two-sided matching markets under strict preferences.

Let $M = \{m_1, \dots, m_k\}$ and $W = \{w_1, \dots, w_l\}$ be two non-empty, finite and disjoint sets of agents (e.g. men and women).

The mechanism is designed for, as he calls, the *semi-centralized* market form. While the matching process is centralized for one side of the market, it is decentralized for the other.

For a given matching game $\mathbb{R} = (R_i)_{i \in M \cup W}$, one side is assigned as *the Restrictor*, and the other side as *the Chooser*. The preferences restricters are regarded as the restrictions or the priorities on the chooser side. The choosers make decisions simultaneously at the preferences of the restricter side.

Without loss of generality, he assigns M as the restricter and W as the chooser sides and he shows that set of the outcomes does not depend on which set is the restricter or the chooser side.

We start with the best women in view of some men. These women are called to decide; either to say "yes" or "no" to men who propose to them. If a woman says "yes" to a man and accepts his offer, then they form a pair and both of them are deleted from the profile; if she says "no", she loses that man forever and waits for her turn for other men. At the end of the first step, all agents are informed about the results.

The second step continues with same scenario. And, so on.

Now we give the game scenarios of his mechanism.

Definition 1 *Let $w \in W$ be any chooser agent and $m_i, m_j \in M$ be any two restricter agents with $r_{m_i}(w) > r_{m_j}(w)$ and $w_i R_m w_j$. If at the step $k = r_{m_j}(w)$ none of m_i and m_j have been taken by other choosers yet, then we say the agent w experiences a **conflict** between agents m_i and m_j .*

The definition says that for a chooser if the offer of a restricter comes before any better one, then the chooser agent experiences a conflict.

Definition 2 *If a chooser agent $w \in W$ does not experience any conflict, then we say w has a **smooth game**.*

The focus of his paper is for the special case of the conflicts.

Definition 3 Let $\{m_1, \dots, m_r\} \subset M$ be a set of restricters and $\{w_1, \dots, w_r\} \subset W$ be a set of choosers. If we have such a case;

- $w_1 R_{m_1} w_2, w_2 R_{m_2} w_3, \dots, w_r R_{m_r} w_1,$
- $m_1 R_{w_2} m_2, \dots, m_{r-1} R_{w_r} m_r, m_r R_{w_1} m_1,$
- $r_{m_1}(w_1) = r_{m_2}(w_2) = \dots = r_{m_r}(w_r) = k$ (for at least one side),

Then, agents of $\{w_1, \dots, w_r\}$ experience a **cyclical conflict** with each other for the agents of set $\{m_1, \dots, m_r\}$ at step k .

His first existence result is on the relationship between the cycles and multi stability.

Theorem 4 For any given preference profile $\mathbb{R} = (R_i)_{i \in M \cup W}$, there exists only a unique stable matching if and only if there exists no cyclical conflict for the choosers.

This theorem explains the reason of multi stability for any preference profile. He then presents the types of cycles under his game scenarios in a preference profile. In the proof of theorem, he also showed that Nash Equilibria for a single cycle generates two stable matchings.

Definition 5 Let M and W be the sets of restricters and choosers, respectively. Let $W_1, W_2 \subset W$ be the set of the agents of two cycles. If $W_1 \cap W_2 = \emptyset$, then we say that the cycles are **independent**. Otherwise, they are **(inter) dependent**.

His next result is on the relationship between the number of stable matchings and the number of cycles in a profile.

Proposition 6 *Two (inter) dependent cycles generate three stable matchings.*

From *Proposition 6*, the idea saying that "each cycle produces two stable matchings" fails. Unfortunately we cannot know further about this relationship between the number of cycles and the number of stable matchings, because of the common agents in the existence of several dependent cycles for a preference profile.

And, finally next theorem explains the "partial order structure" of stable matchings.

Theorem 7 *For any preference profile, there exist independent cyclical conflicts which occur at the same step k if and only there exist **incomparable** stable matchings.*

His next theorem is on the implementation of stable matchings.

Theorem 8 *If Nash Equilibria of the cycles are chosen, mechanism γ implements the full set of stable matchings for any preference profile. In other words, we always end up with one of the stable matchings for any profile.*

He also shows that truth telling is weakly dominant for the choosers.

3.2 The Refinement of the Mechanism

Mechanism γ in [Evci \(2014\)](#) is designed for the two-sided matching markets. Therefore, we should modify either the mechanism or the structure of the roommate problem. In this paper, we stick to the mechanism and we convert the roommate problem into a two-sided matching problem. Thus, we need a method to separate the set of boys N into two disjoint sets. For this purpose, we benefit from a well-known social welfare function.

Definition 9 Let A be a set of alternatives, with $\text{Card}(A) = m$, and N be a set of agents, with $\text{Card}(N) = n$. Each agent has a strict preference ordering R over A , e.g. $\forall i \in N$, R_i is the (strict) preference ordering of agent i over the set of alternatives A . Let $\mathbb{R} = R^N = (R_i)_{i \in N}$ be a preference profile and \mathfrak{R} be the set of all preference profiles.

A Social Welfare Function (SWF) $f : \mathbb{R} \rightarrow R$ gives the social preference of the society N over the alternative set A , where $\mathbb{R} \in \mathfrak{R}$.

This is the usual definition of a social welfare function. Next, we give the definition of a famous SWF, which is one of the Scoring Rules.

Definition 10 In a preference profile, the Borda Score $BS(a)$ of an alternative $a \in A$ is $BS(a) = \sum_{i \in N} [(m + 1) - r_i(a)]$. In a voting system, the Borda Rule as a SWF, ranks the alternatives according to their Borda Scores. We allow weak orders in the social preference.

And, this is the usual the Borda Rule definition. In the roommate problem, since there is no alternative set, we modify the definition of the Borda Rule to this game. Now, we will show this modification with an example.

Example 11 Let $N = \{a, b, c, d, e, f\}$ be the set boys with preference profile \mathbb{R}_2 ,

<u>a</u>	<u>b</u>	<u>c</u>	<u>d</u>	<u>e</u>	<u>f</u>	
b	c	d	a	a	b	5 points
c	d	a	b	b	a	4 points
d	a	b	c	c	c	3 points
e	e	f	f	d	d	2 points
f	f	e	e	f	e	1 point

Now, we shall compute the Borda scores of the agents.

$$B(a) = 0 + 3 + 4 + 5 + 5 + 4 = 21$$

$$B(b) = 5 + 0 + 3 + 4 + 4 + 5 = 21$$

$$B(c) = 4 + 5 + 0 + 3 + 3 + 3 = 18$$

$$B(d) = 3 + 4 + 5 + 0 + 2 + 2 = 16$$

$$B(e) = 2 + 2 + 1 + 1 + 0 + 1 = 7$$

$$B(f) = 1 + 1 + 2 + 2 + 1 + 0 = 7$$

From these scores, we get the Borda ranking of the set of boys and together with the preference profile we have,

$$\mathbb{R}_2^* = \begin{array}{cccccc} \underline{a} & \underline{b} & \underline{c} & \underline{d} & \underline{e} & \underline{f} & \underline{B(\mathbb{R}_2)} \\ b & c & d & a & a & b & ab \\ c & d & a & b & b & a & c \\ d & a & b & c & c & c & d \\ e & e & f & f & d & d & ef \\ f & f & e & e & f & e & \end{array}$$

Since the mechanism in [Evcı \(2014\)](#) is called γ , we shall denote ours by β . β is defined over any preference profile \mathbb{R} and its Borda ranking $B(\mathbb{R})$ into the set of matchings Π^N . Formally,

$$\beta : (\mathbb{R}, B(\mathbb{R})) \longrightarrow \Pi^N.$$

Now, we describe how our mechanism works here. We use the Borda ranking of the preference profile to convert the game into two-sided case; that is we use it to generate two sides of the market. The Borda ranking gives the order of the agents that will be the restricters in all successive stages of the game.

We assign the first agent in the Borda ranking as the restricter of the first stage. Then, all the other agents take place in the chooser side. If there is more

than one agent at the top of the Borda ranking, we randomly break the tie and assign the top agent as the restricter. Then, we run mechanism γ .

Claim 12 *At the end of the first stage, we get a pair which consists of the restricter and one of the choosers.*

Proof. The proof is easy. Since the restricter is (one of) the top agent(s) in the Borda ranking, he is (one of) the favorite agent(s). If there exists a chooser agent whose best agent is the restricter, then they form a pair which is trivial to show.

So, let us assume that there is no agent whose best agent is this restricter. This is possible under the Borda rule. If all the choosers reject, then we get an unstable matching (in the existence of stability) which is against the rationality of the agents. Since there is only a single ordering, the choosers do not confront any conflict or cyclical conflict as they do in the games for γ under two-sided framework. Thus (that is for serial dictatorship), it is easy to show that in subgame perfect Nash equilibrium there exists a chooser that accepts the offer because the better alternatives than this restricter agent are not achievable for him. ■

Then, we delete the agents of this pair from the preference profile and the Borda ranking. In the second stage of the game, we assign the best agent among remains in the Borda ranking as the restricter. Then, we run our mechanism. And, so on.

Now, we demonstrate the mechanism β with an example.

Example 13 We will study \mathbb{R}_2 in Example 11.

	a	b	c	d	e	f	$B(\mathbb{R}_2)$
$\mathbb{R}_2^* =$	b	c	d	a	a	b	ab
	c	d	a	b	b	a	c
	d	a	b	c	c	c	d
	e	e	f	f	d	d	ef
	f	f	e	e	f	e	

Since there is tie between a and b , we randomly choose one of them.

Firstly, let us pick a as the restricter. Firstly, b is called to make a choice. If he accepts the offer, then he forms the pair (a, b) . In the second stage c becomes the restricter. d is called for an offer and d accepts the offer since a and b are deleted from the profile and so there is no better mate remained. Then, he forms the pair (c, d) . The final pair (e, f) is automatically formed. Now, let us assume that b rejects the offer in the first step of the first stage. Then, in the second step of the first stage a offers to c . If c rejects the offer, then d will be called in the third step and surely he will accept the offer which means c will lose his chance for both of a and d . So, c accepts the offer and forms the pair (a, c) . In the second stage, b will be the restricter and he offers to d . d will definitely accept and form the pair (b, d) . And, the last pair is (e, f) . Now, let us back to the beginning of the first stage. If b accepts the offer of a , he forms the pair (a, b) . If he rejects a 's offer, then he forms the pair (b, d) . Since, b prefers d over a , he rejects the offer of a at the first stage and so we end up with matching $\mu = \{(a, c), (b, d), (e, f)\}$, which is the only stable matching for \mathbb{R}_2 .

Secondly, let us pick b as the restricter of the first stage. It is easy to show that for c rejecting b 's offer is a dominant strategy and we end up with the same matching μ .

Example 13 showed that mechanism β finds a stable matching for \mathbb{R}_2 . From Theorem 8 in section 3.1, this result is not unexpected. \mathbb{R}_2 has a stable matching and our mechanism finds it.

But, what about \mathbb{R}_1 , the example described by Gale and Shapley? What do we observe if we apply our mechanism to some profile that does not have any stable matching?

Example 14 *We will study \mathbb{R}_1 .*

$$\mathbb{R}_1^* = \begin{array}{ccccc} \underline{a} & \underline{b} & \underline{c} & \underline{d} & \underline{B(\mathbb{R}_1)} \\ b & c & a & c & c \\ c & a & b & b & b \\ d & d & d & a & a \\ & & & & d \end{array}$$

In the first stage, c will be the restricter. a is called to make a decision. If a accepts the offer, he forms the pair (a, c) . Then, the other pair will be (b, d) . If a rejects, then b will be called. Definitely b accepts the offer and forms (b, c) . Then, the other pair will automatically be (a, d) . So, at the beginning of the stage, if a accepts the offer, then he matches with c . If a rejects, then his mate will be d . Since a prefers c over d , he accepts the offer of c at the first stage of the game. Then, we end up with matching $\mu_2 = \{(a, c), (b, d)\}$, which is unstable.

We applied mechanism β to \mathbb{R}_1 and we ended up with an unstable matching. This result is not a surprise; we knew that there is no stable matching for this profile. The unexpected point is the behaviour of our mechanism. In section 3.1, Theorem 8 (Theorem 13 in [Evcı \(2014\)](#)) shows that γ is a stable mechanism; it always finds a stable matching. The surprise part is that as if there was some stable matching for \mathbb{R}_1 , the procedure was very smooth. But, it gave an unstable

matching in the end. Then, what is the mystery of \mathbb{R}_1 ? We will answer this question later in this section.

Now, we shall work on another example.

Example 15 Let $N = \{a, b, c, d, e, f\}$ be the set boys with preference profile \mathbb{R}_3 ,

	a	b	c	d	e	f	$B(\mathbb{R}_3)$
$\mathbb{R}_3^* =$	c	c	d	a	a	b	a
	b	d	a	c	b	a	c
	d	a	b	b	c	c	b
	e	e	f	f	d	d	d
	f	f	e	e	f	e	ef

In the first stage, a will be the restricter. c is called to make a decision. If c accepts the offer, he forms the pair (a, c) . In the second stage b becomes the restricter. d is called for an offer and d accepts the offer since a is deleted from the profile and so there is no better mate remained. Then, he forms the pair (b, d) . The final pair (e, f) is automatically formed. And, we end up with matching $\mu_x = \{(a, c), (b, d), (e, f)\}$.

Now, let us assume that c rejects the offer in the first stage. Then, a offers to b in the second step of the first stage. If b accepts the offer, he forms the pair (a, b) . In the second stage c becomes the restricter. Then, c offers to d and d accepts the offer since a is deleted from the profile. Then, he forms the pair (c, d) . The final pair (e, f) is automatically formed. Then, we end up with matching $\mu_y = \{(a, b), (c, d), (e, f)\}$.

If b rejects the offer in the second step of the first stage, then d will be called and surely he will definitely accept the offer and forms the pair (a, d) . In the second stage c becomes the restricter. Then, c offers to b and b definitely accepts the offer. Then, he forms the pair (b, c) . The final pair (e, f) is automatically formed. Then, we end up with matching $\mu_z = \{(a, d), (b, c), (e, f)\}$.

In the second step of the first stage (after c rejects a 's offer), if b accepts a 's offer, we end up with μ_y . If he rejects a 's offer, we end up with μ_z . Since $\mu_z R_b \mu_y$, b rejects the offer of a .

In the first step of the first stage, if c accepts a 's offer, we end up with μ_x . If c rejects a 's offer, we end up with μ_z . Since $\mu_x R_c \mu_z$, c accepts the offer by a .

Hence, we end up with matching μ_x , which is unstable for \mathbb{R}_3 since it is blocked by (c, d) .

We applied our mechanism β to \mathbb{R}_3 and we got an unstable matching. Is \mathbb{R}_3 one of the preference profiles which do not have any stable solution?

The answer is "No!". \mathbb{R}_3 has absolutely and only one stable matching and it is μ_y .

As [Evcı \(2014\)](#) has stated and proved, the mechanism γ is stable. β is stronger than γ , since there is only one queue and the chooser agents never experience any conflict. Then, why cannot β end up with a/the stable matching while there exist some?

The definition below will help us to figure out the reason.

Definition 16 Let N be a set of agents. Let \mathbb{R} be a preference profile and $B(\mathbb{R})$ is the corresponding Borda ranking. Let $M \subset N$ be a proper subset of N . The preference profile $\hat{\mathbb{R}}$ of M is constructed by deleting the agents $i \in N/M$ in \mathbb{R} . Namely, $\hat{\mathbb{R}}$ is the profile of M **Purified from Irrelevant Alternatives (PIA)** of N/M and $B(\hat{\mathbb{R}})$ is the corresponding Borda ranking.

In the next example, we will examine \mathbb{R}_3 with purified Borda ranking.

<u>a</u>	<u>b</u>	<u>c</u>	<u>d</u>	<u>e</u>	<u>f</u>	<u>$B(\mathbb{R}_3)$</u>	<u>$B(\hat{\mathbb{R}}_3)$</u>
c	c	d	a	a	b	a	c
b	d	a	c	b	a	c	ad
d	a	b	b	c	c	b	b
e	e	f	f	d	d	d	ef
f	f	e	e	f	e	ef	

Example 17 $\hat{\mathbb{R}}_3 =$

In \mathbb{R}_3 , agents e and f are the worst alternatives for the rest of the society and for each other they are same. If we purify \mathbb{R}_3 by excluding e and f , we get the relationships of $\{a, b, c, d\}$ with each other, as seen in profile $\mathbb{R}_3^{\{a, b, c, d\}}$ below,

<u>a</u>	<u>b</u>	<u>c</u>	<u>d</u>	<u>$B(\hat{\mathbb{R}}_3^{\{a, b, c, d\}})$</u>
c	c	d	a	c
b	d	a	c	ad
d	a	b	b	b

The comparison of $B(\mathbb{R}_3)$ and $B(\hat{\mathbb{R}}_3)$ tells us that eventhough a is not the most favorite member of the set $\{a, b, c, d\}$, the support from an "irrelevant" set $\{e, f\}$ makes him the best of the profile, which falsifies the result of the game. The fake position of a makes him get a better mate which leads the game to an unstable solution.

If we start the game with a , then we get $\mu_x = \{(a, c), (b, d), (e, f)\}$. On the other hand, starting the game with c gives $\mu_y = \{(a, b), (c, d), (e, f)\}$, which is the only stable matching for \mathbb{R}_3 .

Now, we will check the case for \mathbb{R}_2 with purified ranking.

<u>a</u>	<u>b</u>	<u>c</u>	<u>d</u>	<u>e</u>	<u>f</u>	<u>$B(\mathbb{R}_2)$</u>	<u>$B(\hat{\mathbb{R}}_2)$</u>
b	c	d	a	a	b	ab	abcd
c	d	a	b	b	a	c	ef
d	a	b	c	c	c	d	
e	e	f	f	d	d	ef	
f	f	e	e	f	e		

Example 18 $\hat{\mathbb{R}}_2 =$

Since $B(\mathbb{R}_2)$'s top cycle $\{a, b\}$ is included by the one $\{a, b, c, d\}$ of $B(\hat{\mathbb{R}}_2)$, \mathbb{R}_2^* with unpurified Borda ranking gave the stable matching. It is easy to show that starting the game with c or d would give the same (only) stable matching.

Now, we will examine profile \mathbb{R}_1 , the example described by Gale and Shapley, but with purified orderings.

<u>a</u>	<u>b</u>	<u>c</u>	<u>d</u>	<u>$B(\mathbb{R}_1)$</u>	<u>$B(\hat{\mathbb{R}}_1)$</u>
b	c	a	c	c	abc
c	a	b	b	b	d
d	d	d	a	a	
				d	

Example 19 $\hat{\mathbb{R}}_1 =$

Since there is a cycle between $\{a, b, c\}$, we randomly pick one of the agents and assign him as the restricter. If we start with a , we end up with $\mu_1 = \{(a, b), (c, d)\}$. Starting with b gives $\mu_3 = \{(a, d), (b, c)\}$. And, finally if c is the restricter of the first stage, the game reaches $\mu_2 = \{(a, c), (b, d)\}$. As we have already said, none of them is stable.

From above example, the following question arises; what is the stability condition of the roommate problem (in terms of β)?

Tan (1991) already stated the stability condition for the roommate problem. The following theorem provides the necessary and sufficient condition of stability in our game scenario.

Theorem 20 *Let N be a society and $\hat{\mathbb{R}}$ be their preference profile with purified ordering. The profile \mathbb{R} does not have any stable matching if and only if in the game for $\hat{\mathbb{R}}$, mechanism β confronts a top cycle with odd number of agents in a subgame.*

Proof. (\Leftarrow). Let $M \subset N$ be a set of agents of a cycle with odd number of cardinality $(2k - 1)$. Every agent prefers other agents in the cycle to the agents outside of the cycle. From the structure of the game and the cycles, $(2k - 2)$ number of agents of M match with each other. So, one of them forms a pair with an agent from the bottom set. The existence of an agent in the bottom set is guaranteed by the number of agents in the set N , that is $Card(N) = 2n$. Let $i \in M$ be that agent and $j \in N/M$ be his mate. From the definition of a Borda top cycle, it is clear $lR_i j$ where $\forall l \in M/\{i\}$. And, again from definition of a cycle, there is an agent $k \in M/\{i\}$ such that $iR_k \mu(k)$. Otherwise, agents of M would not construct a full cycle. So, the pair (k, i) blocks the matching.

(\Rightarrow). We suppose that \mathbb{R} does not have any stable matching and there is no cycle with odd number of agents in any subgame of β . We will show that this leads to a contradiction.

Firstly, let us assume that there is no cycle at all. So, $B(\hat{\mathbb{R}})$ is a one-ranking sequence of agents. Let (i, j) be a blocking pair. Without loss of generality, let us assume i has a higher ranking than j does in $B(\hat{\mathbb{R}})$. Since (i, j) blocks the matching, i has a mate k such that $jR_i k$ and also j has a mate l such that $iR_j l$. (i, j) being a blocking pair means that until j 's turn, i has not been taken. Since, we have (j, l) , j has not chosen any agent until his turn for i , because better agents are not achievable for him. And, finally, when it is his turn, he does not choose i and in a later stage he matches with l . Eventhough j has a chance, he does not choose i , which contradicts to the rationality axiom.

Secondly, let us assume that the game consists of cycle(s) with even number of agents. From the game scenario of γ and the definition of a (top Borda) cycle, in such cycles agents match with other in the same cycle in the way the tie is broken. This contradicts to instability.

Finally, if the game is combination of two above cases, same arguments work.

■

The final topic of this paper is multi stability. As [Knuth \(1976\)](#) showed that for the roommate problem some preference profiles have more than one stable matching like for the marriage problem. In the following pages we will analyze multi stability.

Next example is on a profile with multi stability and the outcome of mechanism β for this profile.

Example 21 Let $N = \{a, b, c, d\}$ be any set of boys with the profile \mathbb{R}_4 ,

	<u>a</u>	<u>b</u>	<u>c</u>	<u>d</u>	<u>$B(\mathbb{R}_4)$</u>	<u>$B(\hat{\mathbb{R}}_4^{\{a,b,c\}})$</u>	<u>$B(\hat{\mathbb{R}}_4^{\{a,b,d\}})$</u>	<u>$B(\hat{\mathbb{R}}_4^{\{a,c,d\}})$</u>	<u>$B(\hat{\mathbb{R}}_4^{\{b,c,d\}})$</u>
$\mathbb{R}_4 =$	b	d	a	c	$abcd$	a	b	c	d
	c	a	d	b		b	d	a	c
	d	c	b	a		c	a	d	b

$\Pi^N = \{\mu_1, \mu_2, \mu_3\}$ be the set of all possible matchings, where

$$\mu_1 = \{(a, b), (c, d)\},$$

$$\mu_2 = \{(a, c), (b, d)\},$$

$$\mu_3 = \{(a, d), (b, c)\}.$$

As seen from the table above, profile $\hat{\mathbb{R}}_4$ has no fixed purified Borda ranking like $\hat{\mathbb{R}}_3$ does. So, any agent could be the restricter of the first stage.

For profile \mathbb{R}_4 , the set of stable matchings is $\{\mu_1, \mu_2\}$. It is easy to show that if we start the game with b or c , we end up with matching μ_1 . On the other

hand, starting with a or d gives us matching μ_2 . This is because of the **cycle conflict** between the sets $\{a, d\}$ and $\{b, c\}$.

In the existence of multi stability, we need to run the mechanism for different (all) purified orderings in order to find all of stable matchings for the preference profile. Giving priorities in Borda ranking to different agents changes the outcome from one stable solution to another.

The reason of multi stability is, not surprisingly, the existence of the cyclical conflicts between two disjoint sets of agents. We refer to [Evcı \(2014\)](#) for an exhaustive analysis of cycles.

In the game scenario of γ of [Evcı \(2014\)](#), choosers have to decide simultaneously at the same step. Having the same decision of the agents in a cycle generates two stable matchings. At such nodes, Nash equilibrium is a binding criterion for the choosers in cycles.

In the game scenario of β on the other hand, cyclical conflicts are broken since the game path (tree) has only one single ranking. The choosers are called one by one to make a decision and so they do not experience any cyclical conflict. Therefore, β generates a bias for the chooser agents and the outcome is always chooser-optimal, as also seen in Example 21.

[Evcı \(2014\)](#) also proposed a refinement mechanism α of his original mechanism γ . While γ implements the full set of stable matchings for any given preference profile, α partitions the full domain of preference profiles. For some profiles the mechanism implements the full set of stable matchings; for some profiles, it gives a proper subset of stable matchings and for some of them, it induces (Chooser-Optimal) Gale and Shapley's algorithm. Thus, while mechanism α generates "**partial bias**" for the chooser side as compared to mechanism γ , mechanism β generates "**full bias**".

Now, we state our most general result.

Theorem 22 *Let N be a set of agents and \mathbb{R} be their preference profile. Let $B(\hat{\mathbb{R}})$ be the corresponding purified ordering(s). The mechanism β defined over $\hat{\mathbb{R}}$ and $B(\hat{\mathbb{R}})$, formally*

$$\beta : (\mathbb{R}, B(\hat{\mathbb{R}})) \longrightarrow \Pi^N,$$

implements the full set of stable matchings in the existence of stability and gives a Pareto Efficient matching for the instances of instability.

Proof. The stability part has been proved by the examples, claims, proposition and theorems so far.

Pareto efficiency is proved from the definition of a cycle with odd number of agents. In a cycle, increasing the "payoff" of an agent, not a member of the blocking pair, reduces the payoff of another agent.

Secondly, we shall specifically analyze the blocking pair. Let (i, j) be the blocking pair such that i is matched with agent k from the bottom set. Re-matching of agents in a way that in the new matching we have (i, j) , such that jR_ik and j is a member of the top cycle, generates a new blocking pair since another agent, say l , who is another member of the top cycle, will be matched with k . And, this means the payoff of agent l reduces. Hence, the matching from the procedure is Pareto efficient. ■

4 Conclusion

In this paper, we propose a simple mechanism for the roommate problem. The mechanism is a refinement of the mechanism described in [Evcı \(2014\)](#). While applying his mechanism to this problem, we benefit from a famous Social Choice Rule (SCR), the Borda rule in welfare function form. Then, we analyze the effect of this SWF in two scenarios by simply separating the raw and purified orderings.

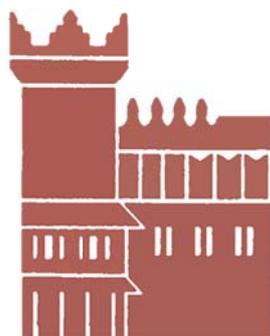
First of all, as we show, the mechanism β is quite successful under the purified orderings in checking stability for any given preference profile. The success of β for the raw orderings depends on whether it coincides with the purified orderings or not. As long as the top set of the raw orderings is a subset of the one generated under purified orderings, we end up with a/the stable matching.

We have showed that in the absence of stability, the mechanism β ends up with a Pareto efficient matching.

And finally, we have showed that β is also an easy and strong mechanism to find all stable matchings for a given preference profile in the existence of multi stability.

References

- [1] Cechlarova, K. & Fleiner, T. (2005), "On a Generalization of the Stable Roommates Problem", *ACM Transactions on Algorithms*, 1(1): 143-156.
- [2] Evcı, B. (2014), "A New Dynamic Mechanism to the Marriage Problem with a variant", Unpublished manuscript, Available online: <http://www2.dse.unibo.it/wp/WP973.pdf>
- [3] Gale, D. & Shapley, L.S. (1962), "College Admissions and the Stability of Marriage", *The American Mathematical Monthly*, 69: 9-15.
- [4] Irving, R.W. (1985), "An Efficient Algorithm for the Stable Roommates Problem", *Journal of Algorithms* 6, 577-595.
- [5] Knuth, D.E. (1976), "Mariages Stables et leurs relations avec d'autres problèmes combinatoires", *Les Presses de l'Université Montréal, Montréal*.
- [6] Tan, J.J.M. (1991), "A Necessary and Sufficient Condition for the Existence of a Complete Stable Matching", *Journal of Algorithms*, 12:154-178.
- [7] Tan, J.J.M. & Hsueh, YC. (1995), "A Generalization of the Stable Matching Problem", *Discrete Applied Mathematics*, 59:87-102.



Alma Mater Studiorum - Università di Bologna
DEPARTMENT OF ECONOMICS

Strada Maggiore 45
40125 Bologna - Italy
Tel. +39 051 2092604
Fax +39 051 2092664
<http://www.dse.unibo.it>