ThE Papers

Dpto. Teoría e Historia Económica
Universidad de Granada

Working Paper n. 14/04

A new solution for the roommate problem: The Q-stable matchings

Péter Biró, Elena Iñarra and Elena Molis

September, 2014

A new solution for the roommate problem: The $\mathcal{Q}$-stable matchings*

Péter Biró†, Elena Iñarra‡, and Elena Molis§

September 16, 2014

Abstract

The aim of this paper is to propose a new solution for the roommate problem with strict preferences. We introduce the solution of maximum irreversibility and consider almost stable matchings (Abraham *et al.* [2]) and maximum stable matchings (Tan [30] [32]). We find that almost stable matchings are incompatible with the other two solutions. Hence, to solve the roommate problem we propose matchings that lie at the intersection of the maximum irreversible matchings and maximum stable matchings, which are called Q-stable matchings. These matchings are core consistent and we offer an efficient algorithm for computing one of them. The outcome of the algorithm belongs to an absorbing set.

*This research is supported by the Spanish Ministry of Science and Innovation (ECO2010-17049 and ECO2012-31346), co-funded by ERDF, by Basque Government IT-568-13 and by the Government of Andalusia Project for Excellence in Research (P07.SEJ.02547). Péter Biró also acknowledges the support from the Hungarian Academy of Sciences under its Momentum Programme (LD-004/2010), and the Hungarian Scientific Research Fund, OTKA, grant no. K108673.

†Institute of Economics, Research Centre for Economic and Regional Studies, Hungarian Academy of Sciences, H-1112, Budaörsi út 45, Budapest, Hungary. Email: peter.biro@krtk.mta.hu. He a visiting professor at Department of Economics, Stanford University in 2014.

‡*BRiDGE,* University of the Basque Country (UPV/EHU). and Instituto de Economía Pública. Email: elena.inarra@ehu.es

§GLOBE, Departamento de Teoría e Historia Económica, Universidad de Granada.. Email: emolis@ugr.es

1 Introduction

Gale and Shapley [12] introduce one-to-one stable matching problems. They first define, the marriage problem, a two-sided matching problem in which agents are divided into two disjoint groups (*e.g.* men and women) and any agent can only be matched to an agent in the other group. Then, they proceed to the roommate problem, a one-sided matching problem in which all agents belong to a single group and any agent can be matched to any other. The authors propose stable matchings as a solution for these problems. A matching is stable if no two agents prefer one another to their current partners. They show that for the marriage problem there is always a stable matching, but there may not be for the roommate problem. The following example, slightly modified from the original in Gale and Shapley [12], illustrates this case.

Example 1 *Consider the following 4-agent problem:*

a_1	a_2	a_3	a_4
a_2	a_3	a_1	a_1
a_3	a_1	a_2	a_2
a_4	a_4	a_4	a_3
a_1	a_2	a_3	a_4

To see that there is no stable matching, assume that one of the first three agents, say agent a_3, is either unmatched or matched to agent a_4. Then, agents a_2 and a_3 form a blocking pair. Similar arguments can be applied to agents a_1 and a_2.

In the last few decades an extensive literature on one-to-one matching problems has emerged in both Economics and Computer Science. However, it focuses mostly on the marriage problem: the roommate problem has been much less widely studied. This can be explained by two reasons: First, there are more economic issues that can be modeled as two-sided problems than as one-sided ones. Second, the impossibility of finding a stable matching and the more complex structure of the roommate problem may have discouraged researchers from analyzing it.

Pairing police officers on patrols, pilots on flights (see Cechlárová and Ferková [8]), students to share double rooms in colleges or marriages between agents of the same-sex poses significant problems worthy of analysis. In sports competitions the way in which players are paired in events such as tennis or paddle-tennis doubles may affect the final result. The kidney exchange problem has

been modeled as a roommate problem (see Roth *et al.* [27]). Furthermore there are centrally coordinated programs such as odd shoe exchanges,[1] holidays home exchanges[2], and centralized pairing methods used in chess competitions (see Kujansuu *et al.* [21]), which also suggest potential applications of the roommate problem. Moreover, as Klaus *et al.* [19] point out, the roommate problem has interest in itself since it boils down to hedonic coalition formation (see Bogomolnaia and Jackson [6]) and network formation problems (see Jackson and Watts [17]).

In the scarce literature about one-sided matching problems it is common practice to restrict the analysis to those problems in which a stable matching exists (solvable problems) (see for instance, Gusfield and Irving [14], Chung [9], Diamantoudi *et al.* [10], Klaus and Klijn [19] and Gudmundsson [13]). However, restricting attention to solvable roommate problems means ignoring a significant subclass of problems without stable matchings (unsolvable problems). This is corroborated by Pittel and Irving [25], who observe that as the number of agents increases the probability of a roommate problem being unsolvable also increases fairly steeply.

The aim of the current paper is to propose a new solution for the roommate problem with strict preferences.[3] Indeed it is essential to require a solution which provides a stable matching when dealing with solvable problems and some matching otherwise. Hence we focus on core consistent solutions.[4] At the interface between Economics and Computer Science several solutions have been proposed explicitly for dealing with unsolvable problems, but there has yet to be any in-depth discussion regarding comparisons between solutions and scope for new ones.

Two interesting core consistent solutions have been analyzed in the literature on unsolvable problems: *Almost stable matchings*, proposed by Abraham *et al.* [2], form a subclass of Pareto optimal matchings with the minimum number of blocking pairs.[5] The notion of *maximum internal stability* introduced by Tan [30], singles out matchings with the largest set of pairs that are stable one

[1] http://www.oddshoe.org/

[2] http://www.exchangeholidayhomes.com/

[3] Except for small number of references, we have deliberately avoided the analysis of a variety of roommate problem reappraisals.

[4] Other solutions have been proposed in the literature, for instance popular matchings. However, this solution is not core-consistent and for that reason we have not considered it in our analysis. For more details of such solution see Biro et al. [5].

[5] See Biermann [3] for a critical evaluation of this criterion in the marriage problem.

with another. However, as far as we know, the following basic proposal has been overlooked. Consider the case in which two agents are top choices for each other. Once this pair is formed it never splits. A less extreme case is the existence of a set of agents who form pairings so stable that they are stably paired within them and none of them prefers an agent other than his/her current partner no matter how the outside agents are matched. Hence, once these pairs are formed they never break. Thus, we believe that a *maximum irreversible* set of pairs should form part of the matching selected to solve any roommate problem.

All three of the solutions mentioned show sufficient grounds for consideration as good candidates for solving roommate problems. It thus makes sense to consider a proposal that could conciliate most if not all of those solutions.

By studying the relationship between these solutions we find that the almost stable solution is incompatible with the other two. Moreover the problem of finding a matching with the minimum number of blocking pairs happens to be NP-hard. Hence, our next move is to search a solution that could conciliate the notions of maximum internal stability and maximum irreversibility. Accordingly, we select the set of matchings that lie at the intersection of the two solutions and refer to them as Q-stable matchings. Since our ultimate motive is to select a single matching to solve the roommate problem, an essential criterion to take into account is the possibility of determining a Q-stable matching. We offer an efficient algorithm for computing such a matching.

Finally, we seek to extend what we have learnt from two-sided matching problems to one-sided matching problems. In two-sided matching problems if agents interact freely and decide systematically after a match what to do next then they eventually reach a stable matching. It is also known that market frictions may prevent a stable matching from being reached.[6] This justifies the presence of clearinghouses[7] where agents submit preference lists to a policy maker who, following a procedure, implements the desired matching. Similarly in one-sided matchings it could be interesting to identify a matching resulting from a decentralized process. For the roommate problem it is known that the blocking dynamic between agents leads to an absorbing set of matchings (see Inarra *et al.* [16] and Klaus *et al.* [20]). In fact, once one of these matchings

[6] A well-known documented episode of unraveling in matching markets for medical interns shows that contracts for interns were signed two years earlier than students' graduation (see, for instance, Echenique and Pereyra [11] and the references therein).

[7] For instance, the National Resident Matching Program (NRMP) matches physicians and residency programs in the United States.

has been reached the blocking dynamic of the agents does not allow them to abandon that set. Hence, we proceed by studying whether our proposal is one of the elements of an absorbing set. We find that not all Q-stable matchings belong to an absorbing set, but the matching determined by the algorithm does. Therefore we are providing policy makers with a procedure that implements a Q-stable matching for solving roommate problems.

The rest of the paper is organized as follows: Section 2 contains the preliminaries. Section 3 presents and discusses the notion of maximum irreversibility and the other two solutions found in the relevant literature for unsolvable roommate problems, and proceeds to compare these three core-consistent solutions. Section 4 introduces Q-stable matchings and an algorithm for computing one of them. We also show that such a matching belongs to an absorbing set. Section 5 concludes.

2 Preliminaries

In a roommate problem, a finite set of agents $N = \{a_1, \ldots, a_n\}$ has to be partitioned into pairs and singletons. Each agent has strict preference over potential roommates with the possibility of having a room to herself/himself. Formally, a *roommate problem*, or a *problem* for short is a pair $(N, (\succeq_{a_i})_{a_i \in N})$ (or $(N, \succeq)$ for short) where N is a finite set of agents and for each agent $a_i \in N$, $\succeq_{a_i}$ is a complete, transitive preference relation defined on N. Preferences are strict, *i.e.*, $a_k \succeq_{a_i} a_j$ and $a_j \succeq_{a_i} a_k$ if and only if $a_j = a_k$. The strict preference relation associated with $\succeq_{a_i}$ is denoted by $\succ_{a_i}$. Agent a_j is *acceptable* for agent a_i if $a_j \succ_{a_i} a_i$. Otherwise he/she is said to be *unacceptable*. A solution to a problem, a *matching*, is a function $\mu : N \to N$ such that if $\mu(a_i) = a_j$ then $\mu(a_j) = a_i$. Thus, a matching is a set of disjoint pairs and singletons formed by the agents in N. Let $\mu(a_i)$ denotes the partner of agent a_i in matching μ. If $\mu(a_i) = a_i$, then agent a_i is unmatched in μ. A matching μ with all its agents paired is called *complete*. Given $S \subseteq N$, $S \neq \emptyset$, let $\mu(S) = \{\mu(a_i) : a_i \in S\}$. That is, $\mu(S)$ is the set of partners of the agents in S under matching μ. Let $\mu \mid_S$ denotes the restriction of μ to agents in S. If $\mu(S) = S$, then $\mu \mid_S$ is a matching in $(S, (\succ_{a_i})_{a_i \in S})$.

A matching μ is *blocked by a pair* $\{a_i, a_j\} \subseteq N$ if $a_j \succ_{a_i} \mu(a_i)$ and $a_i \succ_{a_j} \mu(a_j)$, that is a_i and a_j prefer each other to their current partners (if any) in

μ. If pair $\{a_i, a_j\}$ blocks matching μ then $\{a_i, a_j\}$ is called a *blocking pair* of μ. Let $\{a_i, a_j\}$ blocks a matching μ. A matching μ' is obtained from μ by *satisfying* $\{a_i, a_j\}$ if $\mu'(a_i) = a_j$, their partners (if any) under μ are alone in μ', and the remaining agents are matched as in μ. A matching without blocking pairs is called a *stable* matching. A problem is called *solvable* if the set of stable matchings is non-empty and *unsolvable* otherwise.

We extend each agent's preferences over his/her potential partners to the set of matchings in the following way. We say that agent a_i prefers μ' to μ, and denote it by $\mu' \succ_i \mu$ if and only if agent a_i prefers his/her partner at μ' to her partner at μ, $\mu'(a_i) \succ_{a_i} \mu(a_i)$. (We say that agent a_i is indifferent between matchings μ' and μ, denoted by $\mu' \sim_{a_i} \mu$ if he/she is matched to the same partner in both matchings).

Stable partitions

Tan [31] establishes the necessary and sufficient condition for the solvability of a problem with strict preferences using the notion of stable partition which is formally defined as follows:

Let $A = \{a_1, ..., a_k\} \subseteq N$ be an ordered set of agents. The set A is a *ring* if $k \geq 3$ and for all $i \in \{1, ..., k\}$, $a_{i+1} \succ_{a_i} a_{i-1} \succ_{a_i} a_i$ (subscript modulo k). The set A is a *pair of mutually acceptable agents* if $k = 2$ and for all $i \in \{1, 2\}$, $a_{i-1} \succ_{a_i} a_i$ (subscript modulo 2). The set A is a *singleton* if $k = 1$.

A *stable partition* is a partition $\mathcal{P}$ of N such that:
(i) For all $A \in \mathcal{P}$, the set A is a ring, a pair of mutually acceptable agents or a singleton, and
(ii) For all $A, B \in \mathcal{P}$ where $A = \{a_1, ..., a_k\}$ and $B = \{b_1, ..., b_l\}$ (possibly $A = B$), the following condition holds:

$$\text{if } b_j \succ_{a_i} a_{i-1} \text{ then } b_{j-1} \succ_{b_j} a_i,$$

for all $i \in \{1, ..., k\}$ and $j \in \{1, ..., l\}$ such that $b_j \neq a_{i+1}$.

Thus, a stable partition is a partition of the set of agents such that each set in a stable partition is either a ring, a pair of mutually acceptable individuals, or a singleton, and the partition satisfies the (usual) stability condition between any two sets and also within each set.[8] The following assertions are proven by

[8]Stable partitions are also called *stable half-matchings* in some recent papers, such as Biró *et al.* [4]. A half-matching is a well- known notion in graph theory that also helps to understand the meaning of this notion.

Tan [31].

Remark 1 *(i) A problem $(N, \succ)$ has no stable matchings if and only if there exists a stable partition with an odd ring. (ii) All stable partitions have exactly the same odd rings and singletons. (iii) All even rings in a stable partition can be broken into pairs of mutually acceptable agents without upsetting stability.*

W.l.o.g. hereafter we suppose that the even parties are always pairs in any stable partition we are working with.

The notion of stable partition plays a significant role in the present work and it is not that easy to interpret. Hence, in the appendix, we informally describe the algorithm introduced by Tan and Hsueh [33] for computing stable patitions, and illustrate it with a numerical example which we believe clarifies its meaning.

3 Core consistent solutions

In this section we first introduce a notion of strong stability that we believe is suitable for consideration in the search for a matching that is as stable as possible. Then we consider two existing proposals for dealing with unsolvable problems.

3.1 Maximum irreversibility

Consider the case in which two agents are top choices for each other. Once this pair is formed it never splits.[9] A less extreme case is the existence of a set of agents forming a pairing so strongly stable that they are stably paired within them and none of them prefers an agent outside to her current partner no matter how the outside agents are matched. Therefore, once these pairs are formed they never break. We call this set of pairs "irreversible".

A problem may have matchings with different irreversible sets but it seems natural to require that some of the largest ones be contained in the proposed matching. Formally,

Definition 1 *(i) A set of agents $S \subseteq N$ form an* irreversible *set of pairs μ_S if there is no pair $\{a_i, a_j\}$ (possibly $a_i = a_j$) such that $\{a_i, a_j\} \cap S \neq \emptyset$ such that $\{a_i, a_j\}$ blocks μ_S. (ii) Matching μ is* maximum irreversible *if it contains the largest irreversible set of pairs.*

[9] This property, called "mutually best" property was introduced by Toda [34] for the marriage problem and by Can and Klaus [7] for the roommate problem.

The set of maximum irreversible matchings is core-consistent and the larger the set of irreversible pairs is, the more selective this criterion will be. To see the robustness of this solution consider the following example:

Example 2 *Consider the following* 10*-agent problem:*

a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}
a_2	a_3	a_1	a_8	a_9	a_4	a_6	a_5	a_7	a_1
a_3	a_1	a_2	a_9	a_1	a_5	a_8	a_7	a_4	a_2
a_4	a_{10}	a_4	a_3	a_6	a_7	a_9	a_4	a_5	a_3
a_6	a_5	a_5	a_6	a_8	a_8	a_5	a_6	a_1	a_4
a_5	a_6	a_6	a_7	a_7	a_9	a_1	a_1	a_2	a_5
a_7	a_7	a_7	a_{10}	a_4	a_2	a_2	a_2	a_3	a_6
a_8	a_8	a_8	a_1	a_2	a_3	a_3	a_9	a_6	a_7
a_9	a_4	a_9	a_2	a_3	a_1	a_4	a_{10}	a_8	a_8
a_{10}	a_9	a_{10}	a_5	a_{10}	a_{10}	a_{10}	a_3	a_{10}	a_9
a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}

Matching $\mu_1 = \{\{a_1, a_2\}, \{a_3\}, \{a_4, a_8\}, \{a_5, a_9\}\{a_7, a_6\}, \{a_{10}\}\}$ *is maximum irreversible, with an irreversible set of three pairs. The pairing* $\{\{a_4, a_8\}, \{a_5, a_9\}, \{a_7, a_6\}\}$ *is stable and no agent in it prefers any other outside agent to his/her current partner. Hence, once this pairing is formed these pairs stay together.*

However, for unsolvable problems the set of irreversible pairs of a problem might be empty. Hence, this criterion by itself might be indefinite. In what follows, we present two core consistent solutions that have been proposed in the relevant literature.

3.2 Two previous solution concepts

Almost stability

Pareto optimality, one of the most relevant criteria in Economics, has also been applied to the roommate problem and to a number of extensions of it. In our setup it can be defined as follows:

Definition 2 *A matching* μ *is* Pareto optimal *if there is no matching* μ' *such that* $\mu'(a_i) \succeq_{a_i} \mu(a_i)$ *for all* $a_i \in N$ *and* $\mu'(a_i) \succ_{a_i} \mu(a_i)$ *for some* $a_i \in N$.

There is no doubt that Pareto optimality is an appealing criterion related to stability. A stable matching is Pareto optimal (see *Proposition 5* in Abraham and Manlove [1]) and a non-Pareto optimal matching is always unstable.

This is because it is blocked by a set of agents who are better off in another matching.[10] However, Pareto optimality by itself is not a convincing criterion for selecting matchings in the roommate problem for two reasons. First, it requires that when any two agents block a matching by forming a new pair their partners, if any, must not be worse off in the new matching. In our setting, however, it suffices for two agents to become better off by forming a blocking pair, without considering the wellbeing of their abandoned partners in the new matching formed.[11] Secondly, it can select too many matchings: For solvable problems, the Pareto optimal solution is core-inclusive, that is, it selects all stable matchings and some unstable ones. For unsolvable problems, it suffers from a similar drawback; it can select matchings with different numbers of blocking pairs.

The idea of refining the set of Pareto optimal matchings is taken up by Abraham and Manlove [1], who prove the following:

Remark 2 *Let $bp(\mu)$ denote the set of blocking pairs of matching μ, that is $bp(\mu) = \{\{a_i, a_j\} \subseteq N : \{a_i, a_j\}$ blocks $\mu\}$ (i) If matching μ Pareto dominates matching μ' then $bp(\mu) \subset bp(\mu')$ (ii) If μ is a matching with the minimum number of blocking pairs of a problem then μ is Pareto optimal.*

Following the idea behind the previous results, Abraham *et al.* [2] study matchings with the minimum number of blocking pairs and call them almost stable matchings. Formally,

Definition 3 *A matching μ is* almost stable *if $|bp(\mu)| \leq |bp(\mu')|$ for all $\mu' \neq \mu$, where $|bp(\mu)|$ denotes the number of blocking pairs of matching μ.*

Maximum internal stability

A matching μ is *maximum stable* if it excludes the minimum number of agents such that the non excluded ones form a complete stable matching see Tan [30], [32]. Given a stable partition, Tan [30] proposes that to compute a maximum stable matching one agent be deleted from each odd ring of the partition as well as all singletons. Then he defines the problem restricted to the set of non-deleted agents, keeping their original preferences. This new problem is solvable and the computation of a stable matching gives a maximum stable matching.

[10] If a matching is not Pareto optimal then it admits an improving coalition. See Proposition 6.24 in Manlove [22]) and Chapter 6 in this book for a survey on Pareto optimal matchings.

[11] This is not the case in problems in which bilateral approval is required to dissolve partnership see Morrill [23].

Example 2 (cont.) *In this problem,* $\mathcal{P} = \{\{a_1, a_2, a_3\},\{a_4, a_6\},\{a_5, a_8\},\{a_7, a_9\}, \{a_{10}\}\}$ *is a stable partition and all maximum stable matchings have four stable pairs. Matching* $\mu_1 \supseteq \{\{a_1, a_3\}, \{a_4, a_6\}, \{a_5, a_8\}, \{a_7, a_9\}\}$ *is maximum stable and it can be derived from Tan's proposal by isolating one agent from the odd ring, agent* a_2*, and the singleton, agent* a_{10}*. Apart from the maximum stable matchings obtained in this manner there may be others. For instance,* $\mu_2 \supseteq \{\{a_2, a_{10}\}, \{a_4, a_6\}, \{a_5, a_8\}, \{a_7, a_9\}\}$ *is also maximum stable.*

Tan's solution is applied to a setting in which a matching is defined as a set of pairs while isolated agents never form part of that matching. To adapt Tan's definition of a maximum stable matching to our setup, where a matching is a set of disjoint pairs and singletons formed by all the agents of a given set N, we must add the deleted agents as singletons to the maximum stable matching, so that all agents in the problem form part of that matching. Other possibility is to match some (or all) the singletons by forming pairs among them. All these matchings are equally close to stability in the sense that they contain the same number of stable pairs. Hence we extend the idea of maximum stability to our setting:

Definition 4 *(i) A set of agents* $T \subseteq N$ *form an* internally stable *set of pairs* μ_T *if there is no pair* $\{a_i, a_j\} \subseteq T$ *such that* $\{a_i, a_j\}$ *blocks* μ_T*. (ii) Matching* μ *is* maximum internally stable *if it contains the largest number of pairs which are internally stable.*

To see that our definition is an extension of Tan's definition let us consider the following example:

Example 2 (cont.) *Matching* $\mu = \{\{a_1, a_3\},\{a_2\},\{a_4, a_6\},\{a_5, a_8\},\{a_7, a_9\}, \{a_{10}\}\}$ *is maximum stable and maximum internally stable, while matching* $\mu' = \{\{a_1, a_3\},\{a_2, a_{10}\},\{a_4, a_6\},\{a_5, a_8\},\{a_7, a_9\}\}$ *is maximum internally stable but not maximum stable.*

Considering the notion of stable partition, say $\mathcal{P}$, introduced above, Inarra *et al.* [15] define matchings associated with that partition, called $\mathcal{P}$*-stable matchings*. These matchings are formally defined as follows:

Let $\mathcal{P}$ be a stable partition. A $\mathcal{P}$*-stable matching* is a matching such that for each set $A = \{a_1, ..., a_k\} \in \mathcal{P}$, agent a_i is paired with either a_{i+1} or a_{i-1} for all $i \in \{1, ..., k\}$ except for a unique agent j who remains unmatched if A

is odd or a singleton.[12] Hence, for each odd ring one agent is left out, and the rest of the agents in the ring are matched following its order, that is, they are matched to the subsequent or preceding agent. The reader may have noticed some similarities between $\mathcal{P}$-stable matchings and maximum stable matchings as defined above. It turns out that the set of $\mathcal{P}$-stable matchings coincides with the set of maximum internally stable matchings that can be computed by Tan's algorithm.

Remark 3 *A $\mathcal{P}$-stable matching is maximum internally stable.*

The following result states some features of $\mathcal{P}$-stable matchings obtained from the same stable partition $\mathcal{P}$.

Given two matchings μ and μ', we denote $\mu' R \mu$ if and only if μ' is obtained from μ by satisfying a blocking pair of μ (direct domination). We denote $\mu' R^T \mu$ if and only if there is a sequence of matchings $\mu = \mu_1, \ldots, \mu_k = \mu'$ such that for all $l \in \{1, \ldots, k-1\}$ μ_{l+1} is obtained from μ_l by satisfying a blocking pair of μ_l (indirect domination).

Remark 4 *Let $\mathcal{M}$ be the set of all matchings and let $\mathcal{P}$ be a stable partition. Consider $\mathcal{M}|_{\mathcal{P}} = \{\mu \in \mathcal{M} : \mu$ is a $\mathcal{P}$-stable matching of $\mathcal{P}\}$. Then (i) For any $\mu, \mu' \in \mathcal{M}|_{\mathcal{P}}$, $\mu R^T \mu'$. (ii) For all $a_i \in N$ belonging to an odd ring of $\mathcal{P}$, there exists a matching $\mu \in \mathcal{M}|_{\mathcal{P}}$ such that $\mu(a_i) = a_i$.*

$\mathcal{P}$-stable matchings play a significant role in Section 4 below.

3.3 Incompatibilities between solutions

In this subsection we analyze whether there is a solution that could reconcile all the solutions presented above. Unfortunately we find that there is not. In what follows we show that the solution of almost stable matchings is incompatible with both the other solutions.

To prove the incompatibility between almost stable matchings and maximum internally stable matchings, we start by showing the incompatibility between the latter and the family of Pareto optimal matchings in the following example.

Example 3 *Consider the following 8-agent problem:*

[12] Inarra *et al.* [15] show that from any matching there exists a sequence of blocking pairs converging a $\mathcal{P}$-stable matching. See Roth and Vande Vate [26] and Diamantoudi *et al.* [10] for similar approaches to convergence to stability.

a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8
a_2	a_3	a_1	a_1	a_1	a_7	a_8	a_6
a_3	a_1	a_2	a_2	a_2	a_8	a_6	a_7
a_4	a_4	a_4	a_3	a_3	a_4	a_4	a_4
a_5	a_5	a_5	a_6	a_6	a_5	a_5	a_5
a_6	a_6	a_6	a_7	a_7	a_1	a_1	a_1
a_7	a_7	a_7	a_8	a_8	a_2	a_2	a_2
a_8	a_8	a_8	a_5	a_4	a_3	a_3	a_3
a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8

This example has a unique stable partition $\mathcal{P}$= $\{\{a_1, a_2, a_3\},\{a_4, a_5\},\{a_6, a_7, a_8\}\}$ *with two odd rings* $\{a_1, a_2, a_3\}$ *and* $\{a_6, a_7, a_8\}$. *Every maximum internally stable matching has three stable pairs: one pair from each odd ring and the pair* $\{a_4, a_5\}$. *First, note that those matchings with singletons are not Pareto optimal since any other matching that joins them will Pareto dominate the original one. Hence we restrict our attention to those maximum internally stable matchings without singletons. There are nine such matchings with three stable pairs each:* μ_1= $\{\{a_1, a_2\},\{a_3, a_6\},\{a_4, a_5\}, \{a_7, a_8\}\}$; μ_2= $\{\{a_1, a_2\},\{a_3, a_7\}, \{a_4, a_5\}, \{a_6, a_8\}\}$; μ_3= $\{\{a_1, a_2\},\{a_3, a_8\},\{a_4, a_5\},\{a_6, a_7\}\}$; μ_4= $\{\{a_1, a_3\}, \{a_2, a_6\},\{a_4, a_5\},\{a_7, a_8\}\}$; μ_5= $\{\{a_1, a_3\},\{a_2, a_7\},\{a_4, a_5\},\{a_6, a_8\}\}$; μ_6= $\{\{a_1, a_3\}, \{a_2, a_8\},\{a_4, a_5\},\{a_6, a_7\}\}$; μ_7= $\{\{a_1, a_6\},\{a_2, a_3\},\{a_4, a_5\},\{a_7, a_8\}\}$; μ_8= $\{\{a_1, a_7\}, \{a_2, a_3\},\{a_4, a_5\},\{a_6, a_8\}\}$; μ_9= $\{\{a_1, a_8\},\{a_2, a_3\},\{a_4, a_5\},\{a_6, a_8\}\}$. *None of these matchings is Pareto optimal since in each of them agents in the pair* $\{a_4, a_5\}$ *and agents in the pair coming from different odd rings can become better off by changing their partners. Hence the following proposition can be established.*

Proposition 1 *The intersection of Pareto optimal matchings and maximum internally stable matchings may be empty.*

Since almost stable matchings are Pareto optimal the following corollary can be established.

Corollary 2 *The set of maximum internally matchings and the set of almost-stable matchings may have an empty intersection.*

This implies that the idea of finding a matching with the minimum number of blocking pairs conflicts with the idea of finding a matching with the maximum number of stable pairs. In fact the example shows the well known trade-off between Pareto optimality and stability.

Next, the incompatibility between almost stability and maximum irreversibility is shown in the following example.

Example 4 *Consider the following 8-agent problem:*

a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8
a_2	a_3	a_4	a_5	a_4	a_5	a_8	a_6
a_3	a_1	a_1	a_3	a_6	a_7	a_6	a_7
a_8	a_4	a_2	a_1	a_3	a_8	a_4	a_1
a_5	a_5	a_5	a_6	a_2	a_4	a_5	a_5
a_4	a_6	a_6	a_7	a_7	a_1	a_1	a_4
a_7	a_7	a_7	a_8	a_8	a_2	a_2	a_2
a_6	a_8	a_8	a_2	a_1	a_3	a_3	a_3
a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8

Matching $\mu = \{\{a_1, a_2\}, \{a_3, a_4\}, \{a_5, a_6\}, \{a_7, a_8\}\}$ *is the only one that is almost stable, with* $\{a_4, a_5\}$ *being its only blocking pair. However, pair* $\{a_4, a_5\}$ *is the maximum irreversible set and it is not contained in matching* μ*. Hence the following proposition can be established.*

Proposition 3 *The intersection of almost stable matchings and maximum irreversible matchings may be empty.*

Therefore almost stable matchings are incompatible with the other two solutions. Moreover Abraham *et al.* [2] show that the problem of finding a matching with the minimum number of blocking pairs is NP-hard and even hard to approximate. These results suggest that a proposal could be drawn up for the roommate problem by conciliating maximum internally stable and maximum irreversible matchings.

4 $\mathcal{Q}$-stable matchings

Given the incompatibility between almost stable matchings and the other two solutions demonstrated in the previous section, a natural question to ask is whether the intersection of the other three solutions is non-empty. The answer is yes. Matchings lying at the intersection comprise are called Q-stable matchings, which is analyzed in this section.

Definition 5 *A matching is* Q*-stable if it is maximum internally stable, and maximum irreversible.*

As mentioned in the Introduction our aim in this paper is to provide policy makers with a procedure for computing a Q-stable matching. In what follows we introduce an algorithm that does this job efficiently. The algorithm starts with a stable partition of a roommate problem and, by means of an iterative process, removes from the preference lists those agents who are unable to form irreversible pairs. Then a stable partition with a maximum set of irreversible pairs is derived from which a Q^*-stable matching is finally obtained.

Some additional notation is needed.

Given a stable partition $\mathcal{P}_t$. Let D_t be the set formed by the agents in the odd sets of $\mathcal{P}_t$, i.e. odd rings or singletons. Let S_t be the set of agents in pairs so that $N = D_t \cup S_t$. Let $\mathcal{P}_t|_{S_t}$ denote the set of pairs of partition $\mathcal{P}_t$. Notice that $\mathcal{P}_t|_{S_t}$ is a stable pairing of the agents in S_t.

The algorithm

Stage 1: Finding a maximum irreversible set of pairs

Step 1. Let $(N_1, (\succeq_{a_i}^{R_1})_{i\in N_1})$ be a problem where $N_1 = N$ and $(\succeq_{a_i}^{R_1}) = (\succeq_{a_i})$. Compute a stable partition $\mathcal{P}_1$ for $(N_1, (\succeq_{a_i}^{R_1})_{a_i\in N_1})$.[13]
Let $N_1 = D_1 \cup S_1$. If $S_1 = \emptyset$ then STOP and set $\mu_I = \mathcal{P}_1|_{S_1} = \emptyset$. If $S_1 = N_1$ then STOP and set $\mu_I = \mathcal{P}_1|_{S_1}$. Otherwise, for every agent $a_i \in S_1$ remove from $(\succeq_{a_i}^{R_1})$ every agent $a_k \in D_1$ and every agent $a_j \in S_1$, $a_j \neq a_i$, such that $a_k \succ_{a_i} a_j$ (a_i prefers a_k to a_j) for some $a_k \in D_1$.

Step t. Define a reduced problem $(N_t, (\succeq_{a_i}^{R_t})_{a_i\in N_t})$ where $N_t = S_{t-1}$ and $(\succeq_{a_i}^{R_t})$ is agent a_i's preference list after the clearing process over $(\succeq_{a_i}^{R_{t-1}})$. If no agent is removed from agent a_i's preference list, set $(\succeq_{a_i}^{R_t}) = (\succeq_{a_i}^{R_{t-1}})$. Compute a stable partition $\mathcal{P}_t$ for $(N_t, (\succeq_{a_i}^{R_t})_{a_i\in N_t})$.
Let $N_t = D_t \cup S_t$. If $S_t = \emptyset$ then STOP and set $\mu_I = \mathcal{P}_t|_{S_t} = \emptyset$. If $S_t = N_t$ then STOP and set $\mu_I = \mathcal{P}_t|_{S_t}$. Otherwise, for every agent $a_i \in S_t$ remove from $(\succeq_{a_i}^{R_t})$ every agent $a_k \in D_t$ and every agent $a_j \in S_t$ $a_j \neq a_i$, such that $a_k \succ_{a_i} a_j$ (a_i prefers a_k to a_j) for some $a_k \in D_t$. Increase t and repeat this step.

Stage 2: Build a stable partition for $(N, \succeq)$.
Let I denote the set of matched agents in μ_I and let D denote the remaining agents. Join $\mathcal{P}_1\ |_D$ with μ_I to determine a stable partition $\mathcal{P}^*$ on N. That is, $\mathcal{P}^* = \mathcal{P}_1\ |_D \cup \mu_I$.

[13] An algorithm which computes a stable partition in linear time can be found in Tan [31].

Stage 3: Build a matching from stable partition $\mathcal{P}^*$
From stable partition $\mathcal{P}^*$ derive a $\mathcal{P}^*$-stable matching. This matching is called Q^*-stable.

If the problem is solvable then the output of the algorithm is a stable matching and therefore it is immediate that it is maximum internally stable and maximum irreversible. That is,

Remark 5 *A stable matching is a $\mathcal{Q}$-stable.*

The previous remark enable us to focus on unsolvable problems. We now present some claims to prove that the algorithm provides a $\mathcal{Q}$-stable matching.

First, we present some claims, which are needed to show that the algorithm provides a matching with a set of irreversible pairs of maximum size.

Claim 1 *Assume that μ_I is an irreversible set of pairs formed by a subset of agents $I \subseteq N$ and $\mathcal{P}\mid_D$ is a stable partition for the problem restricted to $D = N \backslash I$. Then $\mathcal{P} = \mathcal{P}\mid_D \cup \mu_I$ is a stable partition on N.*

Proof. No pair $\{a_i, a_j\} \subseteq I$ can block[14] $\mathcal{P}$ by the stability of μ_I, no pair $\{a_i, a_j\} \subseteq D$ can block by the stability of $\mathcal{P}\mid_D$, and no pair $\{a_i, a_j\}$ with $a_i \in I$ and $a_j \in D$ can block by the irreversibility of μ_I. ■

Claim 2 *If agent a_i either belongs to an odd ring or is a singleton in a stable partition $\mathcal{P}$ then he/she can never be part of an irreversible set of pairs.*

Proof. By contradiction assume that a_i is part of an irreversible matching μ_I. Then, by Claim 1, μ_I could be extended to a stable partition $\mathcal{P}' = \mathcal{P}_{N\backslash I} \cup \mu_I$. But then the set of odd rings and singletons would not be the same in $\mathcal{P}$ and $\mathcal{P}'$, contradicting Remark 1 (ii). ■

Claim 3 *The set of pairs μ_I derived in Stage 1 of the algorithm, is maximum irreversible.*

Proof. μ_I is irreversible by construction since there is no agent $a_i \in I$ such that a_i prefers an agent a_k outside I to her current partner. It remains to prove the maximality of μ_I. That is, if there exists another irreversible set of pairs

[14] Abusing notation we say that a pair of agents $\{a_i, b_j\}$ block partition $\mathcal{P}$ if $b_j \succ_{a_i} a_{i-1}$ then $b_{j-1} \succ_{b_j} a_i$.

$\mu_{I'}$ with $I' \subseteq N$, then $I' \subseteq I$. Suppose for a contradiction that there exists $\mu_{I'}$ such that $I' \setminus I \neq \emptyset$. This means that there must be a step t in the algorithm where $I' \subseteq N_t$, but $S_t \setminus I' \neq \emptyset$. However, if $\mu_{I'}$ is irreversible for N then it is also irreversible for N_t, therefore no agent in I' can be a singleton or part of an odd ring in a stable partition $\mathcal{P}_t$ for N_t by Claim 2, a contradiction. ■

The above claim and the argument in its proof imply the following corollary.[15]

Corollary 4 *For a roommate problem the set of agents matched in the largest irreversible set of pairs is the same.*

If matching $\mathcal{Q}^*$-stable matching is maximum irreversible by Claim 3 and maximum internally stable by Remark 3 then the following theorem can be established.

Theorem 5 *There always exists a $\mathcal{Q}^*$-stable matching for any roommate problem.*

Regarding the complexity of the algorithm the following result is established.

Proposition 6 *A $\mathcal{Q}^*$-stable matching can be computed in $O(mn)$ time, where n is the number of agents and m is the total length of the preference lists.*

Proof. Stage 1 can be invoked at most n times since the set of agents in pairs in the initial partition $\mathcal{P}_1$ can only shrink, and it is stopped when it does not shrink. The execution of each step takes linear time in m, which is the total length of the preference lists, since a stable partition can be found with Tan's algorithm [31] in $O(m)$ time, and the clearing process in Stage 1 can also be conducted in linear time. Therefore the algorithm terminates in $O(mn)$ time.[16]

■

The algorithm and the results above can be illustrated with an example:

Example 5 *Consider the following 10-agents problem:*

[15] This corollary is closely related to Proposition 3 in Inarra *et al.* [16] although they have proven in different manner.

[16] Tan and Hsueh [33] propose another algorithm, which constructs a stable partition incrementally and whose complexity is $O(n^3)$ where n is the number of agents. This algorithm can be seen as a generalization of the procedure of convergence used by Roth and Vande Vate [26] to a stable marriage.

a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}	a_{11}	a_{12}
a_2	a_3	a_1	a_3	a_4	a_7	a_5	a_9	a_{10}	a_{11}	a_8	a_1
a_3	a_1	a_2	a_5	a_6	a_5	a_6	a_4	a_8	a_9	a_{10}	a_2
a_4	a_6	a_4	a_6	a_7	a_1	a_1	a_5	a_1	a_4	a_1	a_3
a_5	a_4	a_5	a_2	a_1	a_9	a_2	a_{11}	a_6	a_5	a_2	a_4
a_6	a_5	a_6	a_1	a_2	a_2	a_3	a_1	a_3	a_1	a_3	a_5
a_7	a_7	a_7	a_8	a_3	a_3	a_4	a_2	a_4	a_2	a_4	a_6
a_8	a_8	a_8	a_7	a_8	a_8	a_8	a_3	a_5	a_3	a_5	a_7
a_9	a_9	a_9	a_9	a_9	a_4	a_9	a_6	a_7	a_6	a_6	a_8
a_{10}	a_{10}	a_{10}	a_{10}	a_{10}	a_{10}	a_{10}	a_7	a_2	a_7	a_7	a_9
a_{11}	a_{11}	a_{11}	a_{11}	a_{11}	a_{11}	a_{11}	a_{10}	a_{11}	a_8	a_9	a_{10}
a_{12}	a_{12}	a_{12}	a_{12}	a_{12}	a_{12}	a_{12}	a_{12}	a_{12}	a_{12}	a_{12}	a_{11}
a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}	a_{11}	a_{12}

Stage 1: *Finding a maximum irreversible set of pairs*

Step 1. *Computing a stable partition for the problem* $(N_1, (\succeq^{R_1})_{a_i \in N_1})$, $\mathcal{P}_1 = \{\{a_1, a_2, a_3\}, \{a_4, a_5\}, \{a_6, a_7\}, \{a_8, a_{11}\}, \{a_9, a_{10}\}, \{a_{12}\}\}$ *is obtained, where* $D_1 = \{a_1, a_2, a_3, a_{12}\}$ *and* $S_1 = \{a_4, a_5, a_6, a_7, a_8, a_9, a_{10}, a_{11}\}$. *Remove from the list of preferences of each agent in* S_1 *all agents in* D_1 *and those that are less preferred, except for herself.*

Step 2. *A reduced problem* $(N_2, (\succeq^{R_2})_{i \in N_2})$, *is defined where* $S_1 = N_2$:

a_4	a_5	a_6	a_7	a_8	a_9	a_{10}	a_{11}
a_4	a_4	a_7	a_5	a_9	a_{10}	a_{11}	a_8
	a_6	a_5	a_6	a_4	a_8	a_9	a_{10}
	a_7	a_6	a_7	a_5	a_9	a_4	a_{11}
	a_5			a_{11}		a_5	
				a_8		a_{10}	

Computing a stable partition for this reduced problem, $\mathcal{P}_2 = \{\{a_4\}, \{a_5, a_6, a_7\}, \{a_8, a_{11}\}, \{a_9, a_{10}\}\}$ *is obtained, where* $D_2 = \{a_4, a_5, a_6, a_7\}$ *and* $S_2 = \{a_8, a_9, a_{10}, a_{11}\}$. *Remove from the list of preferences of each agent in* S_2 *all agents in* D_2 *and those that are less preferred, except for herself.*

Step 3. *A reduced problem* $(N_3, (\succeq^{R_3})_{i \in N_3})$ *is defined where* $S_2 = N_3$:

a_8	a_9	a_{10}	a_{11}
a_9	a_{10}	a_{11}	a_8
a_8	a_8	a_9	a_{10}
	a_9	a_{10}	a_{11}

Computing a stable partition for the previous reduced problem $\mathcal{P}_3 = \{\{a_8, a_9\}, \{a_{10}, a_{11}\}\}$ *is obtained, where* $D_3 = \emptyset$ *and* $S_3 = \{a_8, a_9, a_{10}, a_{11}\}$. *Since* $S_3 = N_3$ *STOP. Set* $\mu_I = \{\{a_8, a_9\}, \{a_{10}, a_{11}\}\}$.

Stage 2: Let $I = \{a_8, a_9, a_{10}, a_{11}\}$ *and* $D = \{a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_{12}\}$. *The following stable partition* $\mathcal{P}^* = \mathcal{P}_1 \mid_D \cup \mu_I = \{\{a_1, a_2, a_3\}, \{a_4, a_5\}, \{a_6, a_7\}, \{a_8, a_9\}, \{a_{10}, a_{11}\}, \{a_{12}\}\}$ *is determined.*

Stage 3: From $\mathcal{P}^*$*-stable partition let agent* a_1 *to be left out from the ring while agents* a_2 *and* a_3 *are matched preserving the ordering of the and the remaining agents are matched as in* $\mathcal{P}^*$. *The resulting matching* $\{\{a_1\}, \{a_2, a_3\}, \{a_4, a_5\}, \{a_6, a_7\}, \{a_8, a_9\}, \{a_{10}, a_{11}\}, \{a_{12}\}\}$ *is a* $\mathcal{Q}^*$*-stable matching.*

Note that we have started with partition $\mathcal{P}_1 = \{\{a_1, a_2, a_3\}, \{a_4, a_5\}, \{a_6, a_7\}, \{a_8, a_{11}\}, \{a_9, a_{10}\}, \{a_{12}\}\}$ *and have found that the* $\mathcal{P}_1$*-stable matching derived from it is not maximum irreversible. With the algorithm partition* $\mathcal{P}^* = \{\{a_1, a_2, a_3\}, \{a_4, a_5\}, \{a_6, a_7\}, \{a_8, a_9\}, \{a_{10}, a_{11}\}, \{a_{12}\}\}$ *is reached and a* $\mathcal{Q}^*$*-stable matching* $\{\{a_1\}, \{a_2, a_3\}, \{a_4, a_5\}, \{a_6, a_7\}, \{a_8, a_9\}, \{a_{10}, a_{11}\}, \{a_{12}\}\}$ *is obtained given that it is maximum internal stable and it contains the maximum irreversible set of pairs* $\{\{a_8, a_9\}, \{a_{10}, a_{11}\}\}$.

So far we have not addressed the importance of maximizing the number of pairs matched in a matching. However, in many applications an essential objective is to match as many agents as possible. Consider for example the problem of dividing agents in a fixed number of two-person rooms, or the kidney exchange problem.[17] In those cases we can join the single agents from the Q^*-stable matching outcome of the algorithm. In the example above matching $\{\{a_1, a_{12}\}, \{a_2, a_3\}, \{a_4, a_5\}, \{a_6, a_7\}, \{a_8, a_9\}, \{a_{10}, a_{11}\}\}$ is obtained.

Remark 6 *Every matching formed by joining mutually acceptable unmatched agents from a* $\mathcal{Q}^*$*-stable matching is also a* $\mathcal{Q}^*$*-stable matching.*

We give an algorithm for computing a matching which lies at the intersection of the set of maximum internal stable matchings and the set of maximum irreversible stable matchings. It can be checked that neither of these two

[17] For more details see the survey on market design for kidney exchange by Sönmez and Ünver [28].

solutions contain the other. In Example 6, $\mathcal{P}_1$-stable matching is maximum internally stable and not maximum irreversible while, for instance, matching $\{\{a_1\},\{a_2\},\{a_3\},\{a_4,a_5\},\{a_6,a_7\},\{a_8,a_9\},\{a_{10},a_{11}\},\{a_{12}\}\}$ is maximum irreversible and not maximum internally stable.

4.1 $\mathcal{Q}$-stable matchings and the absorbing sets

For the roommate problem Inarra *et al.* [16] seek to determine which matchings a decentralized process may lead to. They consider a dynamic process in which a matching is adjusted when a blocking pair of agents mutually decide to become partners. Either this change gives a stable matching or a new blocking pair of agents will generate another matching and so on. If there are stable matchings the process eventually converges to one of them. Otherwise it leads to a set of matchings (an absorbing set) such that any matching in the set can be obtained from any other and it is impossible to escape from the matchings in that set. As mentioned in the Introduction, it is important to investigate whether our proposal, the Q^*-stable matching, is achievable from a free interactions of agents, i.e. whether it belongs to an absorbing set. That is the task that we undertake in this subsection.

A non-empty set of matchings $\mathcal{A}$ is an *absorbing set* if the following conditions hold: *(i) For all* $\mu,\mu' \in \mathcal{A}$ $(\mu \neq \mu')$, $\mu' R^T \mu$. *(ii) For all* $\mu \in \mathcal{A}$ there is no $\mu' \notin \mathcal{A}$ such that $\mu' R M$.

Condition (i) says that every matching in an absorbing set is (directly or indirectly) dominated by another matching in the same set. Condition (ii) says that no matching in an absorbing set is directly dominated by a matching outside that set.

The following remark states some properties of absorbing sets and their matchings.

Remark 7 *(i) Absorbing sets satisfy the property of* outer stability*, which requires that every matching not belonging to an absorbing set be (indirectly) dominated by the matchings of an absorbing set (Kalai et al. [18]). (ii) Every absorbing set contains a $\mathcal{P}$-stable matchings but not all $\mathcal{P}$-stable matchings belong to an absorbing set (Inarra et al. [16]).*

Next, consider the relationship between the $\mathcal{Q}$-stable matchings and absorbing sets. We find that not all $\mathcal{Q}$-stable matchings belong to an absorbing set.

In Example 1 all matchings with at least one pair of agents are Q-stable, however, matching $\mu_1 = \{\{2\}, \{3\}, \{1, 4\}\}$ does not belong to the unique absorbing set. To see this let $\mu_2 = \{\{2, 3\}, \{1, 4\}\}$. It is immediate that $\mu_2 R^T \mu_1$ but not $\mu_1 R^T \mu_2$. Hence condition 1 of the definition of absorbing set is not satisfied. However, the matching provided by the algorithm belongs to an absorbing set.

Let $\mathcal{P}^*$ be a maximum irreversible stable partition (one of such a partition can be obtained at Stage 2 of the algorithm) and let I be the set of agents irreversibly matched in $\mathcal{P}^*$. Denote by S^* the set of pairs of $\mathcal{P}^*$ and $D^* = N \backslash S^*$.

Claim 4 *Let μ be a $\mathcal{P}^*$-stable matching and let $V = S^* \setminus I$. Then $\mu' R^T \mu$ where μ' is a matching such that $\mu'|_V(a_i) = a_i$ for all $a_i \in V$ and $\mu'|_{N \setminus V}(a_j) = \mu|_{N \setminus V}(a_j)$ for all $a_j \in N \setminus V$.*

Proof. We show that there is a sequence of matchings from μ to μ', in which all pairs in $\mu|_V$ become singles while the remaining agents are paired as in μ. Notice $\mu|_V$ is blocked by an agent in $_V$ and either a single agent from an odd ring or another agent in V who has previously become single. To do it consider the following iterative process:

For $t = 1$. Let $\upsilon_1 = \{\{a_i, \mu(a_i)\} \subseteq V : \mu$ is blocked by $\{a_j, b_j\}$ where $a_j \in \{a_i, \mu(a_i)\}$ and $b_j \in D^*\}$ and let $V_1 = \cup \upsilon_1$. Let $\mathcal{M}|_{\mathcal{P}^*}$ be the set of $\mathcal{P}^*$-stable matchings of $\mathcal{P}^*$ so that $\mu \in \mathcal{M}|_{\mathcal{P}^*}$. Thus, V_1 is the set of agents who block some matching in $\mathcal{M}|_{\mathcal{P}^*}$ with a single agent of an odd ring of $\mathcal{P}^*$ and their partners under μ.[18] Note that $V_1 \neq \emptyset$ otherwise $S^* = I$ and we are done. Set $\mu = \mu^1$ and consider $\{a_i, a_{i+1}\} \in \mu^1|_{V_1}$ such that $b_i \succ_{a_i} a_{i+1}$ and $a_i \succ_{b_i} b_i$ with $b_i \in D^*$. W.l.o.g. assume that $\mu^1(b_i) = b_i$, otherwise by Remark 4 there exists another matching $\hat{\mu} \in \mathcal{M}|_{\mathcal{P}^*}$ such that $\hat{\mu} R^T \mu^1$ and $\hat{\mu}(b_i) = b_i$. Matching μ^1 is blocked by $\{a_i, b_i\}$ forming matching μ_1^1 in which $\mu_1^1(a_{i+1}) = a_{i+1}$. By the stability of partition $\mathcal{P}^*$, $b_{i-1} \succ_{b_i} a_i$ and $b_i \succ_{b_{i-1}} \mu^1(b_{i-1})$. Thus, matching μ_1^1 is blocked by $\{b_{i-1}, b_i\}$ forming matching μ_2^1 in which $\mu_2^1(a_i) = a_i$ and $\mu_2^1(a_{i+1}) = a_{i+1}$. Repeat this step for all pairs in $\mu|_{V_1}$ until agents in V_1 become singles.[19] Hence, we achieve a matching μ_k^1 such that $\mu_k^1(a_i) = a_i$ for all $a_i \in V_1$ and $\mu_k^1(a_j) = \mu^1(a_j)$ for all $a_j \in N \setminus V_1$ and $\mu_k^1 R^T \mu^1$. Then go to next step.

[18] By definition of stable partition no agent in V prefers a singleton of the partition to her partner in the partition and therefore this type of pairs cannot block any $\mathcal{P}$-stable matching.

[19] Remark 4 can be extended to any set of matchings such that agents in odd rings are paired as in the set of $\mathcal{P}$-stable matchings and the remaining agents are equally paired.

For $t > 1$**.** Let μ_l^{t-1} be the matching obtained at the end of Step $t-1$ such that $\mu_l^{t-1} R^T \mu$. Set $\mu_l^{t-1} = \mu^t$. Let $\upsilon_t = \{\{a_i, \mu(a_i)\} \subseteq V : \mu$ is blocked by $\{a_j, a_k\}$ where $a_j \in \{a_i, \mu(a_i)\}$ and $a_k \in V_{t-1}\}$ and let $V_t = \cup \upsilon_t$. If $V_t = \emptyset$ then $\mu^t = \mu'$ and we are done. Otherwise there is a pair $\{a_l, a_{l+1}\} \in \mu|_{V_t}$ such that $a_t \succ_{a_l} a_{l+1}$ and $a_l \succ_{a_t} a_t$ for some $a_t \in V_{t-1}$. Then matching μ^t is blocked by $\{a_l, a_t\}$ forming a new matching μ_1^t in which $\mu_1^t(a_{l+1}) = a_{l+1}$. By the stability of partition $\mathcal{P}^*$, $\mu(a_t) = a_{t-1} \succ_{a_t} a_l$ and $a_t \succ_{a_{t-1}} a_{t-1}$ so that μ_1^t is blocked by pair $\{a_t, a_{t-1}\}$ forming matching μ_2^t in which $\mu_2^t(a_l) = a_l$, $\mu_2^t(a_{l+1}) = a_{l+1}$ and $\mu_2^t(a_t) = a_{t-1}$. Thus pair $\{a_t, a_{t-1}\}$ is rejoined and it must be unmatched. Repeat the followed reasoning backwards from Step $t-1$ to Step 1 until we reach a matching in which all agents in $V^{t-1} \cup \{a_l, a_{l+1}\}$, where $V^{t-1} = \cup_{i=1}^{t-1} V_i$, are alone and the remaining agents are paired as in μ. Then repeat this step for all pairs in $\mu|_{V_t}$ until all agents in V^t become singles. Hence, we achieve a matching μ_m^t such that $\mu_m^t(a_i) = a_i$ for all $a_i \in V^t$ and $\mu_m^t(a_j) = \mu(a_j)$ for all $a_j \in N \backslash V^t$ and $\mu_m^t R^T \mu$. Then increase t and repeat this step.

Since the number of agents in V is finite, the process finishes in finite time.

■

Proposition 7 *Every P-stable matching which is maximum irreversible belongs to an absorbing set.*

Proof. Let μ^* be a P-stable matching which is maximum irreversible. Assume that μ^* does not belong to an absorbing set. By Remark 7 and from the definition of absorbing set there exists a $\mathcal{P}$-stable matching μ in an absorbing set $\mathcal{A}$ such that $\mu R^T \mu^*$ but not $\mu^* R^T \mu$.

Since $\mu R^T \mu^*$ then $\mu_I^* = \mu_I$ since μ_I^* is a maximum irreversible set of pairs. By Remark 1 $\mu^*|_{N \backslash U} = \mu|_{N \backslash U}$, hence $\mu^*|_U \neq \mu|_U$ otherwise μ and μ^* coincide.

By Claim 4 $\mu' R^T \mu$ such that $\mu'(a_i) = a_i$ for all $a_i \in U$ and $\mu'(a_j) = \mu(a_j)$ for all $a_j \in N \setminus U$. But since $\mu^*(a_i) \succ a_i$ for all $a_i \in U$, then it is easy to see that $\mu^*(a_i) R^T \mu' R^T \mu$, contradicting the initial assumption. ■

Since the matching obtained as the output of the algorithm is a P-stable matching which is maximum irreversible, we can state the following:

Corollary 8 *The $\mathcal{Q}^*$-stable matching obtained as the output of the algorithm belongs to an absorbing set.*

To finish this section the following remark can be established.

Remark 8 *Every matching formed by joining mutually acceptable unmatched agents in a $\mathcal{Q}^*$-stable matching belongs to an absorbing set.*

5 Concluding remarks

In this paper we have claimed that Q-stable matchings are good proposals in the class of roommate problems with strict preferences and we have presented an efficient algorithm for finding one of these matchings. To conclude we discuss our approach and point out some open questions for future analysis.

Even though the solution of absorbing sets is core-consistent for roommate problems with strict preferences, we have not considered it in our pool of core-consistent solutions (Section 3). The main reason is that we propose Q-stable matchings as a "static" solution for problems where a policy maker has to provide a single matching in a centralized market. However, absorbing sets is a "dynamic" solution that rules out those matchings which are not achievable from the free interaction of agents. That is, they are the resulting matchings from a decentralized market. In any case, it is interesting to understand the relation between absorbing sets and the solutions that compose Q-stable matchings. On the one hand, the intersection between absorbing sets and maximum internally stable matchings is non-empty (this is easy to see given that an absorbing set always contains a P-stable matching). On the other hand, it is easy to check that absorbing sets are included in the set of maximum irreversible matchings.[20] Given the inclusive relation, one may wonder why we have not considered instead the intersection between absorbing sets and the set of maximum internally stable matchings. The reason is that it is not clear whether the outcomes of the algorithm are the only reasonable proposals. In Example 2 matching $\mu_1 = \{\{1\}, \{2,3\}, \{10\}, \{4,8\}, \{5,9\}, \{6,7\}\}$ belongs to an absorbing set while matching $\mu_2 = \{\{1\}, \{3\}, \{2,10\}, \{4,8\}, \{5,9\}, \{6,7\}\}$ does not and, however, both are maximum internal stable and maximum irreversible. Thus,

[20] Here we provide a sketch of the proof. First, following a similar reasoning as in Claim 4 it can be deduced that from any P-stable matching which is not maximum irreversible there is a sequence of blocking pairs to a P-stable matching which is maximum irreversible. Next, assume that there is a matching μ in an absorbing set which is not maximum irreversible. Then $\mu^* R^T \mu' R^T \mu$ where μ^* is a P-stable matching which is maximum irreversible and μ' is a a P-stable matching of the same absorbing set as μ (possibly $\mu^* = \mu'$). However, not $\mu R^T \mu^*$, contradicting that matching μ belongs to an absorbing set.

in qualitative terms there are not sufficient reasons for a policy maker to consider one matching better than the other.

A natural question to be addressed is whether an extension of the Q-stable matchings could be a core-consistent solution for other models where the core may be empty. Consider for instance hedonic games, Bogomolnaia and Jackson [6], which constitute an immediate generalization of the roommate problem. In these games each agent has a strict preference over coalitions of agents that contain him. Following analogous patterns of the definitions of maximum irreversibility and maximum internal stability we find that the set of the Q-stable coalitional structures for an hedonic game may be empty as the following example shows:

Example 6 *Consider the following* 11*-agents hedonic formation problem:*

1	2	3	4	5	6	7	8	9	10	11
$\{1,2,3\}$	$\{1,2,3\}$	$\{3,4\}$	$\{4,5\}$	$\{1,5\}$	$\{6,7\}$	$\{7,8\}$	$\{8,9\}$	$\{7,9\}$	$\{10,11\}$	$\{6,11\}$
$\{1,5\}$	$\{2,6\}$	$\{1,2,3\}$	$\{3,4\}$	$\{4,5\}$	$\{2,6\}$	$\{7,9\}$	$\{7,8\}$	$\{8,9\}$	$\{10,6\}$	$\{10,11\}$
					$\{6,10\}$	$\{6,7\}$				
					$\{6,11\}$					

Here, coalition structures containing the set of coalition $\{\{1,2,3\},\{4,5\}\}$ are the only maximum irreversible, whilst $\{\{1,5\},\{3,4\},\{2,6\},\{8,9\},\{10,11\}\}$, $\{\{1,5\},\{3,4\},\{2,6\},\{7,8\},\{10,11\}\}$ and $\{\{1,5\},\{3,4\},\{2,6\},\{7,9\},\{10,11\}\}$ are the only maximum internally stable sets. Therefore their intersection is empty. This negative result indicates that the roommate problem has a very particular structure that makes difficult the extensions of the results in this paper to more complex models. Hence more work is needed to find a core consistent solution for such a more general model.

References

[1] Abraham, D. J., and Manlove, D. F. (2004). Pareto optimality in the roommates problem. Technical Report No. TR-2004-182 of the Computing Science Department of Glasgow University.

[2] Abraham, D. J., Biró, P., and Manlove, D. F. (2006). "Almost stable" matchings in the Roommates problem. In Approximation and Online Algorithms, 3879: 1-14.

[3] Biermann F. M. (2011). A measure to compare matchings in marriage markets. The Hebrew University of Jerusalem. Discussion Paper # 575.

[4] Biró, P., Cechlárová, K., and Fleiner, T. (2008). The dynamics of stable matchings and half-matchings for the stable marriage and roommates problems. International Journal of Game Theory, 36: 333-352.

[5] Biró, P., Irving, R., and Manlove, D.F. (2010). Popular matchings in the marriage and roommates problems. Proceedings of the seventh International Conference on Algorithms and Complexity (CIAC), 97-108.

[6] Bogomolnaia, A., and Jackson, M. O. (2002). The stability of hedonic coalition structures. Games and Economic Behavior, 38: 201-230.

[7] Can, B. and Klaus, B. (2013) Consistency and population sensitivity properties in marriage and roommate markets. Social Choice and Welfare, 41: 835-862.

[8] Cechlárová, K., and Ferková, S. (2004). The stable crews problem. Discrete Applied Mathematics, 140: 1-17.

[9] Chung, K. S. (2000). On the existence of stable roommate matchings. Games and Economic Behavior, 33: 206-230.

[10] Diamantoudi, E., Miyagawa, E., and Xue, L. (2004). Random paths to stability in the roommate problem. Games and Economic Behavior, 48: 18-28.

[11] Echenique F., and J. Pereyra (2013). Strategic uncertainty and unraveling in matching markets. Caltech SS Working Paper 1373.

[12] Gale, D., and Shapley, L. S. (1962). College admissions and the stability of marriage. The American Mathematical Monthly, 69: 9-15.

[13] Gudmundsson, J. (2014). When do stable roommate matchings exist? A review. Review of Economic Design, 18: 151-161.

[14] Gusfield, D., and Irving, R. W. (1989). The stable marriage problem: structure and algorithms (Vol. 54). Cambridge: MIT press.

[15] Inarra, E., Larrea, C., and Molis, E. (2008). Random paths to P-stability in the roommate problem. International Journal of Game Theory, 36: 461-471.

[16] Inarra, E., Larrea, C., and Molis, E. (2013). Absorbing sets in roommate problems. Games and Economic Behavior, 81: 165–178.

[17] Jackson, M. O., and Watts, A. (2002). On the formation of interaction networks in social coordination games. Games and Economic Behavior, 41: 265-291.

[18] Kalai, E., Pazner, E.A., and Schmeidler, D. (1976). Admissible outcomes of social bargaining processes as collective choice correspondence. Econometrica, 63: 299-325.

[19] Klaus, B., and Klijn, F. (2010). Smith and Rawls share a room: stability and medians. Social Choice and Welfare, 35: 647-667.

[20] Klaus, B., Klijn, F., and Walzl, M. (2010). Stochastic stability for roommate markets. Journal of Economic Theory, 145: 2218-2240.

[21] Kujansuu, E., Lindberg, T. and Makinen, E. (1999). The stable roommates problem and chess tournament pairings. Divulgaciones Matemáticas, 7: 19-28.

[22] Manlove, D.F. (2013). Algorithms of Matching Under Preferences. World Scientific Publishing.

[23] Morrill, T. (2010). The roommates problem revisited. Journal of Economic Theory, 145: 1739-1756.

[24] Özkal-Sanver, İ. (2010). Impossibilities for roommate problems. Mathematical Social Sciences, 59: 360-363.

[25] Pittel, B. G., and Irving, R. W. (1994). An upper bound for the solvability probability of a random stable roommates instance. Random Structures & Algorithms, 5: 465-486.

[26] Roth, A. E., and Vande Vate, J. H. V. (1990). Random paths to stability in two-sided matching. Econometrica, 58: 1475-1480.

[27] Roth, A. E., Sönmezc, T. and Ünverd, M. U. (2005) Pairwise kidney exchange Journal of Economic Theory, 125: 151 – 188

[28] Sönmez, T., and Ünver, M. U. (2010). Market design for kidney exchange. Oxford Handbook of Market Design, Oxford University Press.

[29] Sotomayor, M. (2011). The Pareto-stability Concept is a Natural Solution Concept for Discrete Matching Markets with Indifferences. International Journal of Game Theory, 40: 631-644.

[30] Tan, J. J. (1990). A maximum stable matching for the roommates problem. BIT Numerical Mathematics, 30: 631-640.

[31] Tan, J. J. (1991a). A necessary and sufficient condition for the existence of a complete stable matching. Journal of Algorithms, 12: 154-178.

[32] Tan, J. J. (1991b). Stable matchings and stable partitions. International Journal of Computer Mathematics, 39: 11-20.

[33] Tan, J. J., and Hsueh, H-C. (1995). A generalization of the stable matching problem. Discrete Applied Mathematics, 59:, 87-102.

[34] Toda, M. (2006). Monotonicity and consistency in matching markets. International Journal of Game Theory, 34: 13-31.

6 Appendix

Let $(N, \succeq)$ be a problem where $N = \{a_1, ..., a_n\}$ is the set of agents. Let $(N_k, (\succeq)^k)$ be the restricted problem where $N_k = \{a_1, ..., a_k\}$ and $(\succeq)^k$ is the preference list of agents in N_k in which $N \backslash N_k$ agents have been deleted. Assume that we have already found a stable partition, say $\mathcal{P}_k$, a for $(N_k, (\succeq)^k)$, $1 \leq k \leq n-1$ and that one additional agent a_{k+1} is added. The question is whether a stable partition $\mathcal{P}_{k+1}$ for the enlarged problem $(N_{k+1}, (\succeq)^{k+1})$ can be determined. The answer in in the affirmative. The following acceptance-rejection procedure determines it:

Let problem $(N_k, (\succeq)^k)$ and let a_{k+1} be the newcomer. By embedding agent a_{k+1} into the existent lists and adding her own list to problem $(N_k, (\succeq)^k)$, problem $(N_{k+1}, (\succeq)^{k+1})$ is constructed. (Note that $\{a_1\}$ is the unique stable partition for $(N_1, (\succeq)^1)$). Given a stable partition $\mathcal{P}_k$ for $(N_k, (\succeq)^k)$ let agent a_{k+1} propose to the set of agents in N_k according to her preference order:

1. If nobody accepts her proposal, then a_{k+1} is alone and stable partition $\mathcal{P}_{k+1} = \mathcal{P}_k \cup \{a_{k+1}\}$ is obtained.

2. If a_{k+1} is accepted by agent x there are three possible cases:

(i) If x is unmatched in $\mathcal{P}_k$ then $\mathcal{P}_{k+1} = (\mathcal{P}_k \backslash \{x\}) \cup \{x, a_{k+1}\}$ and stable partition $\mathcal{P}\backslash_{k+1}$ for $(N_{k+1}, (\succeq)^{k+1}$ is obtained.

(ii) If x is currently in an odd ring, say $(a_1, a_2, ..., a_{2m}, x)$, then the arrival of a_{k+1} decomposes the set into pairs and $\mathcal{P}_{k+1} = \{\{a_1, a_2\}, \{a_3, a_4\}....\{a_{2m-1}, a_{2m}\}\} \cup \{\{a_{k+1}, x\}\}$ becomes a stable partition for $(N_{k+1}, (\succeq)^{k+1})$.

(iii) If x is in a mutually acceptable pair say $\{x, y\}$ in $\mathcal{P}_k$ then y becomes single and $\mathcal{P}_{k+1} = (\mathcal{P}_k \backslash \{x, y\}) \cup \{x, a_{k+1}\}$ is a stable partition for $N_{k+1} \setminus \{y\}$. Now y reenters the market as a new proposer. In this phase a proposal-rejection sequence takes place which may stops in 1 in 2 (i) or 2 (ii). In both cases the desired stable partition comes out. Otherwise an agent who made a proposal once receives a proposal later and repetition takes place. Then stop. All agents involved in the subsequent cycle form an odd ring and stable partition P_{k+1} for problem $(N_{k+1}, (\succeq)^{k+1})$ is constructed.

Example 7 *Consider the following 7-agent problem:*

a_1	a_2	a_3	a_4	a_5	a_6	a_7
a_3	a_1	a_4	a_2	a_6	a_7	a_5
a_2	a_4	a_2	a_5	a_4	a_5	a_6
a_4	a_3	a_1	a_3	a_7	a_2	a_1
a_5	a_6	a_7	a_7	a_2	a_3	a_2
a_6	a_7	a_5	a_6	a_3	a_1	a_3
a_7	a_5	a_6	a_1	a_1	a_4	a_4
a_1	a_2	a_3	a_4	a_5	a_6	a_7

W.l.o.g assume that agents arrive to the process in the following arbitrary order: $a_1, a_2, a_3, ..., a_7$.

- a_1 *arrives and forms a stable partition* $\{a_1\}$ *for* $(N_1, (\succeq)^1)$.
- a_2 *arrives and proposes to* a_1. *Since* a_1 *rather matches* a_2 *than being alone, accepts the proposal and* $\{a_1, a_2\}$ *is stable partition for* $(N_2, (\succeq)^2)$ *(see Case 2 (i)).*
- a_3 *arrives and proposes to* a_2 *who rejects then proposes to* a_1 *who accepts. Pair* $\{a_1, a_3\}$ *forms and* a_2, *now alone, proposes to* a_1 *who rejects and then to* a_3 *who accepts. Pair* $\{a_2, a_3\}$ *and* a_1 *is abandoned. This agent proposes* a_3 *who rejects and then to* a_2 *who accepts and pair* $\{a_1, a_2\}$ *is formed and* a_3 *is a proposer again. The cycling agents get together in the set* $\{a_1, a_2, a_3\}$ *and stable partition* $\{\{a_1, a_2, a_3\}\}$ *for* $(N_3, (\succeq)^3)$ *is obtained (see Case 2 (iii)).*
- a_4 *arrives and proposes to* a_2 *who accepts. Stable partition* $\{\{a_1, a_3\}, \{a_2, a_4\}\}$ *for* $(N_4, (\succeq)^4)$ *is obtained (see Case 2 (ii)).*
- a_5 *arrives and is rejected for all agents in the process. Hence stable partition*$\{\{a_1, a_3\}, \{a_2, a_4\}, \{a_5\}\}$ *for* $(N_5, (\succeq)^5)$ *is obtained. (See Case 1)*
- a_6 *arrives and proposes to* a_5 *who accepts. Stable partition* $\{\{a_1, a_3\}, \{a_2, a_4\}, \{a_5, a_6\}\}$ *for* $(N_6, (\succeq)^6)$ *is obtained (see Case 2 (i)).*
- a_7 *arrives and proposes to* a_5 *who rejects, then to* a_6 *who acepts and form:* $\{\{a_1, a_3\}, \{a_2, a_4\}, \{a_6, a_7\}\{a_5\}\}$. a_5 *proposes to* a_4 *who rejects and then to* a_7 *who accepts forming* $\{\{a_1, a_3\}, \{a_2, a_4\}, \{a_5, a_7\}\{a_6\}\}$. a_6 *proposes to* a_7 *who rejects and to* a_5 *who accepts forming* $\{\{a_1, a_3\}, \{a_2, a_4\}, \{a_6, a_7\}\{a_5\}\}$. *Then* a_5 *proposes to* a_6 *who rejects, then to* a_4 *who also rejects and then to* a_7 *who accepts forming* $\{\{a_1, a_3\}, \{a_2, a_4\}, \{a_5, a_6\}\{a_7\}\}$ *and we reach*

a cycle. The cycling agents get together $\{a_5, a_6, a_7\}$ *and stable partition* $\{\{a_1, a_3\}, \{a_2, a_4\}, \{a_5, a_6, a_7\}\}$ *for* $(N_7, (\succeq)^7) = (N, (\succeq))$ *is obtained for (see Case 2 (iii))*

The outcome is stable by construction: there is not a pair of agents belonging to different sets or within a set who block the stable partition.